



TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ



**LOJİSTİK REGRESYON ANALİZİ İLE ELDE EDİLEN BETA
KATSAYISINA, ODDS ORANINA VE MAKİNE ÖĞRENME
ALGORİTMALARI İLE ELDE EDİLEN
AĞIRLIKLANDIRILMIŞ SKORLARA DAYALI KLİNİK
TAHMİN MODELLERİNİN BAŞARILARININ
KARŞILAŞTIRILMASI**

Gülçin AYDOĞDU

BİYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ

DANIŞMAN
Prof. Dr. Yasemin YAVUZ

ANKARA

2023

TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

LOJİSTİK REGRESYON ANALİZİ İLE ELDE EDİLEN BETA
KATSAYISINA, ODDS ORANINA VE MAKİNE ÖĞRENME
ALGORİTMALARI İLE ELDE EDİLEN
AĞIRLIKLANDIRILMIŞ SKORLARA DAYALI KLİNİK
TAHMİN MODELLERİNİN BAŞARILARININ
KARŞILAŞTIRILMASI

Gülçin AYDOĞDU

BİYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ

DANIŞMAN
Prof. Dr. Yasemin YAVUZ

ANKARA

2023

ETİK BEYAN

Ankara Üniversitesi

Sağlık Bilimleri Enstitüsü Müdürlüğü'ne,

Doktora tezi olarak hazırlayıp sunduğum “Lojistik Regresyon Analizi ile Elde Edilen Beta Katsayısına, Odds Oranına ve Makine Öğrenme Algoritmaları ile Elde Edilen Ağırlıklandırılmış Skorlara Dayalı Klinik Tahmin Modellerinin Başarılarının Karşılaştırılması” başlıklı tez; bilimsel ahlak ve değerlere uygun olarak tarafımdan yazılmıştır. Tezimin fikir/hipotezi tümüyle tez danışmanım ve bana aittir. Tezde yer alan deneysel çalışma/araştırma tarafımdan yapılmış olup, tüm cümleler, yorumlar bana aittir.

Yukarıda belirtilen hususların doğruluğunu beyan ederim.

Öğrencinin Adı Soyadı: Gülçin AYDOĞDU

Tarih: 22.12.2023

İmza:

KABUL VE ONAY

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü

Biyoistatistik Anabilim Dalında

Gülçin AYDOĞDU tarafından hazırlanan

“Lojistik Regresyon Analizi ile Elde Edilen Beta Katsayısına, Odds Oranına ve Makine Öğrenme Algoritmaları ile Elde Edilen Ağırlıklandırılmış Skorlara Dayalı Klinik Tahmin Modellerinin Başarılarının Karşılaştırılması” adlı tez çalışması aşağıdaki jüri tarafından DOKTORA TEZİ olarak OY BİRLİĞİ ile kabul edilmiştir.

Tez Savunma Tarihi: 22.12.2023

Prof. Dr. Atilla Halil ELHAN
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Jüri Başkanı

Prof. Dr. Pınar ÖZDEMİR
Hacettepe Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Üye

Prof. Dr. Yasemin YAVUZ
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Üye

Doç. Dr. Serdal Kenan KÖSE
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Üye

Doç. Dr. Emre DEMİR
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Üye

Tez hakkında alınan jüri kararı, Ankara Üniversitesi Sağlık Bilimleri Enstitüsü Yönetim Kurulu tarafından onaylanmıştır.

Prof. Dr. Fügen AKTAN
Sağlık Bilimleri Enstitüsü Müdürü

ÖZET

Klinik tahmin modelleri, yaş, cinsiyet ve biyobelirteçler gibi tahmin edici sayısının birden fazla olduğu durumlarda prognostik veya risk faktörlerini kullanarak bir birey için gelecekteki bir olayın veya hastalığın riskini tahmin etmek amacıyla yaygın olarak kullanılmaktadır. Literatürde klinik tahmin modellerinin oluşturulmasında yaygın olarak β katsayıları, odds oranları ve bunların tamsayıya yuvarlanmış skorları kullanılmaktadır. Bazı çalışmalarda makine öğrenme algoritmaları ile ağırlıklandırılmış skorlar da kullanılmaktadır. Ancak klinik tahmin modellerinde hangi skrolama sisteminin kullanılmasının daha iyi sonuç vereceği hakkında bir fikir birliği bulunmamaktadır.

Bu tez çalışmasında lojistik regresyondan elde edilen β katsayıları, odds oranları, bu değerlerin tamsayıya dönüştürülmesi ile elde edilen skorlar ve makine öğrenme algoritmaları (rastgele orman, destek vektör makineleri vb.) ile ağırlıklandırılarak elde edilen skorlar aracılığıyla farklı senaryolar kullanılarak oluşturulan tahmin modellerinin performanslarının değerlendirilmesi ve en iyi tahmin modelinin oluşturulması amaçlanmıştır. Çalışmamızda iki düzeyli sonuç değişkene sahip klinik tahmin modellerinde risk faktörlerinin skor değerlerini belirlemek için benzetim ile üretilen veri seti ve gerçek veri seti kullanılarak lojistik regresyon analizi gerçekleştirilmiştir. Çalışmada farklı senaryolar için benzetim yöntemi kullanılmıştır. Hangi senaryoda hangi klinik tahmin modelinin daha başarılı olduğu sonucuna ulaşılmıştır. Gerçek veri setinde pediatri alanında 522 çocuktan toplanan kapsamlı bir veri seti ile benzetim çalışmasında kullanılan tüm yöntemler ile malnütrisyon durumunu tahmin eden klinik tahmin modeli geliştirildi.

Model performanslarını karşılaştırmak için duyarlılık, seçicilik, PTD, NTD, doğruluk, F1 skorları, AUC ve MCC değerleri hesaplandı. Araştırma bulgularımız birlikte değerlendirildiğinde çalışmamızda öncelikle doğrudan makine öğrenme algoritmaları ile sınıflama yaparak model kurmak yerine iki düzeyli bir yanıtla sahip tanı konulacak klinik tahmin modellerinde lojistik regresyon modeli ile kurulabilecek tüm modelleri farklı senaryolar ile karşılaştırıldı. Ayrıca lojistik regresyona alternatif olarak klinik skrolama için risk faktörlerine modelde verilecek ağırlıkları veya skorları farklı makine öğrenme algoritmaları ile elde edildi ve tüm model sonuçları karşılaştırıldı.

Benzetim veri seti ile oluşturulan klinik tahmin modellerinde tüm senaryolar AUC, duyarlılık F1 skorları ve MCC değerleri ile değerlendirildiğinde ağırlıklı olarak en başarılı modellerin lojistik regresyon ile elde edilen beta, odds oranları, diğer algoritmalarından bayesci GDM ve L1(lasso) olduğu aynı zamanda bunları tamsayıya çevrildiğinde çok büyük farklılıklar yaratmadığı için kullanım kolaylığından dolayı tam sayılı beta ve tam sayılı odds son model olarak belirlendi ve makine öğrenme algoritmalarından rastgele orman algoritmalarının sınıflama başarısının en küçük olduğu belirlendi. Malnütrisyon durumunu tahmin eden klinik tahmin modelinde En iyi başarı lojistik regresyon analizi ile beta katsayıları kullanılarak elde edildi. Ancak klinik tahmin modellerinin kullanılabilirlik özelliği düşünülerek ikincil olarak en başarılı model tam sayılı odds oranlarına dayalı olarak geliştirilen klinik tahmin modeli oldu.

Bu alıřmada yeni bir yntem olarak klinik tahmin modellerinde makine ğrenme algoritmaları ile modeli oluřturmak yerine risk faktrleri iin verilecek skorların makine ğrenme algoritmaları ile verilebileceğini neriyoruz. Sonu olarak, klinik tahmin modellerinin etkili olması, yalnızca iyi bir performans sergilemekle deėil, aynı zamanda gerek dnya kořullarında uygulanabilir, anlaşılabilir, doėru ve genellemeye aık olma zellikleriyle de iliřkilendirilir.

Anahtar Szckler: Klinik tahmin modeli, Lojistik regresyon, Odds oranı, Tam sayılı skrlama sistem, Makine ğrenme algoritmaları



SUMMARY

Clinical prediction models are widely used to predict the risk of future events or diseases for an individual by utilizing predictive factors such as age, gender, and biomarkers when there are multiple predictive variables. In the literature, the creation of clinical prediction models often involves the use of β coefficients, odds ratios, and their integer-rounded scores derived from logistic regression. Some studies also employ machine learning algorithms to generate weighted scores.

However, there is no consensus on which scoring system is more effective in clinical prediction models. This thesis aims to evaluate the performance of prediction models created using logistic regression-derived β coefficients, odds ratios, scores obtained by rounding these values, and scores weighted using machine learning algorithms (random forests, support vector machines, etc.) under different scenarios. The goal is to identify the best prediction model.

In this study, logistic regression analysis was conducted using simulated and real datasets to determine the score values of risk factors in clinical prediction models with binary outcome variables. Simulation methods were employed for different scenarios, leading to conclusions about which clinical prediction model was more successful in each scenario. A comprehensive dataset from 522 children in the field of pediatrics was used for the logistic regression analysis, and a clinical prediction model predicting malnutrition status was developed.

Various performance metrics, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, F1 score, MCC and AUC values, were calculated to compare model performances. Our research primarily compared logistic regression models for diagnosing clinical prediction models with binary outcomes under different scenarios rather than directly classifying with machine learning algorithms. Additionally, we obtained weights or scores for risk factors in logistic regression as an alternative to logistic regression, using different machine learning algorithms, and compared all model results.

When evaluating AUC, accuracy, F1 score and MCC for clinical prediction models created with the simulated dataset, the most successful models were consistently Bayesian GLM, L1 (lasso) based on scores obtained from logistic regression-derived beta, integer beta, odds, integer odds, and machine learning algorithms. The logistic regression analysis using beta coefficients achieved the highest success in predicting malnutrition status in the clinical prediction model. However, considering the usability feature of clinical prediction models, the second most successful model was the one developed based on integer odds ratios.

As a novel approach, we propose in this study that machine learning algorithms can be used to determine scores for risk factors in clinical prediction models, instead of creating the entire model using machine learning algorithms. In conclusion, the effectiveness of clinical prediction models is associated not only with good performance but also with being applicable, understandable, accurate, and generalizable under real-world conditions.

Keywords: Clinical prediction model, Logistic regression, Odds ratio, Integer scoring system, Machine learning algorithms

İÇİNDEKİLER

Etik Beyan	ii
Kabul ve Onay	iii
Özet	iv
Summary	vi
İçindekiler	vii
Önsöz	ix
Simgeler ve Kısaltmalar	x
Şekiller	xi
Çizelgeler	xiii

1. GİRİŞ	1
1.1. Klinik tahmin modeli	6
1.1.1. Model oluşturma adımları	8
1.1.1.1. Klinik tahmin modelleri oluşturmaya hazırlık (Amaç)	8
1.1.1.2. Veri Kümesi Seçimi	9
1.1.1.2.1. Çalışma tasarımı	9
1.1.1.2.2. Sonuç değişkenin (Outcome) tanımları	10
1.1.2. Model geliştirme	10
1.1.2.1. Tahmin edicilerin tanımlanması	10
1.1.2.2. Değişken türleri (sürekli ve kategorik) ve tahmin edicilerin etkileşimleri	12
1.1.3. Model seçimi	13
1.1.4. Model performansı	13
1.1.5. Model geçerliliği	17
1.1.5.1. Dahili doğrulama	17
1.1.5.2. Harici doğrulama	20
1.1.6. Klinik tahmin modelinin sunumu	21
1.2. Lojistik Regresyon	22
1.2.1. Lojistik regresyon ile oluşturulan klinik tahmin modellerinin sunumları	24
1.2.1.1. Tahmin modelinin lojistik regresyon formülü ile sunumu	24
1.2.1.2. Tahmin modelinin web siteleri ve uygulamalar ile sunumu	25
1.2.1.3. Tahmin modelinin nomogramlar ile sunumu	27
1.2.1.4. Tahmin modelinin skarlama sistemi ile sunumu	29
1.3. Makine Öğrenme	36
1.3.1. Makine Öğrenme Algoritmaları	38
1.3.1.1. Karar Ağaçları (Decision Tree)	38
1.3.1.2. Rastgele Orman (Random Forest)	39
1.3.1.3. Genelleştirilmiş Boosted Regresyon Modelleri (GBM)	41
1.3.1.4. Ölçeklenebilir Gradient Boosting (Extreme Gradient Boosting Models-XGBoost)	42
1.3.1.5. Hafif Gradyan Arttırma Makinesi (Light Gradient Boosting Machine-LightGBM)	43
1.3.1.6. Destek Vektör Makineleri-DVM (Support Vector Machines - SVM)	44
1.4. Diğer Algoritmalar	45

1.4.1. Bayesci Genelleştirilmiş Doğrusal Model (Bayesian Generalized Linear Model)	45
1.4.2. Bilgi Kazancı (Information Gain)	47
1.4.3. L1 Düzenlemeli Modeller (L1 Regularization Models- Lasso)	48
2. GEREÇ ve YÖNTEM	49
2.1. Benzetim verisi	49
2.2. Gerçek Veri Seti Üzerinde Uygulama	54
2.2.1. Klinik tahmin modeli oluşturma ve skorlama sistemi	56
2.2.2. Makine öğrenme yöntemi	57
2.2.3. Klinik tahmin modeli	57
3. BULGULAR	58
3.1. LR, ML ve diğer algoritmalar ile elde edilen skorlar için benzetim sonuçları	58
3.1.1. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve %20-%80 (n=200, n=500 ve n=1000)	58
3.1.2. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve %50-%50 (n=200, n=500 ve n=1000)	64
3.1.3. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve rastgele (n=200, n=500 ve n=1000)	71
3.1.4. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve %20-%80 (n=200, n=500 ve n=1000)	76
3.1.5. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve %50-%50 (n=200, n=500 ve n=1000)	82
3.1.6. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve rastgele (n=200, n=500 ve n=1000)	88
3.1.7. Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve %20-%80 (n=200, n=500 ve n=1000)	94
3.1.8. Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve %50-%50 (n=200, n=500 ve n=1000)	100
3.1.9. Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve rastgele (n=200, n=500 ve n=1000)	106
3.1.10. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve %20-%80 (n=200, n=500 ve n=1000)	112
3.1.11. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve %50-%50 (n=200, n=500 ve n=1000)	118
3.1.12. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve rastgele (n=200, n=500 ve n=1000)	124
3.2. Gerçek veri setine ait sonuçlar	130
3.2.1. Lojistik regresyondan elde edilen skorlar ile klinik tahmin modeli	138
4. TARTIŞMA	143
5. SONUÇ VE ÖNERİLER	147
KAYNAKLAR	150
EKLER	158
EK-1. Lojistik regresyona uygun benzetim verisi türetimi için örnek R kodları	158
EK 2. Etik Kurur Onayı	161
ÖZGEÇMİŞ	164

ÖNSÖZ

Bu tez çalışmasında lojistik regresyondan elde edilen β katsayıları, odds oranları, bu değerlerin tamsayıya dönüştürülmesi ile elde edilen skorlar ve makine öğrenme algoritmaları (rastgele ormanlar, destek vektör makineleri vb.) ile ağırlıklandırılarak elde edilen skorlar aracılığıyla farklı senaryolar kullanılarak oluşturulan tahmin modellerinin performanslarının değerlendirilmesi ve benzetim yöntemi ile belirlenen en iyi tahmin modeli kullanılarak Pediatrik Malnütrisyon Klinik Tahmin Skorunun geliştirilmesi amaçlanmıştır.

Doktora eğitimim ve tez çalışması süresince bilgi ve tecrübeleri ile beni yönlendiren, ilgisini, bilgisini ve desteğini esirgemeyen değerli danışman hocam Prof. Dr. Yasemin YAVUZ'a,

Tez çalışması süresince bilgi ve tecrübeleri ile bana yol gösteren ve yardımlarını esirgemeyen sayın hocam Hitit Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı Öğretim Üyesi Doç.Dr. Emre DEMİR'e, bu süreçte bana destek olan Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı Öğretim Üyelerinden Doç.Dr. Serdal Kenan KÖSE ve Prof.Dr. Atilla Halil ELHAN'a, doktora tez çalışmamın uygulama bölümünde kullandığım verileri toplama süresince her türlü kolaylığı sağlayan ve destek olan sayın hocalarım Hitit Üniversitesi Tıp Fakültesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı Öğretim Üyelerinden Doç.Dr. Atakan COMBA, Dr. Öğr. Üyesi İlknur KABA ve Hitit Üniversitesi Çorum Erol Olçok Eğitim ve Araştırma Hastanesi Çocuk Cerrahisi Dr. Nurcan ÇOŞKUN'a, ayrıca Mühendislik Fakültesi Bilgisayar Bilimleri Anabilim Dalı Öğretim Üyesi Dr. Öğr. Üyesi Ömer Faruk AKMEŞE'ye,

Çalışmalarım boyunca her zaman beni destekleyen, çok değerli aileme sonsuz teşekkürlerimi sunarım.

SİMGELER VE KISALTMALAR

AIC	Akaiki Bilgi Kriteri
AUC	Area Under the Curve
CARET	Classification and Regression Tree
CCS	Charlson Komorbidite Skoru
GA	Güven Aralığı
GBM	Genelleştirilmiş Boosted Model
GDM	Genelleştirilmiş Doğrusal Model
GLMNET	Generalized Linear Model with Penalized Maximum Likelihood –
HR	Hazard Ratio
LİGHTGBM	Light Gradient Boosting Machine
LODS	Lojistik Organ Disfonksiyon sistemi
LOESS	Locally Weighted Smoothing
LR	Lojistik Regresyon
ML	Machine Learning
NIHSS	Pediyatrik NIH İnme Ölçeği
NTD	Negatif Tahmin Değeri
OR	Odds Oranı
PAS	Pediyatrik Apandisit Skoru
PAS	Pediyatrik Astım Skoru
pGCS	Pediyatrik Glasgow Koma Skalası
PIAMA	Astım ve Akar Alerjisinin Önlenmesi ve İnsidansı
PTD	Pozitif Tahmin Değeri
RKÇ	Randomize Kontrollü Çalışma
ROC	Receiver Operating Characteristic
RPART	Recursive Partitioning and Regression Trees
SAPSII	Basitleştirilmiş Akut Fizyoloji Skoru
SOFA	Sıralı Organ Yetmezliği Değerlendirme Skoru
SVM	Support Vector Machine
TREAT	Torasik Araştırma Değerlendirme ve Tedavi Modeli
XGBOOST	Extreme Gradient Boosting Methods
α	Alfa
β	Beta

ŞEKİLLER

Şekil 1.1. Bölünmüş örneklem	18
Şekil 1.2. Çapraz doğrulama	19
Şekil 1.3. Bootstrapping	20
Şekil 1.4. Alvarado skoruna ilişkin web uygulama örneği	27
Şekil 1.5. Ceza uygulamalı lojistik regresyon (penalized lojistik regresyon) modelinde altı tahminciye dayalı benign histoloji için nomogram	29
Şekil 1.6. Yapay zeka, makine öğrenmesi ve derin öğrenme	37
Şekil 1.7. Şiddetli COVID-19 hastalarında ölüm sonucunu tahmin etmek için klinik özellikler ve karar ağacı modeli	39
Şekil 1.8. Rastgele orman şeması	41
Şekil 1.9. Destek vektör makinelerine ilişkin grafikler	45
Şekil 2.1. Benzetim verisindeki bağımlı ve bağımsız değişkenler	50
Şekil 2.2. A senaryosuna ilişkin örnek sonuçlar	53
Şekil 2.3. B senaryosuna ilişkin örnek sonuçlar	53
Şekil 3.1. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	60
Şekil 3.2. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	62
Şekil 3.3. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	64
Şekil 3.4. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	66
Şekil 3.5. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	68
Şekil 3.6. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	70
Şekil 3.7. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	72
Şekil 3.8. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	74
Şekil 3.9. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	76
Şekil 3.10. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	78
Şekil 3.11. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	80
Şekil 3.12. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	82
Şekil 3.13. LR ve ML algoritmaları ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	84
Şekil 3.14. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	86
Şekil 3.15. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	88

Şekil 3.16. LR ve ML algoritmaları ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	90
Şekil 3.17. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	92
Şekil 3.18. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	94
Şekil 3.19. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	96
Şekil 3.20. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	98
Şekil 3.21. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	100
Şekil 3.22. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	102
Şekil 3.23. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	104
Şekil 3.24. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	106
Şekil 3.25. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	108
Şekil 3.26. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	110
Şekil 3.27. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	112
Şekil 3.28. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	114
Şekil 3.29. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	116
Şekil 3.30. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	118
Şekil 3.31. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	120
Şekil 3.32. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	122
Şekil 3.33. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	124
Şekil 3.34. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	126
Şekil 3.35. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	128
Şekil 3.36. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği	130

ÇİZELGELER

Çizelge 1.1. Karışıklık Matrisi (Confusion Matrix)	14
Çizelge 1.2. Çoklu lojistik regresyon sonuçları	34
Çizelge 1.3. Puanları hesaplama adımları	35
Çizelge 2.1. ML algoritmaları ve diğer algoritmalar	49
Çizelge 3.1. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	59
Çizelge 3.2. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	61
Çizelge 3.3. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	63
Çizelge 3.4. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	65
Çizelge 3.5. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	67
Çizelge 3.6. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	69
Çizelge 3.7. Senaryo A; n=200; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	71
Çizelge 3.8. Senaryo A; n=500; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	73
Çizelge 3.9. Senaryo A; n=1000; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	75
Çizelge 3.10. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	77
Çizelge 3.11. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	79
Çizelge 3.12. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	81
Çizelge 3.13. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	83

Çizelge 3.14. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	85
Çizelge 3.15. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	87
Çizelge 3.16. Senaryo A; n=200; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	89
Çizelge 3.17. Senaryo A; n=500; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	91
Çizelge 3.18. Senaryo A; n=1000; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	93
Çizelge 3.19. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	95
Çizelge 3.20. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	97
Çizelge 3.21. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	99
Çizelge 3.22. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	101
Çizelge 3.23. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	103
Çizelge 3.24. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	105
Çizelge 3.25. Senaryo B; n=200; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	107
Çizelge 3.26. Senaryo B; n=500; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	109
Çizelge 3.27. Senaryo B; n=1000; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	111
Çizelge 3.28. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	113
Çizelge 3.29. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	115

Çizelge 3.30. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	117
Çizelge 3.31. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	119
Çizelge 3.32. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	121
Çizelge 3.33. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	123
Çizelge 3.34. Senaryo B; n=200; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	125
Çizelge 3.35. Senaryo B; n=500; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	127
Çizelge 3.36. Senaryo B; n=1000; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması	129
Çizelge 3.37. Çocuğa ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması	132
Çizelge 3.38. Anne ve babaya ilişkin sosyo-demografik özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması	134
Çizelge 3.39. Çocuğun beslenme alışkanlığına ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması	135
Çizelge 3.40. Çocuğun beslenme davranışları ve diğer etkili faktörlerin malnütrisyon olan ve olmayanlarda karşılaştırılması	137
Çizelge 3.41. Malnütrisyon öngörüsünde etkili olan risk faktörlerini belirlemek amacıyla gerçekleştirilen çoklu lojistik regresyon analiz sonuçları	139
Çizelge 3.42. Beta, tamsayıli Beta, OR, tamsayıli OR	139
Çizelge 3.45. Skorun malnütrisyon tahminindeki AUC, duyarlılık, seçicilik, PTD, NTD, doğruluk, F1 skoru ve MCC	141
Klinik tahmin modeli tamsayıli beta kullanılarak oluşturulmuştur.	141
Çizelge 3.46. ML ve diğer algoritmalar ile oluşturulan model performanslarının karşılaştırılması	142

1. GİRİŞ

Tahmin modeli, verilerden elde edilen bilgileri kullanarak gelecekteki olayları veya durumları tahmin etmeye yardımcı olan istatistiksel bir modeldir. Bu modeller, belirli değişkenler arasındaki ilişkileri analiz ederek gelecekteki olayların olasılıklarını belirlemeye çalışır. Tahmin modelleri tıp alanında hastalıkların tanı ve prognozuna ilişkin tahminler yapmak için kullanılmaktadır. Tahmin modelleri, doğru ve güvenilir sonuçlar elde etmek amacıyla özenle tasarlanmalı ve uygun veri setleriyle etkili bir biçimde eğitilmelidir. Aynı zamanda modelin güncel verilere ve değişen koşullara uyum sağlaması için sürekli güncellenmesi ve iyileştirilmesi önemlidir (Steyerberg ve Vergouwe, 2014; Chowdhury ve Turin, 2020).

Tahmin modelleri tıbbi uygulama ve araştırma amaçları için değerlidir. Steyerberg, halk sağlığı, klinik uygulama ve tıbbi araştırma olmak üzere klinik tahmin modellerini üç ana uygulama dalı olarak tanımlamıştır. Halk sağlığında, tahmin modelleri hastalara ve hasta olma riski taşıyan kişilere önleyici müdahaleleri hedeflemekte yardımcı olabilir. Klinik uygulamada, tahmin modelleri hastalığın teşhisi veya prognostik sonuç olasılığı hakkında hastaları ve tedavi eden hekimleri bilgilendirebilir. Tıbbi araştırmalarda ise tahmin modelleri birkaç amaç için kullanılabilir. Randomize kontrollü çalışma (RKÇ) gibi deneysel çalışmalarda, öngörücü temel özellikleri hastaların dahil edilmesine ve sınıflandırılmasına yardımcı olabilir ve istatistiksel analizi geliştirebilir. Gözlemsel çalışmalarda, karıştırıcı faktörlerin yeterli düzeyde kontrol edilmesi esastır (Steyerberg ve Vergouwe, 2014).

Klinik tahmin modelleri genellikle sağlık alanında kullanılır ve çeşitli hastalıkların tanısı, hastalık seyrinin tahmini, tedavi yanıtının öngörülmesi gibi klinik karar süreçlerine destek sağlar (Steyerberg ve Vergouwe, 2014; Wolff vd., 2019; Dogancay, 2023).

Klinik tahmin modelleri, algoritmalar aracılığıyla çeşitli girdi verileri kullanarak sonuçları tahmin etmek için kullanılır. Bu girdi verileri, hastanın demografik bilgileri, laboratuvar sonuçları, görüntüleme bulguları, semptomlar, hastalık öyküsü gibi klinik özellikler olabilir. Model, bu verilere dayalı olarak belirli bir sonuç veya durumu tahmin etmek üzere eğitilir. Klinik tahmin modelleri, örneğin bir hastanın kanser tanısı veya prognozu hakkında bilgi sağlayabilir (Visweswaran vd., 2010; Alanazi, 2017; Norgeot vd., 2019).

Klinik tahmin modelleri, sađlık profesyonellerine hastalarıyla ilgili karar verme süreçlerinde destek olma imkanı sađlar. Ancak, sadece bu modellerin sonuçlarına güvenmekle kalmayıp klinik deđerlendirme ve uzman görüşleriyle birlikte kullanılması önemlidir.

Klinik modeli oluşturmada önce tahmin edilecek olay veya hastalığın ne olduđu, modelin hedef hastasının kim olduđu ve tahmin modelinin hedef kullanıcısının kim olduđu belirlenir. Daha sonra model için toplanan veriler hastaların klinik özelliklerini içermelidir (Örneğin, yaş, cinsiyet, tıbbi geçmiş, laboratuvar sonuçları, semptomlar ve diđer ilgili klinik deđerşkenler vb.). Bir klinik tahmin modelinin geliştirilmesinde iki farklı veri kümesi kullanılır. Bu veri kümelerinin biri geliştirme, diđerı doğrulama veri kümesidir (Lee vd., 2016). Verileri analiz ederek model için uygun olan deđerşkenleri seçmek önemlidir. Klinik öneme sahip ve hastalığın tahmininde etkili olan deđerşkenleri seçmek, modelin başarısını artırabilir (Chowdhury ve Turin 2020).

Araştırmacının amacına bađlı olarak, farklı kullanıcıları hedeflemek için farklı deđerşkenleri analiz eden farklı modeller geliştirilebilir. Laboratuvar deđerşkenleri gerektirmeyen basit bir klinik tahmin modeli ve gerektiren kapsamlı bir model, sıradan kişiler ve sađlık profesyonelleri için tasarlanabilir (Lee vd., 2016).

Tahmin modellerinde kategorik ve sürekli deđerşkenlerin farklı şekilde yönetilmesi gerekir. Deđerşkenleri seçmek ve tahmin modelini geliştirmek için literatürde ortak bir görüş bulunmamaktadır, fakat istatistiksel araçlarla çeşitli stratejiler mevcuttur. Modele ve kullanım amacına göre sıklıkla kullanılan yöntemler arasında lojistik regresyon, doğrusal regresyon, Cox modellerini içeren regresyon analizleri ve makine öğrenimi algoritmaları bulunur (Lee vd., 2016).

Klinik tahmin modellerinin doğru ve etkili bir şekilde sunumu çok önemlidir. Bonnett vd klinik tahmin modellerini sunmak için bir rehber geliştirmiştir. Bu kılavuzda puanlama sistemi (points score system), grafiksel puan tablosu, nomogram ve web siteleri ile uygulamalar olmak üzere dört ana sunum formatı bulunmaktadır (Bonnett vd., 2019).

Puana dayalı risk puanlama sistemleri, bilgisayar veya diđer elektronik cihazlar kullanılmadan hasta riskinin hızlı bir şekilde deđerlendirilmesine izin verdiđi için doktorlar

arasında yaygın olarak kullanılmaktadır. Bu tür puan tabanlı risk puanlama sistemlerinin kullanımı, kanıta dayalı klinik karar vermeyi kolaylaştırır (Bonnett vd., 2019).

Literatürde klinik tahmin modellerinin geliştirilmesinde farklı skorlama sistemlerinin kullanıldığı birçok çalışma bulunmaktadır. Tahmin modellerinde farklı skorlama sistemlerinin kullanıldığı yayın sayısı her geçen gün artmaktadır. Yoğun bakım ünitesinde ölüm riskini değerlendirmek amacıyla kullanılan APACHE skorları, akut apandisit tanısında kullanılan Alvarado skoru, yoğun bakım ünitesindeki bakım kalitesini ölçmek için kullanılan Basitleştirilmiş Akut Fizyoloji Skoru (SAPS II), pulmoner emboli için Wells Kriterleri, bireylerin 10 yıl içinde koroner kalp hastalığı riskini öngörmek amacıyla kullanılan Framingham koroner kalp hastalığı risk skoru ve hastalarda bilinç (nörolojik) durumunun değerlendirilmesinde kullanılan Glasgow Koma Skoru, örnek olarak verilebilir (Lee vd., 2016; Shipe vd., 2019; Janssen vd., 2009; Bonnett vd., 2019). Ayrıca Pediatri araştırma alanında da Pediatrik Astım Skoru (PAS), Pediatrik Glasgow Koma Skalası (pGCS), Pediatrik Apandisit Skoru (PAS), Pediatrik NIH İnme Ölçeği (NIHSS) ve Astım ve Akar Alerjisinin Önlenmesi ve İnsidansı (PIAMA) Risk Skoru gibi birçok hastalık için tanı skorları geliştirilmiştir (Kelly vd, 2000; Borgialli vd. 2016; Samuel, 2002; Ichord vd., 2011; Brunekreef vd., 2002).

Klinik tahmin modelleri literatürde genellikle β katsayıları, odds oranları ve bunların tam sayıya yuvarlanmış skorları kullanılarak oluşturulmaktadır (Knaus vd., 1985; Alvarado, 1986; Le Gall vd., 1993; Bang vd., 2009; Mehta vd., 2016; Zhang vd., 2017; Subramanian vd., 2021). Bazı çalışmalarda makine öğrenme algoritmaları ile ağırlıklandırılmış skorlar da kullanılmaktadır (Demir, 2019; Jacobs vd., 2020, Lynch vd., 2020; Zhang vd., 2019; Jelovsek vd., 2018; Cerqueira vd., 2014; Segar vd., 2019; Lee ark., 2019, Pimentel vd., 2021). Ancak klinik tahmin modellerinde hangi skorlama sisteminin kullanılmasının daha iyi sonuç vereceği hakkında bir fikir birliği bulunmamaktadır.

Tez çalışmamız lojistik regresyondan elde edilen β katsayıları, odds oranları, bu değerlerin tam sayıya dönüştürülmesi ile elde edilen skorlar ve makine öğrenme algoritmaları (rastgele ormanlar, destek vektör makineleri vb.) ile ağırlıklandırılarak elde edilen skorlar aracılığıyla farklı senaryolar kullanılarak oluşturulan tahmin modellerinin performanslarının değerlendirilmesini kapsamaktadır. Ayrıca benzetim yöntemi ile belirlenen en iyi tahmin modeli kullanılarak Pediatrik Malnütrisyon Klinik Tahmin Skorunun geliştirilmesi amaçlanmıştır.

Tezin amacı doğrultusunda; konu kapsamında gerçekleştirilen kapsamlı literatür taraması ile skorlara ilişkin çeşitli çalışmalar incelenmiştir.

Knaus et al. (1985) yaptıkları çalışmada APACHE II skorlarını yoğun bakım ünitesinde ölüm riskini değerlendirmek için kullanmışlardır. Bu çalışmada 12 fizyolojik değişken, yaş ve önceki sağlık durumu değerlerine göre bir skor puanı (point score) kullanmışlardır. 12 fizyolojik değişkene klinisyenin klinik deneyimine ve literatürün gözden geçirilmesine bağlı olarak 0-4 arasında ağırlıklar atanmıştır (Knaus vd., 1985).

Alvarado (1986) çalışmasında Akut apandisit düşündüren karın ağrısı ile hastaneye yatırılan hastaları incelemiştir. Alvarado skoru toplam 8 değişkenden oluşmaktadır. Skorlamada daha önemli olan belirtilere (sağ alt kadranda hassasiyeti ve lökositoz) 2 değeri, kalan belirtilere 1 değeri atanmış toplam 10 tam puana ulaşılmıştır. Önemli belirtileri belirlemek için duyarlılık, seçicilik değerleri, pozitif ve negatif tahmin değerleri hesaplanmış ve daha sonra ortak olasılıklar hesaplanarak tanısal ağırlıklar elde edilmiştir. Her bir belirtinin tanısal ağırlığına küçük bir sayı atanmış ve pratikte kullanılabilecek uygulanabilir bir puan elde edilmiştir (Alvarado, 1986).

Le Gall vd. (1993) yaptıkları çalışmada SAPS II skorunu kullanmışlardır. Yoğun bakıma alınan hastaların mortalite riskini değerlendirmek için kullanılan bir skorlama sistemidir. Bu skor, hastanın yaşam fonksiyonları, organ sistemleri ve laboratuvar bulguları temel alınarak oluşturulmuştur. SAPS II, hastaların genel durumunu değerlendirir ve klinik pratikte kullanılan bir prognostik araçtır. SAPS II skoru 12 fizyolojik faktör, yaş, kabul türü (planlı ve planlanmamış cerrahi) ve hastalıkla ilişkilendirilen üç değişken (kazanılmış immün yetmezlik sendromu, metastatik kanser ve hematolojik kanser) olmak üzere toplam 17 değişkenden meydana gelmiştir. SAPS II'yi oluşturacak değişkenlerin seçilmesinde ve her bir değişken için uygun gruplamalara ve puan atamalarına karar verilmesine yardımcı olmak için lojistik regresyon analizi kullanılmıştır. (Le Gall vd., 1993).

Mehta vd. (2016) yaptıkları çalışmada mortaliteyi tahmin etmede yanlış ağırlıklandırmanın performans üzerindeki etkilerini göstermek için regresyon katsayısına ve risk oranına dayalı puanlama sistemlerini kullanarak Charlson komorbidite skorunun (CCS) türetilmiş farklı versiyonlarını göstermiştir. Regresyon katsayısı tabanlı modellerin (Beta, Beta10/tamsayı, Beta/Schneeweiss, Beta/Sullivan), risk oranı tabanlı modellere (HR/Charlson, HR/Johnson) kıyasla daha düşük AIC ve daha yüksek R^2 ye sahip olduğunu

ve bundan dolayı regresyon katsayıları kullanılarak elde edilen CCS, risk oranına dayalı puanlama sistemlerine kıyasla daha başarılı olduğunu göstermişlerdir (Mehta vd., 2016).

Yapılan bir çalışma R yazılımı kullanılarak simülasyon ile elde edilen veri kümesine dayalı risk puanlarının nasıl geliştirileceği ve doğrulanacağı konusunda bilgi sağlamaktadır. Çalışmada sayısal ve kategorik değişkenleri içeren oluşturulan veri seti ve ilk olarak sayısal değişkenler, LOESS smoother tarafından belirlenen kesme noktalarına göre faktör değişkenlerine dönüştürülür. Daha sonra, lojistik regresyondaki katsayılarla ilişkili olan her bir değişkenin risk puanları, dönüştürülen faktör değişkenlerinin her bir düzeyine ve diğer kategorik değişkenlere atanır. Sonuç değişkeni olasılığının tahminini temsil etmek için her birey için toplam puan hesaplanmıştır (Zhang vd., 2017).

Shipe vd. (2019) yaptıkları çalışmada değişkenleri sınıflandırmak, seçmek, model performansını değerlendirmek, dahili ve harici doğrulama yapmak, modeli yeniden kalibre etmek ve modelin klinik etkisini değerlendirmek için lojistik regresyon kullanarak tahmin modeli geliştirmeye genel bir bakış sağlamayı amaçlamışlardır (Shipe vd., 2019).

Subramanian vd. (2021) yaptıkları çalışmalarda genellikle basitlik adına istatistiksel regresyon modelleri üzerine kurulu tahmin modellerini tamsayı puanlarına ve risk sınıflandırma sistemlerine dönüştürürler. Ancak bu sistem yararlı bilgileri atar ve tahmin doğruluğunu azaltır. Bu nedenle, simülasyon çalışması ve örnek klinik veri seti üzerinden bir regresyon modeli bir tamsayı puanına dönüştürüldüğünde tahmin doğruluğu üzerindeki etkisini araştırmışlardır (Subramanian vd., 2021).

Ayrıca Pediatri alanında birçok hastalık için geliştirilen tanı skorlarına ilişkin çeşitli çalışmalar incelenmiştir.

Kelly vd. (2000) yaptıkları çalışmada yatarak tedavi gören astım klinik yolunun astımlı çocukların bakım giderleri ve kalitesi üzerindeki etkisini incelemek amacıyla Pediatrik Astım Skorunu (PAS) kullanmışlardır (Kelly vd., 2000).

Yapılan bir çalışmada konuşma öncesi çocuklarda pediatrik Glasgow Koma Ölçeği (GCS) skorunun doğruluğunu, künt kafa travmasından sonra travmatik beyin yaralanması (TBH) olanları belirlemek için daha büyük çocuklarda standart GCS skoruyla karşılaştırmışlardır (Borgialli vd., 2016).

Samuel (2002) tarafından karın ağrısı olan 4 ila 15 yaşları arasındaki 1.170 hastanın prospektif bir kohort çalışmasında Pediatrik Apandisit Skoru geliştirilmiştir. Pediatrik Apandisit Skoru (PAS) çalışmasında, Akut apandisit için istatistiksel anlamlılık gösteren 8 değişken bulmak için klinik geçmiş ve laboratuvar sonuçlarından gelen veriler analiz edilmiş ve 10 puanlık puanlama sistemi oluşturulmuştur (Samuel, 2002).

Çok merkezli bir prospektif kohort çalışmasında 2-18 yaşları arasındaki 113 hasta arasında Pediatrik NIH İnme Ölçeği (NIHSS) için tahmin ediciler arası güvenilirliği değerlendirilmiştir. Genel olarak, değerlendiriciler arası güvenilirliği, değerlendiren çocuk nörologları arasında (kullanılan spesifik test maddesine bağlı olarak) 0.63 ile 1.00 arasında oldukça yüksek bulunmuştur (Ichord vd., 2011).

Brunekreef vd. (2002) yaptıkları çalışmanın amacı Astım ve Akar Alerjisinin Önlenmesi ve İnsidansı (PIAMA) Risk Skoru ile Akar-alerjiden kaçınmanın çocukluk çağında astım insidansı üzerindeki etkisini araştırmaktır (Müdahale Çalışması) (Brunekreef vd., 2002).

1.1. Klinik tahmin modeli

Klinik tahmin modellerine verilen bazı diğer isimler: Klinik skorlar; hastaların klinik özelliklerine dayanarak hastalık riskini veya klinik sonuçları tahmin eden puanlama sistemleridir. Klinik skorlama sistemleri; belirli bir hastalık veya hastanın sağlık durumu hakkında bilgi veren puan tablosu veya skorlama sistemleridir. Risk skorları; hastaların bir olayın (örneğin, hastalık teşhisi, komplikasyon, mortalite) gerçekleşme olasılığını tahmin etmek için kullanılan skorlama sistemleridir. Prognostik modeller; bir hastalığın seyrini veya belirli bir sonucu (örneğin, tedaviye yanıt, sağkalım) tahmin etmek için kullanılan modellerdir. Tahmin tabloları; nomogram veya puan tablosu gibi grafiksel araçlar kullanarak hastanın sonucunu tahmin etmek için kullanılan tablolarıdır. Klinik indeksler; belirli bir hastalığın varlığını veya hastanın klinik durumunu değerlendirmek için kullanılan klinik ölçeklerdir. Klinik öngörü modelleri; hastaların sağlık durumları veya belirli olaylar hakkında tahminler yapmak için kullanılan matematiksel modellerdir (Morita vd., 1999; Moons vd., 2015; Han, 2016; Bonnett vd., 2019; Cowley vd., 2019; Gu ve Liu, 2020). Bu terimler, klinik tahmin modellerini ifade etmek için kullanılan genel isimlerdir ve farklı araştırma alanlarında, sağlık kuruluşlarında veya literatürde farklı şekillerde kullanılabilir.

Klinik tahmin modelleri, yaş, cinsiyet ve biyobelirteçler gibi birden fazla tahmin edici değişkenin bulunduğu durumlarda, bireyin gelecekteki bir olayın veya hastalığın riskini tahmin etmek amacıyla prognostik veya risk faktörlerini kullanarak yaygın bir şekilde kullanılmaktadır (Lee vd., 2016; Shipe vd., 2019; Janssen vd., 2009; Bonnett vd., 2019). Doğru bir klinik tahmin modelinin amacı, hastanın risk düzeyini belirleyerek kişiye özel klinik karar alma sürecini desteklemek ve hastanın sonuçlarını ve sağlık hizmetlerinin kalitesini artırmak için kullanılır (Shipe vd., 2019). Klinik deneyimin yerine geçmemekle birlikte, bir bireyin hastalık riski hakkında nesnel veriler sağlayabilir ve klinik karar vermede gözlemlenen bazı yaygın yanlışlıkları önleyebilirler (Shipe vd., 2019). Klinik tahmin modelleri oluşturmak karmaşık ve dikkat gerektiren bir süreçtir. İyi bir model, doğru, güvenilir tahminler yapabilmeli ve klinik uygulamada fayda sağlamalıdır. Bu nedenle, verileri dikkatli bir şekilde analiz etmek, uygun modeli seçmek ve performansını doğrulamak için dikkatli bir çalışma yapmak önemlidir.

Asemptomatik bireylerde gelecekte hastalık oluşumunu tahmin etmek için çeşitli modeller geliştirilmiştir. Bu modellerin iyi bilinen örnekleri arasında kardiyovasküler hastalıklar için Framingham risk fonksiyonları yer almaktadır. Framingham risk fonksiyonları, önleyici müdahaleler için mevcut politikaların temelini oluşturur. Örneğin, kolesterol düşürücü tedaviler sadece kardiyovasküler hastalık riski nispeten yüksek olanlar için düşünülmektedir. Benzer şekilde, meme kanseri için tahmin modelleri geliştirilmiştir, bu modeller sayesinde yüksek risk altındaki kişiler için daha yoğun tarama veya kemoprofilaksi düşünülebilir (Steyerberg, 2009).

Klinik tahmin modelleri genellikle lojistik regresyon ve makine öğrenme algoritmaları gibi istatistiksel ve matematiksel yöntemler kullanılarak geliştirilir. Bu algoritmalarından bazıları: Lojistik Regresyon, Destek Vektör Makineleri ve Karar Ağaçları'dır. Bu algoritmalarından biri veya birkaçı kullanılarak, klinik verilerin analizi ve model eğitimi gerçekleştirilir. Bu sayede klinik tahmin modelleri, hastaların tanı, prognoz veya tedavi yanıtı gibi durumlarını tahmin etmek için kullanılabilir (Mythili vd., 2013; Chao vd., 2014; Collins vd., 2014; Christodoulou vd., 2019).

Klinik tahmin modelleri, tanısal veya prognostik bir sonucu tahmini için birçok özelliği (örn. hasta, hastalık veya tedavi ile ilgili) birleştirir (Collins vd., 2014). Klinik tahmin modellerinde, kaç tahmin edici değişkenin kullanılması gerektiği, modelin karmaşıklığı ve performansı bakımından çok önemlidir. Bir tahmin modelinde kullanılacak

tahmin edici deęiřken sayısı, modele etkisi olan önemli deęiřkenlerin belirlenmesi ve gereksiz karmařıklıktan kaçınılması dikkate alınarak belirlenmelidir (Chowdhury ve Turin 2020). Çoęu zaman çeřitli faktörlere baęlı olduęundan, bir tahmin modeline dahil edilecek deęiřkenlerin sayısına iliřkin belirlenmiř bir kural yoktur. Bir veri setinden kaç deęiřkenin tahmin edilebileceęini öngören bir kural olan 'onda bir kuralı', geleneksel klinik tahmin modelleme stratejisinde (örneğin, lojistik regresyon ve hayatta kalma modelleri) oldukça popülerdir. Bu kurala göre baęımlı deęiřkenin kategorilerinden sayısı küçük olanı 10'a bölerek çıkan sayı kadar deęiřken modele eklenebilir (Lee vd., 2016; Chowdhury ve Turin, 2020). Önceden anlamlı bulunan tahmin ediciler normalde aday deęiřkenler olarak düşünölmelidir (örneğin, diyabet risk skorunda ailede diyabet öyküsü). Tüm önemli tahmin edicilerin nihai modele dahil edilmesi gerekmedięine dikkat edilmelidir (örneğin, $P < 0,05$). İstatistiksel olarak anlamlı olmayan veya etki büyüklüęü küçük olan deęiřkenler de modele katkıda bulunabilir. Arařtırmacının takdirine baęlı olarak, farklı kullanıcıları hedeflemek için farklı deęiřkenleri analiz eden farklı modeller geliřtirilebilir (Lee vd., 2016). Risk tahmin modelleri, belirli bir sonuç veya belirlenmiř bir süre içinde ortaya çıkma olasılıęını tahmin etmek için hasta özelliklerini kullanır. Örneęin, TREAT modelini (Torasik Arařtırma Deęerlendirme ve Tedavi modeli), deęerlendiren cerrahlar mevcut olan bilgileri kullanarak bir akcięer nodölünün kanser olma riskini tahmin edebilir. Bu örnekte lojistik regresyon yöntemi kullanılarak klinik tahmin modeli oluřturulmuřtur. Modelde sonuç deęiřkeni belirsiz pulmoner nodüllerde akcięer kanseri olma durumudur. Bu modelde göęüs cerrahisi kliniklerine bařvuran belirsiz pulmoner nodülleri olan hastalar çalıřmaya dahil edilmiřtir (Shipe vd., 2019).

1.1.1. Model oluřturma adımları

Regresyon analizi ile klinik tahmin modelleri yedi adım da geliřtirilir.

1.1.1.1. Klinik tahmin modelleri oluřurmaya hazırlık (Amaç)

Klinik tahmin modelleri oluřurmaya hazırlık ařaması, bařarılı bir modelin oluřması için ilk adımdır. Bir tahmin modeli oluřturmanın en kritik adımı, modelin genel amacını belirlemektir. Modelin amacını (modelin hedef kitlesini, hedef kullanıcısı ve tahmin edilecek olay veya hastalık) tanımlamak, ürettięi bilgileri hastaya fayda saęlayacak bir klinik eylemle

ilişkilendirir (Steyerberg ve Vergouwe, 2014; Lee vd., 2016; Shipe vd., 2019; Chen L. 2020).

1.1.1.2. Veri Kümesi Seçimi

1.1.1.2.1. Çalışma tasarımı

Klinik tahmin modelinin oluşturulmasında veri kümesi seçimi, modelin doğruluğu ve genelleştirilebilirliği açısından son derece önemlidir. Prognostik çalışmalar çoğu zaman bir sonucun ortaya çıkması için zaman içinde takip edilen hasta gruplarında gerçekleştirilir. Kohort, birçok özelliğin varlığıyla tanımlanır. Bu duruma hastalık, yaş ve sadece canlı doğmak örnek olarak verilebilir. Tanı araştırmaları çoğunlukla, tahmini hasta özelliklerinin altta yatan bir tanıyla ilişkili olduğu için kesitsel bir çalışma olarak tasarlanır. Çalışma grubu, denekleri belirli (hedef) bir hastalığa sahip olduğundan şüphelenilen belirli bir belirti veya işaretin varlığı ile tanımlar. Örneğin acil servise başvuran hastalardan akut miyokard enfarktüsü olanlara tanı koyma amaçlanabilir (Steyerberg ve Vergouwe, 2014; Shipe vd., 2019; Chen, 2020).

Model geliştirmede ileriye dönük olarak toplanan bir kohort kullanıldığında hastalar iyi tanımlanır, ilgilenilen tüm değişkenler toplanır ve kayıp veriler en aza indirilir (Shipe vd., 2019). Prospektif tasarım, gerçek dünya koşullarını daha iyi yansıtır ve modelin gerçek yaşamda nasıl performans göstereceğini tahmin etmede daha güvenilir sonuçlar sağlar. Veri toplama süreci çok pahalı, zaman alıcı ve kayıp veri olması dışında ideal bir yoldur. Önceden var olan veri kümeleri (örn. retrospektif ve/veya büyük veri tabanı çalışmaları, rastgele seçilmiş deneme verilerinin ikincil analizleri, vb.) prospektif tasarıma göre hızlı ve ucuz olduğu için klinik tahmin modellerinde yaygın olarak kullanılır. Ancak kayıp veri sorunu, önemli tahmin edicilerin toplanmaması, sonuçların yanlış değerlendirilmesi gibi bazı sınırlamaları da vardır. Bu nedenle retrospektif tasarım modelin gerçek dünya uygulanabilirliğini tahmin etme konusunda daha sınırlıdır (Shipe vd., 2019; Chen, 2020).

Araştırmacılar, tahmin modelinin amacına bağlı olarak farklı türde veri kümeleri kullanmalıdır (Lee vd., 2016). Örneğin; tanısal bir tahmin modeli için optimal çalışma

tasarımı, kesitsel kohort çalışmasıdır, prognostik tahmin modeli için ise tercih edilen tasarım uzunlamasına kohort çalışmasıdır (Cowley vd., 2019).

1.1.1.2.2. Sonuç değişkenin (Outcome) tanımları

Bir sonuç değişkeni tanımlanırken klinik olarak ilgili ve hastalar için anlamlı olan değişken seçilmelidir. Lojistik regresyon tahmin modelleri için bunlar ölüm, akciğer kanseri teşhisi veya hastalık nüksü gibi ikili sonuçları içerir. Sonuç belirleme yöntemi, tahmin edici toplama yöntemi gibi, doğru ve ilgili hastalık ve klinik uzmanlık spektrumunda tekrarlanabilir olmalıdır (Shipe vd., 2019). Tahmin edicilerin sonucun belirlenmesinde kullanılmaması gerekmektedir. Ancak, tüm mevcut hasta bilgilerine dayalı bir uzlaşma tanısı gerektiren bazı durumlarda bu mümkün olmayabilir. Sonuç değişkeninde yanlış sınıflandırma, tahmin doğruluğunda ciddi sorunlara yol açabileceği bilinmektedir (Cowley vd., 2019).

1.1.2. Model geliştirme

1.1.2.1. Tahmin edicilerin tanımlanması

İyi performans gösteren bir risk tahmin modeli, bazı güçlü tahmin edicilerin mevcut olmasını gerektirdiğinden, model geliştirme veri toplama noktasından başlamaktadır (Steyerberg ve Vergouwe, 2014; Shipe vd., 2019).

Tahmin edicileri belirlemek için iki ana strateji (klinik ve veri odaklı) vardır. Klinik odaklı tahmin edici belirlemede, aday tahmin edici ya araştırma grubundaki klinik uzmanlar tarafından ya da literatür taramasıyla seçilir. Veriye dayalı tanımlamada, tahmin edicilerin seçimi için çok sayıda makine öğrenimi tekniği vardır. Bunlar yaygın olarak tek değişkenli analiz tekniklerini kullanır veya her tahmin edici için tek bir çıktı üretir ya da tahmin edicileri belirli bir istatistiğe göre sıralar ve en üst sıradaki tahmin edicilerden kaçının seçileceğine karar verir (Shipe vd., 2019). Tahmin edici seçiminde kullanılan bir diğer yöntem ise doğrusal, lojistik ve Cox modelleri içeren regresyon analizleridir (Lee vd., 2016).

Tahmin ediciler (demografik değişkenler, klinik öykü, fizik muayene bulguları, hastalığın tipi ve şiddeti, komorbid durumlar ve laboratuvar veya görüntüleme sonuçları vb.) açıkça tanımlanmalı, standardize edilmiş ve tekrarlanabilir bir şekilde ölçülmelidir, aksi takdirde tahmin modeli geliştirilememektedir. Sonuç değişkeni ile güçlü bir şekilde ilişkili olan (sonuç değişkeninde gözlenen değişimi açıklayan veya diğer değişkenlerle kombinasyon halinde etkileşime giren) tahmin ediciler, modele dahil edilmeye aday olurlar (Steyerberg, 2009; Moons vd., 2012).

Aday tahmin edici değişkenler, çok değişkenli bir modele dahil edilmek üzere seçilen değişkenlerdir ve sonunda final modele dahil edilenlerden farklılık gösterirler (Cowley vd., 2019). Önceden anlamlı bulunan tahmin ediciler normalde aday değişkenler olarak düşünülmelidir (örneğin, diyabet risk skorunda ailede diyabet öyküsü). Ancak tüm önemli tahmin edicilerin nihai modele dahil edilmesi gerekmediğine dikkat edilmelidir ($P<0,05$) (Lee vd., 2016). Tahmin edici değişkenler tek değişkenli analizde istatistiksel olarak anlamlı olmasalar da aday değişken olarak alınabilir (Cowley vd., 2019). İstatistiksel olarak anlamlı olmayan veya etki büyüklüğü küçük olan değişkenler de modele katkıda bulunabilir. Araştırmacının takdirine bağlı olarak, farklı kullanıcıları hedeflemek için farklı değişkenleri analiz eden farklı modeller geliştirilebilir (Lee vd., 2016).

Aday tahmin ediciler, klinik deneyim, uzman görüş anketleri, niteliksel çalışmalar ve literatürün sistematik incelemeleri ve meta-analizleri gibi bir kombinasyon temel alınarak seçilebilir (Cowley vd., 2019). Aday tahmincilerin sayısını azaltmak için kullanılabilecek stratejiler arasında, diğer tahmincilerle yüksek bir ilişkiye sahip olanları çıkarmak ve benzer tahmincileri birleştirmek yer alır (Harrell, 2001). Tahmin edicilerin final modeli oluşturulurken nasıl seçileceği konusunda ortak bir fikir bulunmamaktadır. Ancak, iki ana strateji ve bunların çeşitleri bulunmaktadır. Yaygın olarak kullanılan bu iki ana strateji “tam model yaklaşımı” ve “tahminci seçim yaklaşımı”dır (Royston, vd., 2009).

Tam model yaklaşımında, önceden belirlenmiş olan tüm aday tahminciler çok değişkenli analizlere ve dolayısıyla final tahmin modeline dahil edilir; hiçbir aday tahmin edici elenmez. Bu yaklaşım, seçim yanlılığını ve modelin aşırı uyumunu engellemesine rağmen, en uygun aday tahminciler hakkında derinlemesine bilgi sahibi olmayı ve sınırlı sayıda tahmin edicinin kullanılmasını gerektirir (Cowley vd., 2019; Shipe vd., 2019). Ayrıca tüm tahmin edicileri önceden belirlemek pratik olmayabilir ve önceden önemli olan tahmin ediciler yeni bir popülasyonda/örnekte olmayabilir (Lee vd., 2016). Tahminci seçim

yaklaşımında ise, tahminciler "geriye doğru eleme" veya "ileriye doğru eleme" yöntemiyle, önceden belirlenen kriterlere göre seçilir. Geriye doğru eleme yöntemi, modeldeki tüm tahmin edicilerle başlar ve tahmin edicileri aşama aşama çıkarırken, ileriye doğru eleme, boş bir modelle başlar ve tahmin ediciler aşama aşama eklenir (Cowley vd., 2019; Chowdhury ve Turin, 2020). İleriye doğru eleme, tüm aday değişkenlerin etkilerinin eş zamanlı olarak değerlendirilmesini sağlayamaz ve ilişkili değişkenler modelde kalmayabilir. Bu nedenle, tahmin modeli geliştirmede ileriye doğru seçim önerilmemektedir. Çünkü yeniden üretilmesi zor, tahmin edici etkileşimlerini hesaba katmayan ve önemli tahmin edicilerin yanlışlıkla ortadan kaldırılabileceği modellerle sonuçlanmaktadır (Shipe vd., 2019). Geriye doğru eleme, tüm tahmincilerin etkilerini eşzamanlı olarak değerlendirmeye ve tahminciler arasındaki tüm ilişkileri göz önünde bulundurmasına izin verdiği için tercih edilir (Cowley vd., 2019; Chowdhury ve Turin, 2020).

Kullanılan tahmin edici seçim tekniğinden bağımsız olarak, dahil etme için anlamlılık düzeyi seçimi final modeli etkileyecektir (Shipe vd., 2019). Tüm yöntemlerde, anlamlılık düzeyi seçilen tahmincilerin sayısını etkiler; daha küçük anlamlılık düzeylerinin (örn. $P<0,05$) kullanılması, potansiyel olarak önemli tahmincileri dışlamak riskini taşıyan daha az tahminci içeren modellere yol açarken, daha büyük anlamlılık düzeylerinin (örn. $P<0,25$) kullanılması, daha az önemli tahmincilerin dahil edilmesine neden olabilir (Pimentel vd., 2021; Cowley vd., 2019). Literatürde tahmin edici kabiliyetine dair güçlü kanıtlar varsa, modelde anlamsız bir tahmin ediciyi korumak kabul edilebilir olabilir (Setoguchi vd., 2008; Cowley vd., 2019).

1.1.2.2. Değişken türleri (sürekli ve kategorik) ve tahmin edicilerin etkileşimleri

Kategorik değişkenlerde, nadiren karşılaşılan kategoriler ve benzer kategoriler birleştirilebilir. Sürekli değişkenler genellikle doğrusallık varsayımı ile regresyon modeline dahil edilirken, araştırmacılar doğrusal olmayan ilişkilerin olasılığını da göz önünde bulundurmak zorundadır. Doğrusal terimler, J veya U şeklinde olan tahminciler gibi doğrusal olmayan ilişkilere sahip tahminciler için uygun değildir. Bu tür şekiller, sınırlı kübik spline veya kesirli polinomlar kullanılarak verimli bir şekilde modellenabilir (Steyerberg ve Vergouwe ark., 2014; Lee vd., 2016). Sürekli değişkenlere sahip tahmin modelleri risk puanlarını kendi başlarına hesaplamaları gerektiğinden, karmaşık, kullanımı zor ve meslekten olmayan kişiler tarafından anlaşılması zor olabilir. Bu modellerin uygulanması

için genellikle web veya bilgisayar tabanlı bir platform gereklidir. Aksi takdirde, daha sonraki bir aşamada, modelin tahmin kapasitesi korunursa, araştırmacılar bazı tahmin edicileri kategorilere ayırarak modeli kullanıcıların daha kolay anlayabileceği bir formata dönüştürebilirler (Lee vd., 2016). Sürekli değişkenleri kategorilere ayırma sırasında değerli tahmin bilgileri kaybedilebilir (Lee vd., 2016; Lee vd., 2010).

Etkileşim, bir değişkenin sonuca etkisinin başka bir değişkene bağlı olması nedeniyle, iki tahmin edici değişkenin etkisi ayrılamadığı zaman ortaya çıkmaktadır. Etkileşimi test etmek için, ana etkilerin çarpımından oluşan bir terim eklenmelidir. Bu terimin katsayısı daha sonra tahmin edicilerin kombinasyonunun sonuca etkisini görmek için analiz edilir (Shipe vd., 2019).

1.1.3. Model seçimi

Model, verilerde bulunması muhtemel olan ve matematiksel olarak modellenen yanıt değişkeni için uygun olan (doğrusal, ikinci dereceden, vb.) genel yapılara uymalıdır. Model geliştirme süreci başka bir analist tarafından yeniden üretilebilecek kadar şeffaf ve ayrıntılı olmalıdır. Modeller ya bir istatistiksel modelden (regresyon analizi ve yaşam analizi gibi) ya da makine öğrenme teknikleri (destek vektör makine modelleri ve ağaç tabanlı modeller gibi) aracılığıyla seçilir (Shipe vd., 2019).

1.1.4. Model performansı

Geleneksel genel performans ölçüleri, tahmin edilen sonuçlar ile gerçek sonuç arasındaki mesafe ile belirlenebilir. İkili (binary) sonuçlar (outcome) için R^2 (Nagelkerke's R^2) veya Brier skoru, genel performans ölçüleri olarak sunulabilir. Duyarlılık, seçicilik, pozitif ve negatif tahmin değerler gibi performans ölçütleri model performansını değerlendirmek için kullanılır. F1 skoru, doğruluk ölçütlerinden biridir ve sınıflandırma problemlerinde modelin performansını değerlendirmek için kullanılır. Ayrıca tıpta yaygın olan tahmin modelindeki sonuç değişkeni ikili (binary) olduğu durumda, modelin performansının en önemli yönleri kalibrasyon ve ayırt edicilik olarak değerlendirilebilir (Han vd., 2016).

Genel model performansı, tahmin edilen sonuç ile gerçek sonuç arasındaki uzaklığın hesaplanmasıyla ölçülebilir (Cowley vd., 2019). R^2 (Nagelkerke's R^2), lojistik regresyon için olayı açıklama varyasyonunu değerlendirir ve gözlenen sonuç (0 veya 1) ile tahmin edilen risk arasındaki korelasyonun karesidir (Royston vd., 2009). Brier skoru, ikili bir sonucu tahmin eden modellerin performansını değerlendirmek amacıyla sıklıkla kullanılır. Brier skoru, mükemmel bir model için 0'dan, sonucu %50 görülme sıklığına sahip bilgilendirici olmayan bir model için 0,25'e kadar değişen puanlarla, gerçek ve tahmin edilen ikili sonuçlar arasındaki kare farklarını karşılaştırır. Brier skoru, hayatta kalma sonuçlarına uygulanmak üzere de uyarlanabilir (Shipe vd., 2019).

Alıcı işlem karakteristiği (ROC) eğrileri, bir klinik tahmin modelinin tüm kesim noktaları arasındaki duyarlılık ve seçicilik değerlerini gösterir ve optimal bir kesim noktası seçmek için kullanılabilir. Optimal kesim noktası (cutoff) elde edildikten sonra duyarlılık, seçicilik, pozitif ve negatif tahmini değerler, doğruluk, F1 ve MCC gibi performans ölçütleri performansı değerlendirmek için kullanılır. Tanı testi ile referans test sonuçlarından karışıklık matrisi oluşturularak performans ölçütlerini incelemek mümkündür.

Çizelge 1.1. Karışıklık Matrisi (Confusion Matrix)

		Gerçek durum (Referans test)		Toplam
		Pozitif	Negatif	
Tahmin (Tanı testi)	Pozitif	DP	YP	DP + YP
	Negatif	YN	DN	YN + DN
	Toplam	DP + YN	YP + DN	DP + YP + YN + DN

DP= Gerçek pozitiflerin doğru bir şekilde tahmin edilenleri “doğru pozitifler”

YP= Gerçek negatiflerin yanlış tahmin edilenleri “yanlış pozitifler”

YN= Gerçek pozitiflerin yanlış tahmin edilenleri “yanlış negatifler”

DN= Gerçek negatiflerin doğru bir şekilde tahmin edilenleri “doğru negatifler” olarak adlandırılır.

Duyarlılık, bir tanı testinin gerçek pozitif ve seçicilik ise bir tanı testinin gerçek negatif tespit etme yeteneğidir. Diğer bir ifade ile duyarlılık gerçekte pozitif olan

bireylerden yüzde kaçının tanı testinde de pozitif olduğunu belirler. Seçicilik ise gerçekte negatif olan bireylerden yüzde kaçının tanı testinde de negatif olduğunu belirler.

$$\text{Duyarlılık} = DP / DP + YN$$

$$\text{Seçicilik} = DN / YP + DN$$

Pozitif tahmini değer, pozitif tanı testi sonucuna sahip bireylerin gerçekte pozitif olma olasılığıdır.

$$\text{Pozitif tahmini değer} = DP / DP + YP$$

Negatif tahmini değer ise negatif tanı testi sonucuna sahip bireylerin gerçekte negatif olma olasılığıdır. (Trevethan, 2017; Cowley vd., 2019).

$$\text{Negatif tahmini değer} = DN / YN + DN$$

Doğruluk, doğru tahminlerin (DP ve DN) tüm tahminlere oranını ifade eder.

$$\text{Doğruluk} = DP + DN / DP + YP + YN + DN$$

F1 skoru, kesinlik (precision) ve duyarlılık (recall) değerlerini birleştirerek modelin dengeli performansını ölçer.

Kesinlik (precision), modelin pozitif olarak tahmin ettiği verilerin gerçekten pozitif olma oranını ifade eder. Kesinlik, yanlış pozitif sonuçların (false positive) az olmasını önemser.

$$\text{Kesinlik} = DP / (DP + YP)$$

Duyarlılık (Recall), gerçek pozitif verilerin model tarafından doğru bir şekilde tespit edilme oranını ifade eder. Duyarlılık, yanlış negatif sonuçların (false negative) az olmasını önemser.

$$\text{Duyarlılık} = DP / (DP + YN)$$

F1 skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalaması olarak hesaplanmaktadır. F1 skoru, modelin hem yanlış pozitif hem de yanlış negatif tahminlerini azaltarak, dengeli bir performans göstermesine olanak tanır. Formülü aşağıda verilmiştir:

$$F1 \text{ Skoru} = 2 * (\text{Kesinlik} * \text{Duyarlılık}) / (\text{Kesinlik} + \text{Duyarlılık})$$

F1 skoru, 0 ile 1 arasında bir değer aralığında bulunur. 1'e yaklaşan bir F1 skoru, modelin daha iyi bir performans sergilediğini gösterir. Ancak, kesinlik ve duyarlılık arasında bir denge kurmak gerekir, çünkü birini arttırmak diğerini azaltabilir. Bu nedenle, F1 skoru, modelin dengeli bir şekilde sınıflandırma yapma yeteneğini ölçmek için kullanılır (Ramesh vd., 2022; Islam vd., 2020).

Matthews Korelasyon Katsayısı (MCC), 2x2 bir karışıklık matrisine dayanır ve doğru pozitifler (DP), doğru negatifler (DN), yanlış negatifler (YN) ve yanlış pozitifler (YP) içerir. MCC, bu değerleri kullanarak bir sınıflandırma modelinin performansını ölçer ve -1 ile +1 arasında bir değere sahip olabilir. MCC, dengesiz sınıflandırma problemlerinde daha duyarlıdır ve sınıflar arasındaki gözlem sayıları dengesiz olduğunda doğruluk gibi ölçütlerden daha güvenilir bir performans ölçüsü sağlar.

$$MCC = (DP * DN - YP * YN) / \sqrt{(DP + YP) * (DP + YN) * (DN + YP) * (DN + YN)}$$

Bir lojistik regresyon modelinin performansı, kalibrasyon ve ayırt etme yeteneği açısından değerlendirilebilir. Kalibrasyon, uygunluk derecesi ile ilişkilidir ve modelin tahminleri ile gerçek gözlemler arasındaki uyumu değerlendirir (Han vd., 2016). Bireysel tahmin edilen risk aralıklarıyla tanımlanan gruplar için gözlenen olay oranlarını tahmin edilen risklere karşı çizerek görsel olarak değerlendirebilir. İdeal olarak, gözlenen olay oranları (y eksenini) ve tahmin edilen olasılıklar (x eksenini) tüm olasılık aralığında uyumluysa, çizimde 45° bir çizgi görüntülenir (eğim 1'dir) (Royston vd., 2009; Cowley vd., 2019). Bu tür bir grafiksel çizim, ikili sonuçlar için Hosmer-Lemeshow uygunluk testinin veya sağ kalım sonuçları için Nam-D'Agostino testinin görsel bir gösterimidir (Royston vd., 2009; Han vd., 2016). Bu kalibrasyon testlerinde anlamlı bir P değeri ($P < 0,05$), farklı risk kategorileri için farklı performans gösterdiği için modelin iyi kalibre edilmediğini ifade eder. $P > 0,05$ ise uygun kalibrasyona sahip olduğunu gösterir. Hosmer-Lemeshow testinin bazı

dezavantajları vardır, küçük örneklem büyüklükleri için genellikle anlamlı olmaması ve neredeyse her zaman büyük örneklem büyüklükleri için anlamlı olması ve kötü kalibrasyonu değerlendirmek için sınırlı güce sahip olmasıdır (Han vd., 2016). Doğrusal regresyon için bu, bir dağılım grafiği olarak gösterilecektir; ikili bir sonuç için, bir eğim oluşturmak üzere yumuşatma (smoothing) teknikleri kullanılabilir. Bu, tahmin edilen olasılığa göre farklı risk grupları için de yapılabilir. Tahmine dayalı bir model tüm risk gruplarında eşit performans gösteremeyebileceğinden kalibrasyonun görsel olarak değerlendirilmesi yararlı olabilir (Shipe vd., 2019).

Ayrımcılık, olayları ve olay olmayanları ayırt etme yeteneği olarak tanımlanır (Lee vd., 2016). Bir modelin olumlu ve olumsuz durumları ne kadar iyi ayırt edebildiğini değerlendirmek için AUC (Alıcı İşlem Karakteristiği Eğrisi Altındaki Alan) ve C istatistiği yaygın olarak kullanılır. AUC, ROC eğrisinin altındaki alanı ölçer ve farklı kesim noktalarında modelin performansını değerlendirir. AUC, 0,5 ile 1 arasında değer alır. C istatistiği (indeksi) ise sıralı verilerde modelin doğruluğunu değerlendiren bir ölçüdür ve değerler genellikle 0,5 ile 1 arasında olur (Shipe vd., 2019; Carrington vd., 2020). Bu indeks çoğunlukla yaşam analizlerinde Cox regresyon modellerinin performansının değerlendirilmesinde kullanılır. Cox regresyon modelleri, olay zamanlarını (örneğin, ölüm) içeren sürekli zamanlı veri analizi için kullanılır. C istatistiği sansürlenmiş verileri de hesaba katabilir. AUC ise genellikle sınıflandırma problemlerinde, özellikle hastalık tanısı ve benzeri durumlarda kullanılır. C istatistiği, sadece iki sınıflı (binary) sınıflandırma modelleri için kullanılabilir, AUC ise çok sınıflı durumları da değerlendirebilir (Royston vd., 2009; Carrington vd., 2020).

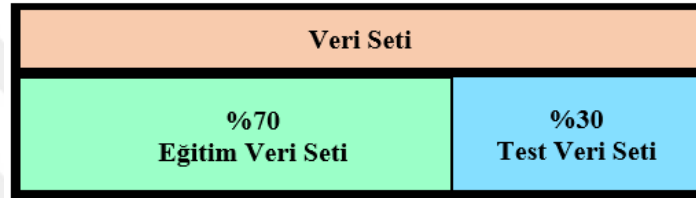
1.1.5. Model geçerliliği

1.1.5.1. Dahili doğrulama

Dahili doğrulama, modelin aynı veri seti üzerindeki performansını değerlendirir ve modelin eğitim verilerine ne kadar iyi uyarlandığını gösterir (Han vd., 2016). Dahili doğrulama, modelin eğitildiği ve test edildiği veri seti içinde gerçekleşir. İlk olarak, veri seti rastgele iki alt küme ayrılır. Birincisi modelin eğitildiği eğitim veri seti, ikincisi ise modelin performansının değerlendirildiği test veri setidir. Eğitim veri seti üzerinde modelin

parametreleri belirlenir ve daha sonra test veri seti üzerinde modelin performansı değerlendirilir. Dahili doğrulamada bölünmüş örneklem (split-sample), bootstrapping ve çapraz doğrulama (cross validation) yöntemleri yaygın olarak kullanılır (Cowley vd., 2019).

Bölünmüş örneklem yönteminde örneklem rastgele iki gruba ayrılır. Dahili doğrulamada, bir grup model geliştirmek için eğitim veri seti, diğer grup ise model performansını ölçmek için test veri seti olarak kullanılır. Yaygın olarak kullanılan bölünme oranları %70:%30, %80:%20 veya 2/3:1/3 şeklindedir (Steyerberg, vd., 2001; Cowley vd., 2019; Steyerberg, 2009).

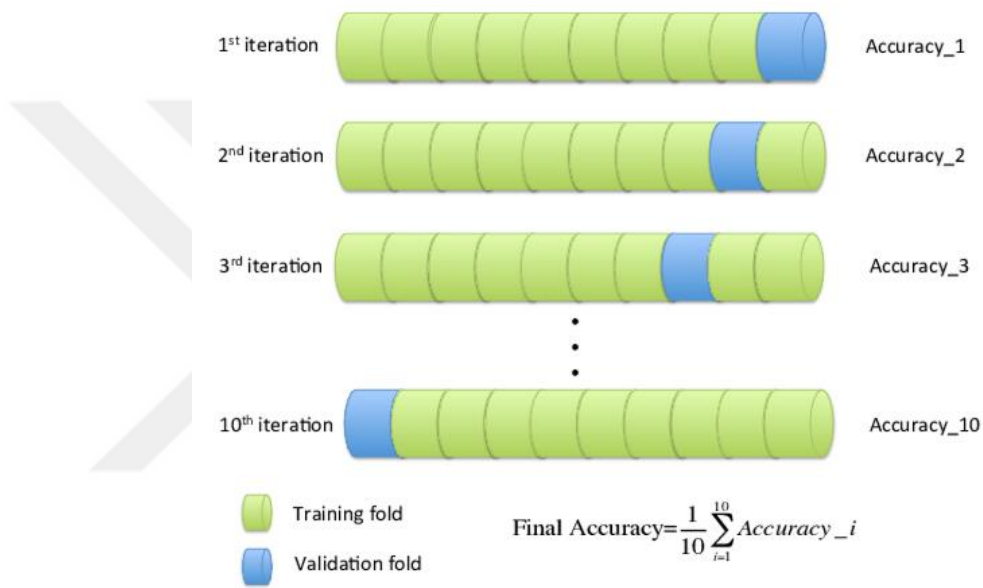


Şekil 1.1. Bölünmüş örneklem (split-sample)

Toplanan tüm veriler tahmin modelini geliştirmek için kullanılmadığından tüm veriler ile oluşturulan modele göre daha az istikrarlı model sonuçları görülür. Ayrıca doğrulama kısmı nispeten küçüktür, bu da model performansının güvenilir bir şekilde değerlendirilmesini zorlaştırmaktadır (Steyerberg, 2009). Doğrulama çalışma popülasyonu içinde yapıldığından dahili doğrulama, kayıp verileri veya ölçüm hatalarıyla seçim yanlılığını ele almamaktadır (Shipe vd., 2019). Bu sorunlardan kaçınmak için, genellikle diğer dahili doğrulama yöntemleri kullanılır (Shipe vd., 2019; Steyerberg, 2009).

Çapraz doğrulamada klinik tahmin modeli, tüm veri kümesini kullanılarak türetilir ve daha sonra performans değerlendirmesi için tüm veri kümesi yeniden kullanılır. K-katlı çapraz doğrulamada, örneklem k alt kümeye bölünür. 5 veya 10 örneklem yaygın olarak kullanılır. Bir alt küme doğrulama alt kümesi olarak seçilir ve geri kalan k-1 alt küme model geliştirme veya “eğitim” için kullanılır. Doğrulama alt kümesi daha sonra modelin tahmin hatasını hesaplamak için kullanılır. Süreç tamamlandıktan sonra, doğrulama alt kümesi olarak farklı bir alt kümeyi tutarak ve kalan alt kümeleri yeni bir model hesaplamak için kullanarak yeniden başlar. Bu k kez tekrarlanır ve her adımda hesaplanan hataların

ortalaması alınarak modelin performansı hesaplanır (Shipe vd., 2019; Steyerberg, 2009). Örneğin; beş örneklem olduğu durumda, model dört örneklemden dört kez yeniden düzenlenir ve performansı beşinci örneklem kullanılarak test edilir; bu işlem, her bir beş örneğin test verisi olarak kullanılana kadar beş kez tekrarlanır ve tahmini performansın ortalama değeri alınır. Stabilliteyi artırmak için, genel prosedür farklı rastgele alt örneklem kullanarak birden çok kez tekrarlanabilir (Cowley vd., 2019). Bu yöntemin avantajı, her gözlemin eğitim ve doğrulama için kullanılması ve her gözlemin yalnızca bir kez doğrulama için kullanılmasıdır.

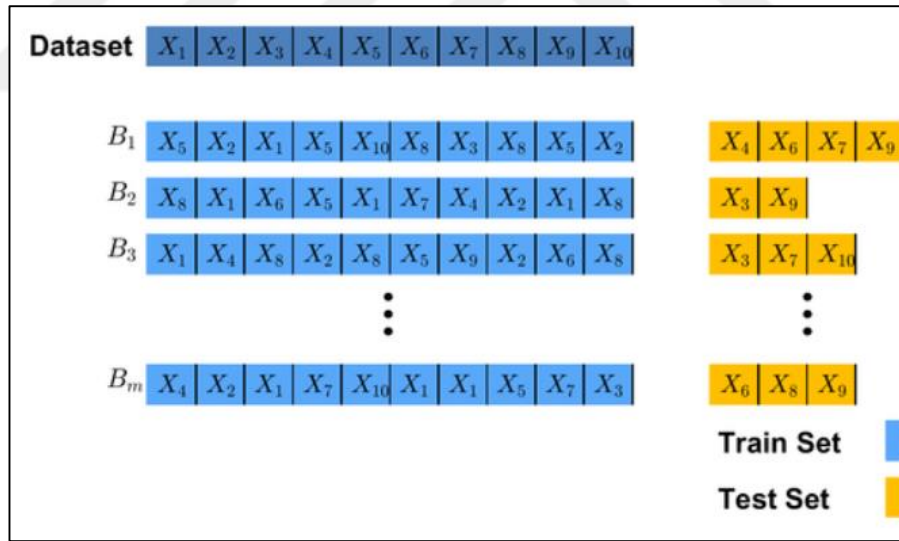


Şekil 1.2. Çapraz doğrulama (Pang vd., 2017)

Bootstrapping başka bir dahili doğrulama yöntemidir. Daha küçük örneklem boyutları veya daha fazla sayıda aday tahmin etmek için ideal bir yöntemdir (Shipe vd., 2019). Bootstrap, temel popülasyondan örnek çıkarma sürecini yansıtır. Bootstrap örnekleri, orijinal örnek boyutunda oluşturulur (Steyerberg, 2009). Bu yöntem, geliştirme veri kümesinden yerine koyarak aynı boyutta birden çok örneklem çekme yoluyla hedef popülasyondan rastgele örnekleme yaparak gerçekleştirilir (Cowley vd., 2019). Bootstrapping adımları model, orijinal veri kümesi üzerinde eğitilir ve bir performans metriği değeri kaydedilir. Daha sonra orijinal veri kümesinden yerine koyarak örnekleme yaparak yeterli sayıda bootstrap örneklemi oluşturulur. Bu işlem genellikle 100 veya 500 kez tekrarlanır (Han vd., 2016). Her bir bootstrap örnekleminde yeni bir model eğitilir ve her bir modelin

performans metriği değeri kaydedilir. Her bir bootstrap örnekleminde eğitilen modeller, orijinal veri kümesine uygulanır ve performans metriği ölçülür. Her bir bootstrap örnekleminde hesaplanan performans metriği değerleri ile bu örneklemlerin orijinal veri kümesinde test edildiğinde elde edilen performans metriği değerlerinin arasındaki farkların ortalaması alınarak modelin eğitildiği örnek veri kümesi üzerinde elde ettiği performansın, gerçek dünya verilerindeki performansından sapma eğilimi tahmin edilir. Bu sapma, modelin eğitim verilerine aşırı uyum (overfitting) yapma olasılığını temsil eder. İlk adımda kaydedilen performans metriği değerinden, tahmin edilen sapma eğilimi değerinin çıkarılmasıyla düzeltilmiş performans metriği hesaplanır. Bu, olası aşırı uyum durumlarını daha iyi hesaba katan bir "düzeltme" değeri sağlar (Lankham ve Slaughter, 2020).

Sonuç olarak bu yöntem, tüm orijinal verilerin model geliştirmede kullanılmasına izin verirken, orijinal modelde ne ölçüde aşırı uyum olduğuna dair bilgi verir (Shipe vd., 2019). Bootstrapping aynı zamanda aşırı uyumu düzelten bir düzeltme faktörünü tahmin eder (Cowley vd., 2019).



Şekil 1.3. Bootstrapping (Tohka ve Van Gils, 2021)

1.1.5.2. Harici doğrulama

Bir tahmin modelinin genel uygulanabilirliğini desteklemek için önemlidir. Dahili doğrulama teknikleri, tümü geliştirme ve test örneklerinin rastgele bölünmesi ile karakterize

edilirken, harici doğrulama, geliştirme hastalarından bazı açılardan farklı olan hastaları dikkate almaktadır (Steyerberg, 2009). Modeli geliştirmek için daha önce kullanılmayan bağımsız bir veri kümesi üzerinde harici doğrulama gerçekleştirilmelidir. Örneğin, veri kümeleri diğer hastanelerden veya merkezlerden (coğrafi doğrulama) veya daha yakın zamanda toplanan bir kohort popülasyonundan (geçici doğrulama) elde edilebilir (Lee vd., 2016). Harici doğrulama, sadece geliştirme aşamasındaki adımları yeni bir örnekle tekrarlamak ve aynı tahmin edicileri ve regresyon katsayılarını elde edip etmediğini gözlemlemek değildir; aynı şekilde, yeni bir örneklemede modeli tekrar uyarlama ve performansı geliştirme örnekleminde gözlenenle karşılaştırmaktır (Cowley vd., 2019). Yeni bireylere uygulandığında, bir tahmin modeli genellikle, iyimserliği ve aşırı uyumu (overfitting) düzeltmek için dahili doğrulama prosedürlerinden sonra bile, orijinal çalışma popülasyonundan daha kötü performans gösterebilir. Bu yüzden yeni bir tahmin modeli uygulamaya alınmadan önce, farklı bireylerle benzer bir popülasyonda tahmin değerini göstermek için harici olarak doğrulanmalıdır (Shipe vd., 2019). Bir tahmin modelini harici olarak doğrulamak için, aynı tahmin edici ve sonuç (outcome) verileri mevcut olan karşılaştırılabilir bir popülasyondan (dahil etme ve hariç tutma kriterleri açısından) bağımsız bir örnekleme ihtiyaç vardır. Model daha sonra dış popülasyondan elde edilen verilere uygulanarak sonuç risklerini tahmin eder, gözlemlenen sonuçla karşılaştırır ve modelin performansını hesaplar (Shipe vd., 2019).

1.1.6. Klinik tahmin modelinin sunumu

Bir klinik tahmin modelinin oluşturulmasındaki son adım, klinik tahmin modelinin sunumudur. Sunum formatına karar vermek için öncelikle “Modele bu formatta kimler erişecek ve onu ne zaman ve nerde kullanacaklar?” sorularını sormak önemlidir. Sunum daha sonra buna göre biçimlendirilebilir. Temel olarak, tam model denklemi sunulabilir (Bonnett, vd., 2019; Steyerberg ve Vergouwe, 2014). Bununla birlikte, sağlık profesyonellerinin modeli belirli bir klinik ortamda örneğin, bilgisayarlara veya mobil cihazlara erişimin sınırlı olduğu durumlarda kullanmalarını sağlamak için ek sunum biçimleri gerekebilir. Benzer şekilde, modeli evde kullanırken (örneğin, ebeveynler) ortak karar vermeyi geliştirmek için formatın sıradan insanlara göre uyarlanması gerekebilir (Bonnet vd., 2019). Sunuldukları ortam (kağıt veya elektronik), modellerin uygulanacağı ortam (örn. klinik veya evde), tahminlerdeki ayrıntı düzeyi (örneğin, yaklaşık veya yuvarlatılmış risk tahminleri veya kesin risk tahminleri) ve kullanım kolaylığı (basit veya

karmaşık formatlar) gibi durumlara dikkat ederek sunum formatına karar verilir (Steyerberg, 2009).

Bonnett vd klinik tahmin modellerini sunmak için geliştirdiği kılavuzda puanlama sistemi (points score system), grafiksel puan tablosu, nomogram ve web siteleri ile uygulamalar bulunmaktadır.

Klinik tahmin modeli, basit bir web tabanlı hesap makinesi veya uygulama olarak geliştirilebilir, böylece klinik tahmin modelinin kullanılabilirliği artırılabilir. Bu, karmaşık klinik tahmin modelleri için faydalı olabilir ve tahmin modelleri elektronik sağlık kayıtlarına entegre edilmesini sağlayarak klinik bakımda kullanılmasını kolaylaştırabilir. Nomogramlar, bir klinik tahmin modelinin yorumlanmasına ve anlaşılmasına yardımcı olabilir; ancak bunlar tam model formülü ile birlikte sunulmalıdır. Grafiksel puan tablosu, puan sistemlerinin oldukça basitleştirilmiş, renk kodlu versiyonlarıdır. Puanlama sistemleri, klinik tahmin modellerini basitleştirmek ve kullanımı kolaylaştırmak için sıkça kullanılır, örneğin tamsayı puanlama sisteminde regresyon katsayıları tamsayı puan değerlerine dönüştürülerek kolayca toplanabilir ve tahmin edilen olasılıklarla ilişkilendirilebilir (Cowley vd., 2019).

1.2. Lojistik Regresyon

Lojistik regresyon, ikili bağımlı bir değişkenin 1'e eşit olma olasılığının (örneğin, bir hastalığın varlığı veya yokluğu) bağımsız değişkenlere (örneğin, laboratuvar test sonuçları, demografik bilgiler) bağlı olarak tahmin edilmesinde kullanılır. Bu yöntem, veri setindeki ilişkileri ve değişkenler arasındaki etkileri değerlendirir (Austin ve Steyerberg, 2012).

Lojistik regresyon modelleri çok yönlüdür ve tıbbi ve tıbbi olmayan araştırmaların çeşitli alanlarındaki olguları tanımlamak için kullanılır. Lojistik regresyon modeli, ikili sonuçlar içeren tanısal (hastalığın varlığı) veya prognostik araştırma soruları için en yaygın kullanılan istatistiksel tekniktir. Model, ikili, kategorik ve sürekli tahmin edicileri, doğrusal olmayan dönüşümleri ve etkileşim terimlerini içerebilen esnek bir yapıya sahiptir (Austin ve Steyerberg, 2012; Uanhoro vd., 2021; Zabor vd., 2022). Lojistik regresyon, genelleştirilmiş bir doğrusal model örneğidir ve doğrusal regresyonun birçok ilkesi burada da uygulanır.

Doğrusal regresyondaki gibi, ikili sonuç y , bir dizi tahmin edici ve regresyon katsayılarının doğrusal bir kombinasyonuna bağlanır (Hilbe, 2016; Zabor, vd., 2022).

Lojistik regresyonda, ikili sonuç değişkeni doğrudan modellemek yerine lojit olarak adlandırılan log-odds işlevi model olarak kullanılır. Bu denklem aşağıda verilmiştir.

$$\text{lojit}(p_x) = \log\left(\frac{p_x}{1-p_x}\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k$$

Bu denklemde, $\hat{\beta}_0$ sabit (intercept) terimidir. $\hat{\beta}_0$ ile temsil edilen referans grubu, $X_1, \dots, X_k$ değişkenlerinin her birinin referans düzeyini sunan bireylerden oluşur. $\hat{\beta}_1 \dots \hat{\beta}_k$ ise $X_1, \dots, X_k$ ortak değişkenleriyle ilişkili değişken tahminlerini gösterir (Sperandei, 2014; Zabor vd., 2022). Bu denklemde $p/(1-p)$ “odds” olarak tanımlanır. Odds, bir olayın gerçekleşme olasılığının, gerçekleşmeme olasılığına bölünmesiyle hesaplanır. Bir olayın gerçekleşme olasılığı ile olayın gerçekleşmeme olasılığının toplamı 1 olmalıdır. Odds birden büyük olduğunda, olayın gerçekleşme olasılığı, gerçekleşmeme olasılığından daha yüksektir ve odds birden küçük ise, olayın gerçekleşme olasılığı, gerçekleşmeme olasılığından daha düşüktür (Sperandei, 2014; Zabor vd., 2022).

$\text{lojit}(p_x)$ denklemi aşağıdaki şekilde yeniden düzenlenebilir:

$$P(Y = 1|X) = \left(\frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k)} \right)$$

İkinci denklem, Y ikili sonuç değişkeninin tam olarak nasıl olasılıklara dönüştürüldüğünü gösterir. Veri kümesindeki her birey için, bireyin hasta olma olasılığı hesaplanabilir. Veri setindeki her birey için, $X_1, \dots, X_k$ 'nin her biri $\hat{\beta}_1, \dots, \hat{\beta}_k$ ile çarpılıp $\hat{\beta}_0$ eklendikten sonra değerlerin doğal üssü $[\exp(x)]$ alınır ve bireyin olayı yaşama olasılığı hesaplanır. Modelde tek bir sürekli değişken, varsa, bireyin hastalık olasılığı, $\left(\frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_k X_k)} \right)$ olarak hesaplanabilir. Bu olasılıklar, 0 ile 1 arasında değişen tahmin edilen değerlerdir. Burada 1 olasılığı, bir bireyin olayı yaşama olasılığının %100 olduğunu gösterir (Zabor vd., 2022).

Sonu olarak, yukarıdaki denklemlerde $\hat{\beta}_1, \dots, \hat{\beta}_k$ olarak lojistik regresyon tarafından kestirilen regresyon deęişkenleri, log-odds oranlarını temsil eder. Odds oranı iki olasılığın oranıdır. Bu nedenle katsayılar, dięer tüm deęişkenler modelde sabitken tahmin edici deęişkeninde bir birimlik deęişiklik olduęunda lojittteki deęişimin ne kadar olacağını gösterir. Bazı durumlarda (örneğin, ilgi duyulan olayın prevalansı çok düşükse), risk oranı odds oranından yaklaşık olarak hesaplanabilir (Zabor vd., 2022).

Bir lojistik model uygun (fit) hale getirildikten sonra, model katsayıları odds oranını (OR) verecek şekilde üstelleştirilebilir. Lojistik regresyonda $OR = \exp(\beta)$ olarak hesaplanır. Örneğin; β_1 'ın üstel deęeri alındığı durumda; yeni tedavi gören ancak daha genç bireylerin ortalama ölüm olabilirliği ($\exp(\beta_1) = \exp(0,454)$) referans bireylere göre 1,58 kat fazladır şeklinde yorumlanabilir. OR, 1'e eşitse, riskli grup ile referans grubun aynı olabilirliğe sahip olduğunu gösterir. $OR > 1$ ise, riskli grubun referans grubuna göre daha yüksek bir olabilirliğe sahip olduğunu gösterir. Son olarak, $OR < 1$ ise, riskli grubun referans grubuna göre daha düşük bir olabilirliğe sahip olduğunu gösterir. Ayrıca, her katsayının (ve odds oranının) standart hatası, güven aralığı oluşturmak ve hipotez testleri yapmak için kullanılabilir. Wald testi, katsayının (yani log odds oranının) sıfırdan önemli ölçüde farklı olup olmadığını veya eşdeęer olarak, odds oranını deęerlendirmek için kullanılır (Sperandei, 2014; Zabor vd., 2022).

1.2.1. Lojistik regresyon ile oluşturulan klinik tahmin modellerinin sunumları

Klinik tahmin modelinin sunumu, modelin başarısını etkileyebilir. Anlaşılır, kabul edilebilir, uygulanabilir ve güvenilir bir sunum, modelin klinik uygulamalarda kullanılmasını sağlar. Bu nedenle, model sonuçlarını iletmek için uygun bir sunum stratejisi oluşturmak önemlidir.

Klinik tahmin modellerinin sunumu, hedef kitlenin (saęlık profesyonelleri, hastalar, yöneticiler vb.) seviyesine ve ihtiyaçlarına göre uyarlanmalıdır. Sonuçları açıklayan ve anlamalarına yardımcı olan etkili bir iletişim yöntemi seçmek önemlidir (Steyerberg, 2009; Chen, 2020).

1.2.1.1. Tahmin modelinin lojistik regresyon formülü ile sunumu

Klinik tahmin modelleri çeşitli formatlarda sunulabilir. En basit haliyle, son regresyon formülü sunulabilir (Steyerberg ve Vergouwe, 2014; Han vd., 2016). Tahminlerin bir regresyon formülüyle hesaplanması iki adım içerir. İlk adım, doğrusal tahmincinin hesaplanmasıdır. Klinik tahmin modellerinde "Doğrusal Tahminci" (Linear Predictor), regresyon analizleri gibi yöntemlerle oluşturulan bir matematiksel ifadedir. Bu tahminci, klinik değişkenlerin (tahmin edici değişkenler) birleşik kombinasyonuna dayanarak, bir bireyin belirli bir olayı (örneğin hastalık gelişimi) gerçekleştirme olasılığını veya zamanı tahmin etmek için kullanılır. Doğrusal tahminci, farklı regresyon modelleri için farklı biçimlere sahip olabilir. Genellikle, doğrusal regresyon, lojistik regresyon ve Cox regresyonu gibi modellerde kullanılır (Steyerberg, 2009; Steyerberg ve Vergouwe, 2014; Bonnett, vd., 2019).

Doğrusal tahminci içinde, regresyon katsayıları, tahmin edici değerlerle çarpılır. Tahmin edicilerin tanımları ve kodlamaları kullanıcı için açık olmalıdır. İkili tahmin ediciler için, 0/1 kodlaması uygun olur, bu da karakteristik olmayan hastaların sıfır puanına sahip olduğu anlamına gelir. Kategorik tahmin ediciler için genellikle kukla değişkenler oluşturulur. Sürekli değişkenler için birimler açık olmalıdır (Steyerberg, 2009; Sperandei, 2014).

İkinci adım, doğrusal tahminci değerlerinin sonuç ölçeğindeki birimlere çevrilmesidir. Lojistik bir model için, sonucun olasılıklarını ($p(Y=1)$) tahmin etmek için lojit dönüşümü (lojistik dönüşümü) kullanılmaktadır (Steyerberg, 2009). Doğrusal tahminci hesaplanır ve ardından lojistik fonksiyon ($1 / (1 + e^{-(\text{doğrusal tahminci})})$) elde edilir.

Doğrusal tahminci, klinik tahmin modellerinde, bireylerin risklerini veya sonuçları tahmin etmek için temel matematiksel yapıyı oluşturan önemli bir bileşendir. Bu tahminci, klinik uygulamada hastaların risklerini anlamak, tedavi stratejilerini belirlemek ve hasta yönetiminde kullanışlı bir araçtır (Steyerberg, 2009; Bonnett, vd., 2019).

1.2.1.2. Tahmin modelinin web siteleri ve uygulamalar ile sunumu

Tahmin modelleri bir web sitesi, hesap makinesi, tablet veya akıllı telefon cihazı aracılığıyla kullanıma sunulabilir (Steyerberg ve Vergouwe, 2014). Web uygulamalarını çalıştırmak için istatistiksel yazılım paketlerini etkinleştiren belirli araçlar da dahil olmak üzere "Shiny for R" ve "SWire for Stata" gibi web sitesi oluşturma platformları vardır. Model web tabanlı bir uygulama ve mobil uygulama ile entegre edilir. Uygulamalar,

kullanıcıların gerekli girişleri yapmasına ve sonuçları almasına olanak tanıyan kolay kullanılabilir bir arayüz kullanmalıdır. Web siteleri ve uygulamalar, sağlık uzmanlarına ve halka açık ve dünyanın her yerinden ilgili kişiler tarafından kullanılabilir olmalıdır (Steyerberg, 2009; Steyerberg ve Vergouwe, 2014; Arostegui vd., 2019; Bonnett vd., 2019).

Web sitelerinin hedef kullanıcı ve hedef kitle hakkında açık olması gerekir. Modelin nasıl kullanılacağını açıkça belirtmelidir. Uygulamaların sonuçları kullanıcılar için anlaşılır ve görsel olarak sunacak şekilde düzenlenmelidir (Steyerberg, 2009; Steyerberg ve Vergouwe, 2014). Öngörülen olasılığın, temel alınan regresyon modelinden gelen tahminlerle uyumlu olup olmadığı kontrol edilmelidir. Sürekli tahmin edicilere sahip modeller için, (geliştirme veri kümesinin) aralığının dışındaki değerlerin girilmemesi için sınırlandırılmalı ve kullanıcıya bir uyarı sağlanmalıdır. Web uygulamalarında arka planda tam model denklemini kullanır ve bu nedenle herhangi bir yaklaşıma gerek yoktur. Web siteleri ve uygulamalar, çok sayıda tahmin edici, doğrusal olmayan terim ve etkileşim içeren karmaşık istatistiksel modellerin önünde kullanıcı dostu bir arayüz sağlayabilmelidir. Dijital uygulamalar, laboratuvar sonuçları ve metre olarak kaydedilen yükseklik gibi antropometrikler için birimler arasında kolayca geçiş yapılmasını sağlar. Elektronik sağlık kayıtlarına entegre edilmesini sağlayarak klinik bakımda kullanılmasını kolaylaştırabilir (Cowley vd., 2019; Arostegui vd., 2019). Ancak dezavantajı web siteleri ve uygulamaları kullanabilmek için internete erişim gereklidir. Veri gizliliği ve depolanması, özellikle bir web sitesi veri toplamak ve sunmak için tasarlanmışsa endişe kaynağı olabilir; bu endişe, web sitesinde veya uygulamada açıkça belirtilmelidir. Özellikle web tabanlı uygulama için, kullanıcı verilerinin güvenliğini sağlamak önemlidir. Kullanıcı bilgilerini ve verilerini koruyacak uygun güvenlik önlemleri alınmalıdır (Zhao vd., 2022; Bonnett vd., 2019).

Akut Apandisit için Alvarado Skoruna ilişkin web uygulama örneği Şekil 1.2.1.1'de gösterilmiştir.

MD
CALC

Search "QT interval" or "QT" or "EKG"

Login

SIGN UP

TAVSİYE

Ayrırcı tanınızda daima sağ alt kadrana ağrısının diğer nedenlerini göz önünde bulundurun; bunlar arasında (ancak bunlarla sınırlı olmamak üzere): ürolojik hastalık (böbrek taşı gibi), yumurtalık patolojisi (burulma, tubo-yumurtalık apsesi) ve diğer GI patolojileri (kolit, divertikülit) .

NOT:

- Sınır değerler çalışmaya göre farklılık gösterir ancak doğrulanmış bir sınıflandırma, erkeklerde ≥ 9 veya kadınlarda ≥ 10 puanlara en yüksek riski, erkeklerde ≤ 1 veya kadınlarda ≤ 2 puanlara en düşük riski atar (Coleman 2018).
- Yüksek riskli hastalarda BT görüntüleme olmadan tedaviyi, düşük riskli hastalarda ise alternatif tanıları değerlendirin.

İşaretler

Sağ alt kadrana hassasiyeti

No 0

Yes +2

Yüksek Ateş (37,3°C veya 99,1°F)

No 0

Yes +1

Rebound hassasiyeti

No 0

Yes +1

Belirtiler

Ağrının sağ alt kadrana doğru yer değiştirmesi

No 0

Yes +1

Anoreksiya

No 0

Yes +1

Mide bulantısı ya da kusma

No 0

Yes +1

Laboratuvar Değerleri

Lökosit >10.000

No 0

Yes +2

Lökosit sola kayma >75 nötrofil

No 0

Yes +1

9 points

Definite appendicitis by the Alvarado Score.

Copy Results

Next Steps

Şekil 1.4. Alvarado skoruna ilişkin web uygulama örneği (MDCalc)

1.2.1.3. Tahmin modelinin nomogramlar ile sunumu

Nomogramlar, klinik tahmin modeli için başka bir grafik sunum formatıdır. Puan skor sistemine benzer şekilde, belirli bir birey için tahmin edici değerlere dayalı olarak puanlar atanır ve bunlar daha sonra bir olay riski veya hayatta kalma olasılığı ile eşitlenir (Collins vd., 2014; Bonnett vd., 2019). Nomogramlar iki adımda oluşurlar. Birinci adım bir regresyon formülü gibi lineer tahmincinin hesaplanmasıdır. Nomogram, puanlama noktalarını okumak için bir referans çizgisine sahiptir (örneğin, 0-100 gibi). Kullanıcı manuel olarak puanları toplar ve bu toplam, doğrusal tahminciyle lineer olarak ilişkilendirilir. İkinci adım, lineer tahmincinin tahminlere dönüştürülmesidir ve bu tahminler nomogramın alt kısmından okunabilir. Tahminler olasılık veya diğer nicelikler şeklinde olabilir (Steyerberg, 2009). Nomogramlar, her bir özellik veya değişkenin bir çizgi üzerindeki bir puanla temsil edildiği bir diyagramdır. Bu puanlar daha sonra toplanır ve bir toplam puanla eşleştirilir (Grobman ve Stamilio, 2006; Shariat vd., 2009).

Nomogramlar, karmaşık tahmin modellerinin sonuçlarını göstermek için kullanışlı bir araçtır. Bir nomogramın amaçlanan kullanıcıları sağlık uzmanlarıdır. Kullanıcılar, belirli

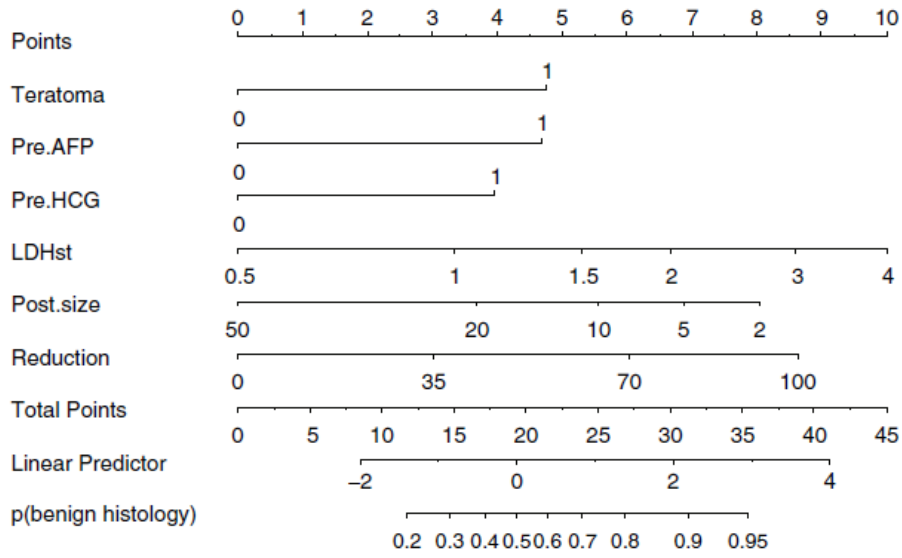
özelliklere sahip bir hasta veya hasta grubunun sonuçlarını hızlı ve kolay bir şekilde tahmin edebilirler (Steyerberg, 2009, Bonnett vd., 2019).

Nomogramların diğer sunum biçimlerine göre başlıca avantajları, sürekli tahmin edicilerin kategorilere ayrılmasına gerek olmaması ve olası toplam puanlara dayalı çoklu olasılık ölçekleri dahil edilerek birden çok zaman noktasının tek bir nomograma dahil edilebilmesidir. Ek olarak, tahmin edicilerin göreceli önemi, nomogram içindeki çizgilerin uzunluğu ile değerlendirilebilir. Etkileşim ve doğrusal olmayan terimler iyi bir şekilde ele alınabilir. Bir hasta veya hasta grubunun sonucunu görmek için puanları toplamak yeterlidir. Klinik uygulamalarda, nomogramlar hastaların tedavi seçeneklerini ve sonuçlarını daha iyi anlamalarına yardımcı olabilir. Bu da daha bilinçli kararlar alınmasına katkı sağlar. Ayrıca nomogramlar, farklı klinik bağlamlara uygun olarak özelleştirilebilir. Özellikle belirli bir hastalık veya durum için özel olarak tasarlanabilirler (Adams ve Leveson, 2012; Bonnett vd., 2019).

Nomogramların dezavantajı yayınlanan şeklin boyutuna ve çözünürlüğüne bağlı olarak hatalı olabilirler; modele dahil edilen tahmin edicilerin sayısı ne kadar fazlaysa, nomogramın yorumlanması o kadar zor olur. Katsayıların yuvarlanması da gerekebilir. Nomogramları kullanabilmek için, her özelliğin bir puanla temsil edildiği bir puanlama sistemi oluşturmak gerekebilir. Bu puanlama sistemi yanlış hesaplandığında veya yanlış kullanıldığında sonuçlar yanıltıcı olabilir. Nomogramlar belirli bir model veya hastalık için tasarlandıklarından, farklı klinik senaryolarda kullanılmaları sınırlı olabilir. Nomogramların yanlış yorumlanması veya yanlış kullanılması sonuçların yanlış anlaşılmasına yol açabilir (Adams ve Leveson, 2012; Bonnett vd., 2019).

Sonuç olarak, klinik tahmin modellerinin sonuçlarını anlamak ve iletmek için nomogramların kullanılması oldukça yaygın bir uygulamadır.

Ceza uygulamalı lojistik regresyon (penalized lojistik regresyon) modelinde altı tahminciye dayalı benign histoloji için nomogram örneği Şekil 1 de sunulmuştur.



Şekil 1.5. Ceza uygulamalı lojistik regresyon (penalized lojistik regresyon) modelinde altı tahminciye dayalı benign histoloji için nomogram (Steyerberg, 2009)

1.2.1.4. Tahmin modelinin skarlama sistemi ile sunumu

Klinik tahmin modellerinde skarlama sistemi, bir hastanın veya bir bireyin belirli bir sonuca (örneğin, bir hastalığın riskine veya bir tedaviye yanıt olasılığına) ilişkin bir tahminde bulunmak amacıyla kullanılır. Bu sistem, hastanın demografik bilgileri, laboratuvar sonuçları, klinik bulguları ve diğer özellikleri gibi değişkenleri kullanarak bir risk veya olasılık puanı üretir (Collins vd., 2014).

Skor sistemlerinde, puanlar bir birey için tahmin edici değerlere göre atanır. Toplam skor puanı daha sonra karşılık gelen bir olay riski veya hayatta kalma olasılığı ile eşleştirilir. Puan sistemleri için hedeflenen kullanıcılar sağlık profesyonelleri ve hastalardır (Bonnett, vd., 2019).

Skarlama sistemi, tipik olarak bir regresyon modeli (lojistik, cox regresyon vb.) kullanılarak geliştirilir. Bu model, bir kişinin sonuç üzerindeki etkilerini değerlendiren katsayıları içerir. Bu katsayılar, her bir özelliğin veya tahmin edicinin önemini ve katkısını yansıtır. Skarlama sistemi, bu katsayıları kullanarak bir kişinin sonuçla ilişkili olasılığını veya riskini hesaplar (Collins vd., 2014).

Skorlama sisteminde bağımsız değişkenlere ağırlıklar, regresyon katsayıları, risk oranı, makine öğrenme algoritmaları ve klinik değerlendirme temelinde bir puanlama algoritması kullanılarak atanır. Bir hastaya özgü ağırlıklar toplanır ve tek bir sayısal skor elde edilir (Grobman ve Stamilio, 2006; Mehta vd., 2016; Bonnett vd., 2019).

Çalışmamızda klinik tahmin modellerinde kullanılan puanlama sistemlerinde lojistik regresyon katsayılarına dayalı ve odds oranına (risk oranı) dayalı puanlama sistemleri kullanılmıştır.

Kısa vadeli klinik tahmin modelleri veya tanı modellerinde;

Sonuç değişkeni sadece iki seçenektan birini alıyorsa (örn. ölüm veya hastalık varlığı gibi) ve belirli bir zamanda tüm hastalar için biliniyorsa, lojistik regresyon kullanılarak bir tahmin modeli geliştirilebilir.

$$\log\left(\frac{p}{1-p}\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 \dots$$

burada p sonucun olma olasılığı ve log (p/1-p) sonucun (outcome) log odds'nu ifade eder. $\hat{\beta}_0$, sabit terimdir. X terimleri, dahil edilen tahmin edicilerin değerlerini gösterir. Örneğin, β_1 , her bir yıllık artış için lojitteki artış olabilir ve β_2 , bir kadınla karşılaştırıldığında bir erkek için lojitteki artış olabilir. Yeni bir birey için risk (sonuç olasılığı) tahminleri, tahmin edici değerleri denkleme girilerek ve ardından tekrar olasılık ölçeğine dönüştürülerek hesaplanabilir (Bonnett vd., 2019).

$$P(Y = 1|X) = \left(\frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 \dots)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 \dots)} \right)$$

Lojistik regresyon analizinde her bağımsız değişken için bir regresyon katsayısı sağlanır. Bu katsayılar, tüm değişkenlerin sonucu tahmin etmedeki etkisini ve yönünü gösterir. Her bir katsayı, bir risk veya olasılık puanı hesaplamak için kullanılır (Grobman ve Stamilio, 2006).

Odds oranına dayalı puanlama sisteminde odds oranları kullanarak puanlar atanır. Olayın gerçekleşme olasılığını etkileyen bir risk faktörüne sahip olanlar ile olmayanlar arasındaki ayrımı ortaya koyar. Odds oranları, lojistik regresyon analizi veya benzer bir analiz sonucunda elde edilir. Tüm risk faktörleri için bir odds oranı tespit edilir ve bu odds oranlarına dayalı olarak her bir faktör için bir ağırlık (puan) atanır. Bir hastanın özelliklerine göre bu puanlar toplanarak bir toplam puan hesaplanır (Valero III vd., 2015).

Her iki puanlama sistemi de bir hastanın sonucunu (örneğin, bir hastalığı olup olmama olasılığını) tahmin etmek için kullanılır. Hangi sistem tercih edilirse edilsin, her iki sistem de lojistik regresyon analizi sonuçlarına dayanmaktadır ve hastaların risklerini değerlendirmede önemli bir rol oynar.

Klinik tahmin modellerinde skorlama sisteminin avantajları; skorlar, klinik uygulamalarda daha kolay anlaşılabilir ve yorumlanabilir. Modelin karmaşıklığına ve dahil edilen tahmin edicilerin sayısına bağlı olarak, kağıt tabanlı puan skor sistemleri, hasta değerlerini doğrudan bir çevrimiçi hesap makinesine veya yayınlanmış bir modele koymaktan daha hızlı bir şekilde risklerin tahmin edilmesini sağlayabilir. Özellikle sağlık profesyonelleri için klinik kararlar alırken anlaşılabilir skorlar kullanmak önemlidir. Skorlama sistemleri, hasta veya durumun hızlı bir şekilde değerlendirilmesine yardımcı olabilir. Hekimler veya sağlık çalışanları, bir skor kullanarak hızlı bir ön değerlendirme yapabilirler. Skorlar, klinik kararlar alırken hekimlere veya sağlık çalışanlarına yardımcı olabilir. Örneğin, bir hasta için belirli bir tedaviye karar verirken skorlar kullanılarak daha bilinçli kararlar alınabilir. Skorlama sistemleri, hastaların veya durumların belirli bir sonuçla ilişkili riskini değerlendirmek için kullanılabilir. Örneğin, bir hastanın belirli bir hastalığı geliştirme riski hesaplanabilir.

Bu avantajlar, klinik tahmin modellerinin geliştirilmesi ve kullanılmasında skorlama sistemlerinin önemini vurgular. Ancak, bu sistemlerin doğru ve güvenilir olmaları için iyi bir şekilde tasarlanmaları ve doğru verilerle eğitilmeleri gerekmektedir (Bonnett vd., 2019).

Klinik tahmin modellerinde skorlama sisteminin dezavantajları; skorlama sistemleri, sürekli değişkenleri kategorilere dönüştürmek zorunda kalabilir. Bu, bilgi kaybına neden olabilir, çünkü sürekli değişkenlerin özgün dağılımı ve ilişkileri kaybedilir. Ek olarak skorlama sistemleri, orijinal regresyon modelinin doğruluğunu sınırlayabilir. Özellikle

kategorize edilen deęiřkenler kullanıldığında, tahminler daha az hassas olabilir (Bonnett vd., 2019).

Skor sistemi için, önce bir tahmin modeli geliştirilir (örneğin, lojistik veya Cox regresyon kullanılarak) ve daha sonra dahil edilen tahmin edicilerin regresyon katsayılarına, negatif veya pozitif olabilen tamsayı puanları atanır. Herhangi bir sürekli tahmin edici kategorize edilmelidir ve bu yüzden bir miktar tahmin doğruluęu feda edilir. Kategorilerin eşit büyüklükte olması gerekmez (Bonnett, vd., 2019).

Framingham çalışmasından yola çıkan Sullivan vd.adařları, risk tahminlerini basit bir şekilde hesaplayabilen bir puanlama sistemi geliřtirmiřtir. Bu skarlama sistemi hem sürekli hem de kategorik risk faktörlerinin dahil edilmesine izin verir. Framingham çalışmasında yařa göre standardize edilmiř puanlama sistemi kullanılır (Han vd., 2016; Sullivan vd., 2004).

Framingham çalışmasından yola çıkan Sullivan'ın puanlama sistemi geliřtirme adımları:

1. Sürekli tahmin edicileri kategoriler halinde düzenlenir ve her kategori için orta nokta belirlenir.

2. Her tahmin edici için bir referans kategorisi seçilir (sürekli, ikili ve kategorik).

3. Sürekli deęiřkenler için, her bir kategorinin referans kategorisinden ne kadar uzakta olduęu belirlenir ve ardından “regresyon birimleri”ndeki farkı belirlemek için her bir farkı o tahmin edicinin regresyon katsayısı ile çarpılır. İkili ve kategorik deęiřkenler için “regresyon birimi” o tahmin edicinin regresyon katsayısıdır.

4. Puanlama sisteminde 1 puana karřılık gelecek regresyon birimlerinin büyüklüęünü tanımlayın; bu tanım genellikle klinisyenin tercihine dayanır. Tahmin edicilerin kategorilerinin her biri ile iliřkili belirlenir.

5. En küçük ve en büyük olası puan toplamaları belirlenir.

6. Tahmin edilen olasılıkları elde etmek için elde edilen puanlar (yuvarlanmış) ile orijinal modeli kullanarak aralık boyunca her bir puan toplamı için risk tahmini hesaplanır. Bu tahmin tam modele yaklaşan yeni bir risk tahmin modelidir (Bonet vd., 2019).

Sonuç olarak puan sisteminde, sürekli risk faktörleri için kategoriler oluşturulur ve her birine bir referans değer atanır. Ardından bir temel risk faktörü kategorisi belirlenir ve bu kategoriye 0 puan atanır. Tüm puanları topladığında, temelde belirli bir kişinin risk faktörü profilinin referans profilden ne kadar farklı olduğu gösterilir. Sonuçlar yorumlanırken dikkatli olunması gereken bazı durumlar görülmektedir. Toplam puana dayalı olarak düşük, orta veya yüksek risk gibi karar verilebilir (Bonet vd., 2019). Puanların tahmin edilen bir riske çevrilebilmesi için olasılıklar tablosu (mutlak risk tahminleri) sunulabilir.

Sullivan'ın puanlama sistemi, bir regresyon modelinin sonuçlarına dayalı olarak risk puanlarını tahmin etmek için kullanılır. Özellikle kardiyovasküler hastalık riskini tahmin etmek için yaygın olarak kullanılır. Bu model cinsiyet, yaş, tansiyon, kan kolesterolü ve sigara içme durumu gibi faktörlere dayalı olarak kalp hastalığı riskini tahmin eder (Sullivan vd., 2004).

Regresyon katsayısına ve odds oranına (OR) dayalı diğer puanlama algoritmaları aşağıda sunulmuştur.

Regresyon katsayısına dayalı puanlama algoritmalarında isteğe bağlı olarak katsayıları daha yorumlanabilir hale getirmek veya belirli bir puanlama aralığına uyacak şekilde ölçeklendirme yapılabilir. Ölçeklendirme, katsayıların bir sabitle çarpılmasını, bölünmesini veya tam sayılara yuvarlanmasını içerebilir. Ağırlıkların elde edilmesi için regresyon katsayısına dayalı puanlama algoritmalarının bazıları aşağıda verilmiştir. Lojistik regresyondan elde edilen orijinal beta değerleri alınabilir. Beta katsayıları bir sabitle çarpılarak tamsayıya yuvarlanabilir. Örneğin; 10 ile çarpılarak en yakın tam sayıya yuvarlanabilir. Beta katsayıları, regresyon katsayısının en küçük mutlak değerine bölünüp daha sonra en yakın tam sayıya yuvarlanabilir (Mehta vd., 2016).

Ağırlıkların elde edilmesi için odds oranına (OR) dayalı puanlama algoritmalarının bazıları aşağıda verilmiştir (Mehta vd., 2016).

- Lojistik regresyondan elde edilen orijinal OR değerleri alınabilir.

- OR değerleri en yakın tam sayıya yuvarlanabilir.
- Johnson'ın puanlama sistemi kullanılabilir.

$OR < 1$ ağırlık = - 1

$1 \leq OR < 1,2$ ağırlık = 1

$1,2 \leq OR < 1,4$ ağırlık = 2

$1,4 \leq OR < 1,6$ ağırlık = 3

$1,6 \leq OR < 1,8$ ağırlık = 4

$1,8 \leq OR < 2,0$ ağırlık = 5

$OR \geq 2,0$, ağırlık = 6

Herhangi bir OR için, eğer $P > 0,10$ ise, ağırlık = 0 olarak atanır.

Johnson'ın puanlama sistemi, odds oranları (OR) kullanarak ağırlıklandırılmış puanlar üretir. Bu yöntem, özellikle lojistik regresyon modellerinin sonuçlarına dayalı olarak risk tahminleri yapmak için kullanılır. Örneğin, kanser riskini tahmin etmek için kullanılabilir (Johnson vd., 2006).

Tıbbi uygulama sırasında ana görevlerden biri risk sınıflandırması ve triyajdır. Kardiyovasküler hastalıklar için risk sınıflandırması, profilaksi ve terapötik müdahalelerin doğru tahsisi için hayati öneme sahiptir. Acil serviste hastaların triyajının yapılması oldukça önemlidir. Puanlama sistemi bu amaç için en kullanışlı araçtır. Kardiyovasküler hastalıklarda iyi bilinen örnekler, 10 yıllık kardiyovasküler olay riskini tahmin etmeye yardımcı olabilecek Framingham risk skorudur. Kritik hastalar için basitleştirilmiş akut fizyoloji skoru (SAPS) II, akut fizyolojik skor (APS) III, Lojistik Organ Disfonksiyon sistemi (LODS), Sıralı Organ Yetmezliği Değerlendirmesi (SOFA), Oxford Akut gibi çeşitli puanlama sistemleri de vardır (Zhang vd., 2017).

Puanlama sistemlerinden Framingham risk skoru örneği aşağıda verilmiştir.

Çizelge 1.2. Çoklu lojistik regresyon sonuçları

Risk Faktörü	Regresyon katsayısı	<i>P</i>	OR	OR için %95 Güven aralığı
β_0	- 10,5161	0,0001	—	
Yaş	0,0575	0,0001	1,059	1,046–1,073
Cinsiyet (Erkek)	1,3078	0,0001	3,698	2,579–4,956
Sistolik kan basıncı	0,0185	0,0001	1,109	1,012–1,025
Sigara içme durumu	0,9456	0,0001	2,574	1,958–3,384

Çoklu lojistik regresyon analizinden sonra puan hesaplama adımlarına devam edilir. B belirlenir. B, yaşı 5 yıl artması ile ilişkili risk artışını yansıtır.

Bu örnek için B değeri; $B = 5 \cdot (0,0575) = 0,2875$ 'dir.

Çizelge 1.3. Puanları hesaplama adımları

Risk Faktörü	Kategoriler	Referans değer (W_{iREF})	β_i	$\beta_i (W_{ij} - W_{iREF})$	Puan $\beta_i (W_{ij} - W_{iREF}) / B$
Yaş			0,0575		
	30-39	34,5 = W_{1REF}		0	0
	40-49	44,5		0,5750	2
	50-59	54,5		1,1500	4
	60-69	64,5		1,7250	6
	70-79	74,5		2,3000	8
Cinsiyet (Erkek)			1,3078		
	Kadın	0 = W_{2REF}		0	0
	Erkek	1		1,3078	5
Sistolik kan basıncı			0,0185		
	<120	107		-0,3330	-1
	120-129	125 = W_{3REF}		0	0
	130-139	135		0,1850	1
	140-159	150		0,4625	2
	≥ 160	170		0,8325	3
Sigara içme durumu			0,9456		
	Hayır	0 = W_{4REF}		0	0
	Evet	1		0,9456	3

*Örneklemdaki yaş aralığı 30-79'dur.

**Sistolik kan basıncının aralığı 78-240'tır. İlk ve son kategoriler için referans değerlerini belirlemek için 1. yüzdilik (94) ve 99. yüzdilik (190) kullanıldı.

Puanlar hesaplandıktan sonra risk tahmini aşağıdaki formüller ile hesaplanır.

$$\sum_{i=0}^p \beta_i X_i \approx -10,5161 + 0,0575 * 34,5 + 0,0185 * 125 + 0,2875 * (\text{toplam puan})$$

$$\hat{p} = \frac{1}{1 + \exp(-\sum_{i=0}^p \beta_i X_i)}$$

Örneğin; 55 yaşında, sigara içmeyen, sistolik kan basıncı 135 olan bir erkek alındığında toplam puan = 4 + 0 + 1 + 5 = 10 olarak hesaplandı.

$$\sum_{i=0}^p \beta_i X_i \approx -10,5161 + 0,0575 * 34,5 + 0,0185 * 125 + 0,2875 * (10) = -3,345$$

$$\hat{p} = \frac{1}{1 + \exp(3,345)}$$

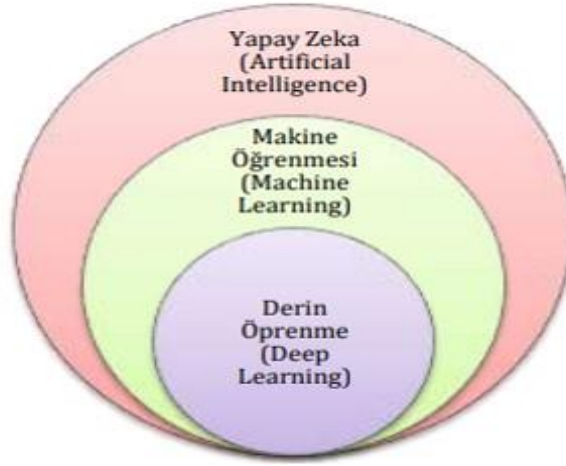
$$\hat{p} = 0,0341$$

Sonuç olarak; 55 yaşında, sigara içmeyen, sistolik kan basıncı 135 olan bir erkeğin 5 yıllık bir risk tahmininin %3 olduğu belirlendi.

1.3. Makine Öğrenme

Makine öğrenimi, bilgisayar bilimi, matematik ve istatistikten oluşan çok disiplinli bir alandır. Yapay zeka belirli görevleri gerçekleştirmek için doğrudan insan uyarısı olmaksızın bilgisayar yazılımı ve algoritmaları kullanan bilgisayar bilimi dalıdır (Boon vd., 2018; Shouval vd., 2021). Yapay zeka, otomasyon, robotik, oyun geliştirme, sağlık, otomotiv endüstrisi ve daha birçok alanda kullanılır. Makine öğrenme, bilgisayar sistemlerinin deneyimlerden öğrenmelerini ve verilere dayalı olarak gelecekteki görevleri daha iyi bir şekilde yerine getirmelerini sağlayan yapay zekanın bir alt dalıdır. Bu, belirli bir görevi el ile programlamak yerine, algoritmalara geniş veri kümesi üzerinden öğrenme yetisi kazandırmak anlamına gelir. Makine öğrenme yöntemleri, veri madenciliği alanında da kullanılır. Makine öğrenme, veriye dayalı kararlar alma, tahminler yapma ve örüntüleri tanıma konularında büyük önem taşır (Das vd., 2015; Eli-Chukwu, 2019).

Makine öğrenme algoritmaları, verilerden öğrenme yeteneğine sahip olan ve belirli bir görevi gerçekleştirmek için bu öğrenmeyi kullanan matematiksel ve istatistiksel işlemlerdir. Temel olarak, bu algoritmalar, verilerdeki örüntüleri ve ilişkileri tespit ederek, gelecekteki verilere dayalı tahminler veya sınıflandırmalar yapmak için kullanılır. Ayrıca makine öğrenme, büyük veri kümelerinden bilgi çıkarmak, desenleri tanımak ve öğrenmek için kullanılır (Hügler vd., 2020; Shouval vd., 2021).



Şekil 1.6. Yapay zeka, makine öğrenmesi ve derin öğrenme (Aylak vd., 2021)

Son zamanlarda araştırmacılar ve klinisyenler, farklı modalitelerdeki verileri kullanarak klinik sonuç tahmin etme, tanı koyma, tedavi ve optimal karar verme sorunlarını çözmek için makine öğrenme algoritmalarını uygulamaya başlamışlardır. Bu yapılan uygulamalar, yapısal laboratuvar ölçümleri, talep verileri, serbest metinler, görüntüler ve fizyolojik sinyaller dahil farklı modalitelerdeki verileri kullanmaktadır (Hügler vd., 2020; Henry vd., 2022; Ghassemi vd., 2014).

Makine öğrenme tabanlı yaklaşımlar veriler içindeki desenlerin tanınmasına yardımcı olur ve daha geliştirilebilir modellerle kişiselleştirilmiş klinik tahminler yapılmasına olanak tanır (Gehrmann vd., 2018). Klinik tahmin modellerinde kullanılan makine öğrenme algoritmaları, hastalıkların teşhis edilmesi, tedavi planlarının geliştirilmesi ve hastaların sağlık sonuçlarının tahmin edilmesi gibi bir dizi farklı uygulama için kullanılır (Hügler vd., 2020; Henry vd., 2022).

Yapay zeka algoritmaları genellikle üç ana kategori altında sınıflandırılmaktadır. Bu kategoriler sırasıyla denetimli öğrenme (supervised learning), denetimsiz öğrenme (unsupervised learning) ve pekiştirmeli öğrenme (reinforcement Learning) yöntemleridir (Shouval vd., 2021).

Denetimli öğrenme yönteminde, model, farklı veri kaynaklarından beslenerek elde ettiği sonuçları hedef çıktılarla mukayese eder. Oluşan hatalar sistem tarafından azaltılmaya çalışılır. Sonucun iyi tanımlandığı pratik makine öğrenme görevleri için en yaygın senaryodur. Bu algoritmaya hasta mortalitesi, hastane kalış süresi veya ilaç yanıtı tahmini

örnek olarak verilebilir (Celi vd., 2020). Denetimli öğrenmede algoritmaları genel olarak Sınıflama ve Regresyon olmak üzere iki başlıkta toplanmaktadır. Tez çalışmamızda iki düzeyli sonuç değişkeni kullanıldığı için sınıflama algoritmaları kullanılmıştır.

1.3.1. Makine Öğrenme Algoritmaları

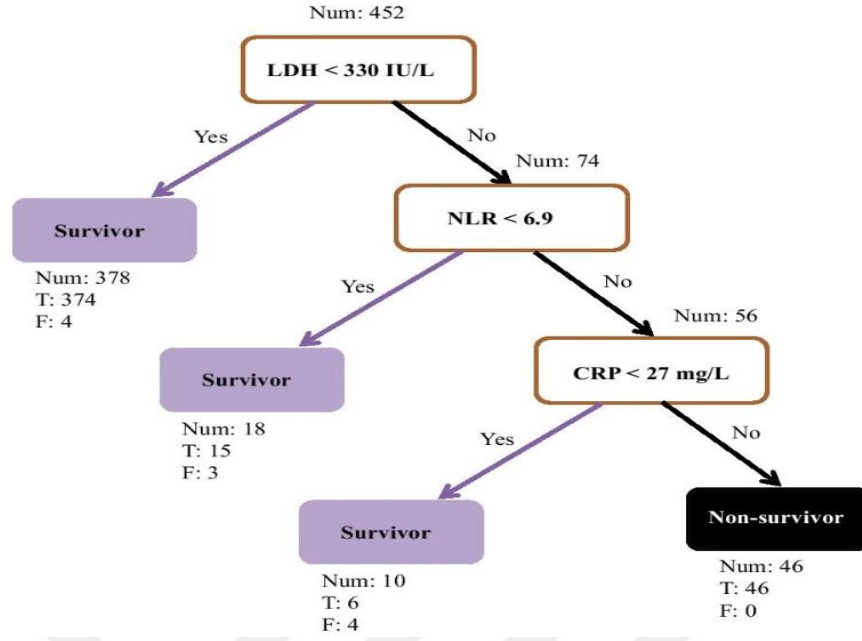
1.3.1.1. Karar Ağaçları (Decision Tree)

Karar ağaçları, sınıflandırma ve regresyon problemlerini çözmek için kullanılabilen modellerdir. Hastalıkları sınıflandırmak ve risk faktörlerini belirlemek için kullanılırlar. Karar ağaçları, karmaşık ilişkileri basit bir şekilde görselleştirebilir. Bu ağaçlar, bir akış diyagramı gibi bir çıktı üretirler. Her bir örnek değerine dayalı olarak gezilebilecek bir yol oluşturur ve sonuçta tahmin edilen bir değer elde edilir (Ronowicz vd., 2015; Mahesh, 2020; Zhou vd., 2021). Karar ağacı modeli, farklı sonuçlar arasında en iyi ayrımı yapan bir özellik seçilerek eğitilir, bu özellik (düğüm) üzerinde ağacı böler ve bu bölümü oluşturan her yeni düğüm üzerinde bu bölünmeyi tekrarlayarak bir ağaç benzeri bir grafik üretir. Bu, bilgiyi en iyi şekilde içeren özelliklerin modele dahil edilmesi nedeniyle içsel bir özellik azalmasıyla sonuçlanır (Richter ve Khoshgoftaar, 2018; Shouval vd., 2014).

Karar ağacı modelinin ne zaman bölünmeyi durduracağını belirleyen çeşitli değişkenler vardır (iterasyon sayısı, düğüm sayısı vb.). Model, her bir düğümde ağacı bölmek için özellikler seçtiği için her düğümde bir özellik seçimi yapar ve bu nedenle içsel bir özellik azalması meydana gelir. Bu, en bilgilendirici özelliklerin modele dahil edilmesi sonucunu doğurur (Zhou vd., 2021; Shouval vd., 2014).

Karar ağaçları, belirli bir hastalığın nüks olasılığını veya başka bir tıbbi olayı tahmin etmek için kullanılabilirler. Karar ağaçları, özellik seçimi için sıklıkla kullanılır. Karar ağaçları, kolay yorumlanabilir olmaları ve modelin içerdiği değişkenlerin önem sırasını belirleme yetenekleri nedeniyle klinik tahmin modellerinde yaygın olarak kullanılır (Rodriguez-Galiano vd., 2015; Zhou vd., 2021; Shouval vd., 2014).

Örnek: Şiddetli COVID-19 hastalarında ölüm sonucunu tahmin etmek için üç biyobelirteç ve bunların mutlak değerdeki eşiklerini kullanan bir karar ağacı modeli



Şekil 1.7. Şiddetli COVID-19 hastalarında ölüm sonucunu tahmin etmek için klinik özellikler ve karar ağacı modeli (Num, bir sınıftaki hasta sayısı; T, doğru sınıflandırılmış hasta sayısı; F, yanlış sınıflandırılan hasta sayısı; NLR, nötrofil-lenfosit oranı; CRP, C-reaktif protein; LDH, laktik dehidrojenaz) (Yang vd., 2021)

1.3.1.2. Rastgele Orman (Random Forest)

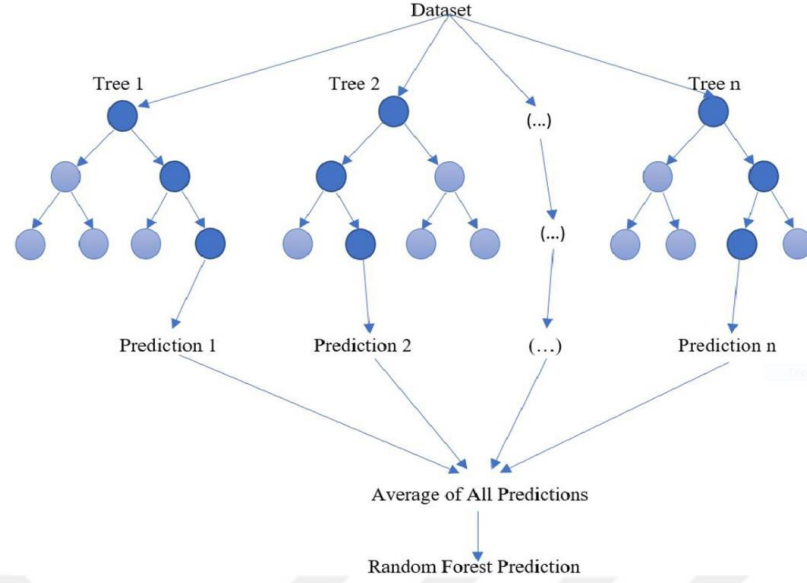
Rastgele orman klinik tahmin modelleri oluşturmak için önemli bir makine öğrenme algoritmasıdır. Birden fazla karar ağacının bir araya getirilmesiyle oluşturulan bir yöntemidir. Bu yöntem, modelin kararlılığını ve doğruluğunu artırabilir. Bu algoritma, özellikle büyük ve karmaşık veri kümesiyle çalışırken başarılı sonuçlar elde edebilir ve klinik uygulamalarda yaygın olarak kullanılır. Rastgele orman bir dizi karar ağacını kullanarak tahminler yapar ve bu ağaçların sonuçlarını bir araya getirerek nihai tahmini verir (Rodriguez-Galiano vd., 2015; Rigatti, 2017).

Rastgele orman, birden fazla karar ağacını kullanır. Her ağaç, rastgele seçilen bir alt küme veri üzerinde eğitilir. Bu, ağaçların çeşitliliğini artırır ve modelin daha iyi genelleme yapmasına yardımcı olur. Klinik tahmin modelleri için çok sayıda özellik veya değişken içerebilecek karmaşık veri kümeleri vardır. Rastgele orman, bu tür verilere uygundur, çünkü verilerin alt kümeleri üzerinde çalışır. Rastgele orman, her özelliğin tahmin sonuçlarına katkısını değerlendirmek için özellik önemi skorları üretebilir. Bu, hangi özelliklerin en önemli olduğunu belirlemeye yardımcı olabilir (Sekulić vd., 2020). Rastgele orman, aşırı

uyuma (overfitting) karşı dirençlidir. Çünkü her ağaç, yalnızca alt küme veri üzerinde eğitildiği için aşırı uyma riski azalır. Model performansını değerlendirmek için deneysel olmayan veriler kullanılabilir. Ayrıca, çapraz doğrulama gibi tekniklerle modelin sağlamlığını test etmek mümkündür (Schonlau ve Zou, 2020; Naji vd., 2021).

Rastgele orman algoritmasında ilk adım, klinik tahmin modeli oluşturmak için verilerin toplanmasıdır. Bu veriler, hastaların demografik bilgileri, laboratuvar sonuçları, tıbbi geçmişleri, ilaç kullanımı ve diğer klinik özelliklerini içerebilir. Veri, eksik değerlerin doldurulması, aykırı değerlerin ele alınması ve özellik seçimi gibi ön işleme adımlarından geçirilir. Bu, veriyi modellemeye hazır hale getirir. Rastgele orman, her ağacı farklı bir alt küme veriyle eğitmek için rastgele bir alt küme seçer. Bu, ağaçların çeşitliliğini artırır. Her alt küme veri ile bir karar ağacı oluşturulur. Karar ağaçları, verileri sınıflandırmak veya regresyon yapmak için kullanılır. Her düğümde, rastgele bir özellik seçimi yapılır. Bu, ağaçların farklı yönlerde büyümesini sağlar (Qi, 2012; Sekulić vd., 2020). Ağaçlar, veriyi bir dizi düğümde bölümlere ayırır. Her düğüm, en iyi bölme noktasını seçmek için veriyi değerlendirir. Birden fazla karar ağacının oylama sonuçlarına dayalı olarak bir sınıf tahmini yapılır (Sekulić vd., 2020). Örneğin, bir hastanın kanser riskini tahmin etmek için kullanılabilir. Regresyon problemi için, her ağaç tarafından yapılan tahminlerin ortalaması alınarak nihai tahmin yapılır (Oshiro, vd., 2012).

Model performansı iyileştirilmesi gerekiyorsa, değişken ayarları ve özellik seçimi gibi optimizasyon adımları uygulanabilir. Elde edilen sonuçlar ve tahminler klinik uzmanlara veya ilgili sağlık profesyonellerine sunulur. Bu, hasta tedavi planlarının ve klinik kararların desteklenmesine yardımcı olabilir. Birden fazla karar ağacının bir araya getirilmesi, daha güçlü ve dengeli tahminler elde edilmesini sağlar ve aşırı uyma riskini azaltır (Probst vd., 2019). Bu nedenle, klinik tahmin modelleri oluşturmak için popüler bir seçenektir. Ayrıca özellikle veriler karmaşık ve çok sayıda özelliğe sahipse, Rastgele orman algoritması tercih edilebilir. Rastgele Orman algoritmasına ilişkin şema Şekil 1.3.1.2’de sunulmuştur.



Şekil 1.8. Rastgele orman şeması (Uyanık vd., 2020)

1.3.1.3. Genelleştirilmiş Boosted Regresyon Modelleri (GBM)

Klinik tahmin modelleri için kullanılan bir makine öğrenimi algoritmasıdır. GBM'nin temel amacı, veri setindeki yapıyı anlamak ve bu yapıyı temsil eden bir tahmin modeli oluşturmaktır. Bu algoritma, zayıf öğrencileri bir araya getirerek güçlü bir model oluşturmayı amaçlar. Zayıf öğrenci veri setindeki karmaşıklıkları ve desenleri tam olarak ifade edemeyen, genellikle düşük doğruluk oranlarına sahip ve tek başlarına yeterince etkili olmayan öğrenci modelleridir. Örneğin, basit bir karar ağacı zayıf öğrenci olarak adlandırılabilir. GBM, regresyon problemleri için özellikle güçlü olan bir tekniktir ve karmaşık, doğrusal olmayan ilişkileri modelleme yeteneğine sahiptir. Ayrıca, sınıflandırma problemleri için de kullanılabilir (Özgür ve Doğanay Erdoğan, 2020; Demir ve Şahin, 2023).

GBM, genellikle bir zayıf öğrenci olan karar ağaçları ile başlar. İlk ağaç, temel ilişkileri modellemeye çalışır. İlk ağaçtan elde edilen tahminlerle gerçek değerler arasındaki hatalar incelenir. Bu hatalar, bir sonraki ağacın odaklanacağı alanlardır. İkinci ağaç, birinci ağaçtan kaynaklanan hataları düzeltmeye çalışır. Bu hataları azaltarak modelin performansını artırmaya yöneliktir. GBM, belirli bir sayıda ağaç ekleyerek modelini güçlendirir. Her ağaç, önceki ağaçların bıraktığı hataları düzeltmeye odaklanır. Her ağaç eklendikten sonra toplam hata analiz edilir. Bu, modelin geliştirilip geliştirilmediğini değerlendirmek için kullanılır (Özgür ve Doğanay Erdoğan, 2020; Touzani vd., 2018). GBM'nin önemli değişkenleri

arasında öğrenme oranı ve toplam ağaç sayısı bulunur. Öğrenme oranı, her ağacın model üzerindeki etkisini kontrol ederken, toplam ağaç sayısı GBM'nin karmaşıklığını belirler. GBM, aşırı uyma riskini azaltmak için düzenleme (regularizasyon) tekniklerini içerir. Bu, modelin genelleme yeteneğini artırabilir (Touzani vd., 2018).

Belirlenen sayıda ağaç eklenip öğrenme tamamlandığında, GBM bu ağaçları birleştirerek final tahmin modelini oluşturur. GBM, veri setindeki karmaşık ilişkileri modelleme yeteneği, direnç ve tahmin performansındaki esneklik nedeniyle klinik tahmin modellerinde yaygın olarak kullanılır (Touzani vd., 2018; Demir ve Sahin, 2023).

1.3.1.4. Ölçeklenebilir Gradient Boosting (Extreme Gradient Boosting Models-XGBoost)

Ölçeklenebilir gradient boosting (XGBoost), klinik tahmin modellerinde başarıyla kullanılan bir makine öğrenimi algoritmasıdır. Bu algoritma, öğrenme sürecindeki zayıf öğrencilerin bir araya getirilmesiyle güçlü ve yüksek doğruluğa sahip modeller oluşturmaya amaçlar (Demir ve Şahin, 2023; Chen ve Guestrin, 2016). Ölçeklenebilir ve büyük veri setlerine uyum sağlayabilen bir yapıya sahiptir. XGBoost, geniş bir hiperparametre ayarı sağlar ve bu sayede modelin karmaşıklığını kontrol etmek mümkündür (Şahin, 2022).

XGBoost'un temelinde, ağaç tabanlı öğrenme yöntemleri yer alır. Ağaçlar, karmaşık veri setlerindeki ilişkileri modellemek için kullanılır. XGBoost, gradient boosting yöntemini kullanır. Bu, modelin önceki ağaçlardan kaynaklanan hataları anlamasına ve bu hataları düzeltmesine olanak tanır (Alabdullah vd., 2022; Mitchell vd., 2018). XGBoost, ağaçları sırayla ekler ve her ağacın katkısını önceki ağaçların hatalarına göre ağırlıklandırır. Bu, her ağacın modelin genel performansına etkisini kontrol etmeye yardımcı olur. Her ağaç, öncekilerin hatalarını düzeltmeye odaklanarak, topluca daha güçlü bir model oluşturur. Algoritmanın performansını etkileyen önemli faktörler arasında öğrenme oranı ve düzenleme (regularizasyon) bulunur. Öğrenme oranı, her ağacın model üzerindeki etkisini belirlerken, düzenleme (regularizasyon) aşırı uyma riskini azaltır (Chen ve Guestrin, 2016; Mitchell vd., 2018). XGBoost'un başarısını belirleyen diğer değişkenler arasında ağaç derinliği ve ağaç sayısı yer alır. Dikkatlice ayarlanmış bu değişkenler, modelin genel doğruluğunu artırabilir. XGBoost, genellikle çapraz doğrulama yöntemleriyle birlikte kullanılır. Bu, modelin farklı veri setlerindeki performansını değerlendirir ve en iyi modelin seçilmesine yardımcı olur.

XGBoost, gradient boosting yöntemini kullanarak ağaçları birleştirir. Bu, bir dizi zayıf öğreniciyi bir araya getirerek güçlü bir tahmin modeli oluşturulmasını sağlar (Chen ve Guestrin, 2016; Alabdullah vd., 2022).

Yüksek doğruluk, hızlı eğitim süreleri ve geniş uygulanabilirlik alanı, XGBoost'u tercih edilen bir algoritma haline getirmiştir. Temeldeki ağaç tabanlı öğrenme yöntemleri ve gradient boosting prensibi, karmaşık ilişkileri anlama ve yüksek doğrulukla tahminlerde bulunma kapasitesini artırır. Yapılandırılmış değişkenler ve çapraz doğrulama ile birleştirildiğinde, XGBoost, klinik uygulamalarda etkili ve güvenilir sonuçlara ulaşmak için kullanılan önemli bir algoritmadır (Alabdullah vd., 2022; Mitchell vd., 2018).

1.3.1.5. Hafif Gradyan Arttırma Makinesi (Light Gradient Boosting Machine- LightGBM)

Hafif gradyan arttırma makinesi (LightGBM), klinik tahmin modellerinde yaygın olarak kullanılan ve yüksek performansa sahip bir makine öğrenimi algoritmasıdır. LightGBM, karar ağacı öğrenimi için eğitim sürecini hızlandırabilen ve hesaplama karmaşıklığını azaltabilen histogram tabanlı algoritmalar kullanır. Düşük bellek kullanımı ve histogram tabanlı özellikleri nedeniyle hızlı çalışır, özellikle büyük veri setlerinde etkilidir (Rufo vd., 2021; Şahin, 2022; Alabdullah vd., 2022). Düşük bellek kullanımı ve histogram tabanlı özellikleri sayesinde XGBoost'a kıyasla daha hızlı olabilir.

LightGBM, ağaç tabanlı öğrenme yöntemlerine dayanır. Ancak, diğer geleneksel ağaç tabanlı yöntemlere göre daha hızlı ve daha verimli çalışma özelliği taşır. Bu algoritma temelde ağaç tabanlı bir öğrenme yöntemi olan gradient boosting üzerine kurulmuştur. LightGBM, öğrenme sürecini örnek grupları üzerinde gerçekleştirir. Büyük veri setlerini küçük gruplara böler ve her bir grup üzerinde ağaç oluşturarak paralel hesaplamalar yapar. Bu durum, hızlı öğrenme süreleri sağlar. Bu özellik, büyük veri setlerinde verimli bir şekilde çalışabilme yeteneği sağlar. Bu yöntem, modelin önceki ağaçlardan kaynaklanan hataları öğrenmesine ve bu hataları düzeltmesine imkan sağlar. Her ağaç, öncekilerin hatalarını düzeltmeye odaklanarak, topluca daha güçlü bir model oluşturur. Öğrenme süreci sırasında belirli bir öğrenme oranı kullanılır. Bu oran, her ağacın model üzerindeki etkisini düzenler (Rufo vd., 2021). Ayrıca, LightGBM'nin düzenleme değişkenleri kullanılarak aşırı öğrenmeyi önlemek için düzenleme (regularizasyon) uygulanabilir. Optimal öğrenme oranı

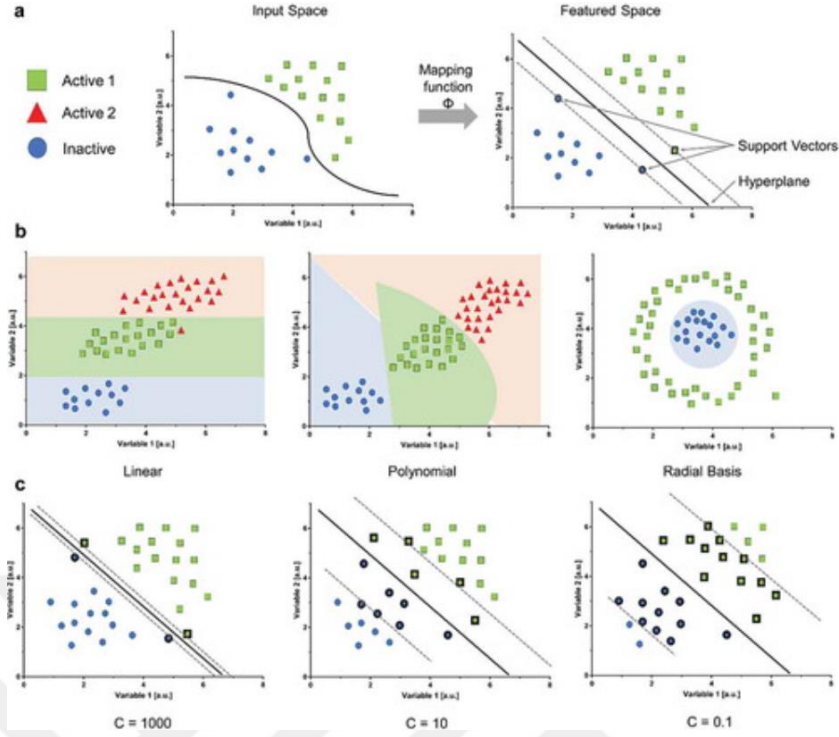
ve düzenleme, modelin genel doğruluğunu artırabilir (Rufo vd., 2021; Dong, 2022). LightGBM, genellikle çapraz doğrulama yöntemleriyle birlikte kullanılır. Bu, modelin farklı veri setlerindeki performansını değerlendirir ve en iyi modelin seçilmesine yardımcı olur. Model, yeni verilere hızlı bir şekilde tahminler yapabilir (Alabdullah vd., 2022).

LightGBM, klinik tahmin modellerinde yüksek doğruluk ve hızlı eğitim süreleri sunan güçlü bir araçtır. Light Gradient Boosting Machine (LightGBM), klinik tahmin modellerindeki ihtiyaçları karşılamak üzere tasarlanmış etkili ve hızlı bir makine öğrenimi algoritmasıdır (Alabdullah vd., 2022; Rufo vd., 2021).

1.3.1.6. Destek Vektör Makineleri-DVM (Support Vector Machines - SVM)

Destek Vektör Makineleri, klinik tahmin modellerinde kullanılan güçlü bir makine öğrenme algoritmasıdır. Temel olarak, DVM, verileri iki veya daha fazla sınıfa ayırmak için kullanılır (Alanazi, 2017). DVM, veri noktalarını sınıflandırma hiperdüzlemiyle ayırarak çalışır. Bu algoritma, farklı sınıflar arasında optimum bir ayırım yapmaya çalışırken, margin (marginal) adı verilen bir kavramı maksimize etmeye odaklanır (Mythili vd., 2013; Shouval vd., 2014).

DVM algoritmasında ilk adım, klinik verilerin toplanması ve gerektiğinde ön işlemedir. Bu veriler, hastaların klinik özellikleri, laboratuvar sonuçları veya diğer ilgili veriler olabilir. Verileri analiz etmek için uygun özellikler seçilmelidir. Bu, gereksiz veya düşük bilgi içeren özelliklerin çıkarılması ve modelin daha iyi performans göstermesini sağlamak için önemlidir. Veri kümesi eğitim ve test alt kümelerine bölünür. Eğitim verileri, DVM modelini oluşturmak için kullanılır. Test verileri, modelin performansını değerlendirmek için kullanılır. DVM'nin temel amacı, verileri iki veya daha fazla sınıfa ayırmaktır. Bu iki sınıf arasında bir karar sınırı oluşturur. Bu karar sınırı, veri noktalarını sınıflandırırken kullanılır. DVM, en iyi karar sınırını bulmaya çalışır. Bu sınır, iki sınıf arasındaki mesafeyi maksimize ederken hata payını minimize etmeyi amaçlar (Mythili vd., 2013). DVM, bu sınırı "destek vektörleri" adı verilen belirli veri noktalarının etrafında oluşturur. Bu destek vektörleri, sınıflandırma işleminde önemli bir rol oynar. DVM, doğrusal ve doğrusal olmayan çekirdek fonksiyonları kullanabilir. Doğrusal bir çekirdek, verileri doğrusal bir düzlemde ayırırken, doğrusal olmayan bir çekirdek daha karmaşık sınırları tanımlamak için kullanılır (Shouval vd., 2014).



Şekil 1.9. Destek vektör makinelerine ilişkin grafikler (Maltarollo vd., 2019)

Eğitim verileri kullanılarak DVM modeli eğitilir. Bu, en uygun karar sınırını ve destek vektörlerini bulmayı içerir. Eğitilen model, test verileri üzerinde değerlendirilir. Bu, modelin ne kadar başarılı olduğunu anlaşılmasını sağlar. DVM'nin bazı parametreleri vardır, örneğin, C parametresi (C değeri, modelin aşırı uyum veya düşük uyum arasındaki dengeyi kontrol eder) ve çekirdek fonksiyonu seçimi gibi. Bu parametreleri dikkatli bir şekilde ayarlayarak modelin optimize edilmesi sağlanabilir (İslam vd., 2020). Klinik tahmin modellerinde kullanılabilir ve iki veya daha fazla sınıfı olan problemleri ele almak için uyarlanabilir (Shouval vd., 2014).

1.4. Diğer Algoritmalar

1.4.1. Bayesci Genelleştirilmiş Doğrusal Model (Bayesian Generalized Linear Model)

Bayesci genelleştirilmiş doğrusal model, klinik tahmin modelleme alanında öne çıkan bir bayesian istatistik yaklaşımdır. Bu model, genelleştirilmiş doğrusal modellerin temel

prensiplerini benimsemekte ve bu prensipleri bayesian çerçeve içinde uygulamaktadır (Kapelner ve Bleich, 2013). Hastalık tanısında, risk faktörlerinin belirlenmesinde ve veri setlerindeki gizli ilişkilerin ortaya çıkarılmasında kullanılır. Regresyon modellerini genişletmek için kullanılır.

Bayesci istatistik, belirsizlikleri ve değişkenlerle ilgili bilgi eksikliklerini ele almak için güçlü bir araçtır. Bu, belirsizlikleri ele almak ve değişkenlerle ilgili olasılık dağılımlarını belirlemek amacıyla Bayesci istatistiklerini kullanır. Değişkenlerin belirli değerler yerine olasılık dağılımları olarak ifade edilmesi, belirsizlikleri daha iyi yönetmeye olanak tanır (Kapelner ve Bleich 2013). Bayesci genelleştirilmiş doğrusal modelin temel adımlarından biri önceki (prior) dağılımların belirlenmesidir. Bu dağılımlar, veri gözlemlenmeden önce değişkenlerle ilgili mevcut bilgileri temsil eder. Araştırmacı, önceki çalışmalardan gelen bilgileri ve uzman görüşlerini kullanarak bu dağılımları belirler. Veri gözlemlendikten sonra, önceki dağılımlar ve yeni veri birleştirilerek güncellenmiş olasılık dağılımları elde edilir. Bu adım, Bayes teoremi kullanılarak gerçekleştirilir. Güncellenmiş dağılımlar, değişkenlerin olasılık dağılımlarını ve bu değişkenlerle ilgili güncellenmiş bilgileri içerir. Güncellenmiş olasılık dağılımlarına sonsal dağılım denir (Grudniewicz ve Ślepaczuk, 2023). Sonsal dağılım, belirli değişken değerleri için olasılık yoğunluğunu belirtir ve modelin değişkenlerine dair güncellenmiş bilgileri sağlar. Bayesci istatistikte, sonsal dağılımlardan örneklem almak için örnekleme yöntemleri kullanılır. Markov Chain Monte Carlo (MCMC) gibi yöntemler, bu dağılımlardan rastgele örnekler elde etmek için kullanılır (Wang vd., 2021; Grudniewicz ve Ślepaczuk, 2023). Bayesci genelleştirilmiş doğrusal model, model seçimi ve değerlendirmesi için bayesci model seçimi kriterleri veya modelin tahmin gücünü değerlendiren diğer metrikler kullanır. Bu, farklı modeller arasında karşılaştırma yapmayı sağlar. En son olarak, modelin öğrenilmiş değişkenleri kullanarak tahmin yapılır ve çıkarımlarda bulunulur. Bu durum, klinik uygulamalarda belirli bir durumun olasılığını tahmin etmek gibi önemli uygulamalar içerir (Kapelner ve Bleich 2013).

Bayesci genelleştirilmiş doğrusal model, belirli bir klinik sorun veya hedefe yönelik olarak uygun bağlantı fonksiyonları ve önceki dağılımlar kullanılarak özelleştirilebilir. Bayesci yaklaşım, belirsizliklerin ve bilgi eksikliklerinin daha etkili bir şekilde yönetilmesine olanak tanır ve klinik uygulamalarda güçlü bir araç olabilir (Kapelner ve Bleich 2013).

1.4.2. Bilgi Kazancı (Information Gain)

Bilgi Kazancı karar ağaçları gibi algoritmaların özellik seçimi süreçlerinde kullanılır. Klinik tahmin modelleri, hastalıkların teşhisi, prognozu veya tedaviye yanıt gibi klinik senaryoların öngörüsünde kullanılan önemli araçlardır. Bilgi kazancı, klinik verilerin analizi için kullanılan bir özellik seçme algoritmasıdır. Temel amacı, veri kümesindeki değişkenlerin, hedef değişkenle ilişkisini değerlendirerek en önemli özellikleri belirlemektir (Jadhav vd., 2018; Win ve Kham 2019).

Algoritmanın işleyişi, öncelikle klinik verilerin toplanması ve düzenlenmesi ile başlar. Bu adım, hastaların demografik bilgileri, laboratuvar sonuçları, semptomatik göstergeler ve diğer ilgili veri noktalarını içerir. Bilgi kazancının temeli, entropi kavramına dayanır. Entropi, bir sistemdeki belirsizliği ölçen bir terimdir. Algoritma, her özellik için entropiyi hesaplar ve bu sayede hangi özelliğin daha fazla bilgi kazandırabileceğini belirler. Her özellik için entropi hesaplandıktan sonra, bilgi kazancı hesaplanır (Omuya vd., 2021; Win ve Kham 2019). Bilgi kazancı, bir özelliğin eklenmesi veya çıkarılmasının entropideki değişikliği ifade eder. Yüksek bilgi kazancına sahip olan özellikler, modelin daha etkili bir şekilde öğrenmesini sağlar. Bilgi kazancı, genellikle karar ağaçları ile birleştirilir. Algoritma, en yüksek bilgi kazancına sahip özelliği seçer ve bu özellik üzerinden karar ağacını oluşturur. Bu süreç, ağacın düğümlerini ve dallarını belirler. Oluşturulan karar ağacı, klinik verileri üzerinde eğitilir (Jadhav vd., 2018; Lee ve Lee, 2006). Model, hastaların belirli durumları hakkında tahminler yapabilir. Eğitilen model, klinik uygulamalarda kullanılarak belirli durumların tahmin edilmesine yardımcı olabilir. Modelin çıktıları, sağlık profesyonelleri tarafından yorumlanarak klinik kararlar alınabilir. Klinik veriler sürekli olarak güncellendiğinden, Bilgi kazancı algoritması sürekli olarak yeniden eğitilebilir. Yeni verilerle güncellenen model, zamanla daha hassas ve güvenilir hale gelir.

Sonuç olarak, Bilgi kazancı algoritması, klinik tahmin modellerinde veri analizi ve özellik seçimi için güçlü bir araçtır. Bilgi kazancı prensibi, modelin en önemli özellikleri öğrenmesine ve klinik kararlara daha fazla katkıda bulunmasına olanak tanır.

1.4.3. L1 Düzenlemeli Modeller (L1 Regularization Models- Lasso)

L1 düzenlemeli modeller, istatistik ve makine öğrenimi alanlarında kullanılan bir regresyon tekniğidir. Lasso, "Least Absolute Shrinkage and Selection Operator" ifadesinin kısaltmasıdır. Bu algoritma, regresyon analizlerindeki aşırı öğrenmeyi önlemek, modelin daha genelleştirilebilir ve anlamlı olmasını sağlamak amacıyla kullanılır (Ayna Altuntaş, 2022; Muthukrishnan ve Rohini, 2016).

Algoritmanın ilk adımı, klinik verilerin toplanması ve düzenlenmesidir. Bu veriler genellikle hastaların demografik bilgilerini, laboratuvar sonuçlarını, semptomatik göstergeleri ve diğer ilgili bilgileri içerir. Lasso, özellik seçimi yaparak gereksiz veya etkisiz özellikleri filtreler. Bu işlem sırasında, her bir özelliğin katsayısı üzerine bir kısıtlama getirilir. L1 regularization, katsayıları sıfıra yaklaştırarak bazı özellikleri tamamen ortadan kaldırır. L1 Regularization'ın etkinliği, "alpha" adı verilen bir hiperparametre ile kontrol edilir. "Alpha" değerinin artması, katsayıların sıfıra daha yakın hale gelmesini sağlar. Hiperparametre ayarı, modelin genelleme yeteneğini ve özellik seçimini etkiler (Ayna Altuntaş, 2022; Vidaurre vd., 2013; Muthukrishnan ve Rohini, 2016; Fernández-Delgado, 2019). Veri seti, Lasso algoritması ile eğitilir. Bu süreçte, regresyon katsayıları belirli bir optimizasyon algoritması kullanılarak ayarlanır. Lasso, ağırlıkları sıfıra çekerek gereksiz özellikleri ortadan kaldırdığı için, daha basit ve anlamlı bir model elde edilir (Ayna Altuntaş, 2022).

Modelin performansı iyileştirilebilirse, hiperparametreler ve özelliklerin seçimi gözden geçirilerek optimizasyon yapılır. Eğitilen Lasso modeli klinik uygulamalarda kullanılabilir. Modelin çıktıları, sağlık profesyonelleri tarafından yorumlanarak klinik kararlar alınabilir (Schmidt, 2005).

Klinik veriler sürekli olarak güncellendiği için, Lasso modeli periyodik olarak yeniden eğitilmelidir. Yeni verilerle güncellenen model, zamanla daha doğru ve güvenilir hale gelir. Lasso'nun klinik tahmin modellerinde kullanılması, gereksiz karmaşıklığı azaltarak ve modelin genelleştirilebilirliğini artırarak etkili bir şekilde bilgi çıkarmayı sağlar (Vidaurre vd., 2013).

2. GEREÇ ve YÖNTEM

2.1. Benzetim verisi

Tez çalışmasının ilk bölümünde benzetim verilerinin türetiminde Monte Carlo benzetim yöntemi kullanılmıştır. Benzetim ve uygulamalarda LR, ML ve diğer algoritmalar (kütüphaneler "caret", "dplyr", "gbm", "xgboost", "lightgbm", "randomForest", "rpart", "glmnet", "CORElearn" "e1071") kullanılmıştır. Ayrıca model performansları için "pROC" kütüphanesi kullanılmıştır.

Tez çalışmamızda uygulamalarda en sık kullanılan makine öğrenme algoritmaları ve diğer algoritmalar iki düzeyli sonuç değişkenin tahmini için uygun görülen ve skorları elde edebilmek için bağımsız değişkenlerin önem sırasını/ağırlığını veren 9 algoritma kullanılmıştır. Bunlar; "gbm", "xgboost", "lgb.Dataset", "randomForest", "rpart", "svm" "glmnet", "attrEval" ve "bayesglm" dir. Bu algoritmalar Çizelge 2.1.1'de sunulmuştur.

Çizelge 2.1. ML algoritmaları ve diğer algoritmalar

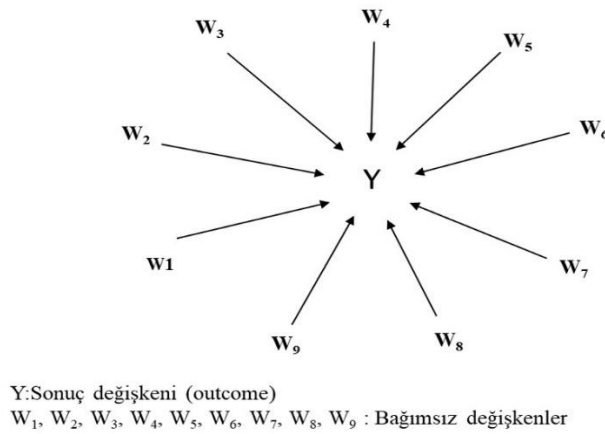
No	Tahminci	Model	R paket
1	"gbm"	Genelleştirilmiş Boosted Regresyon Modeli (Generalized Boosted Regression Models)	gbm
2	"xgboost"	Aşırı gradyan artırma (eXtreme Gradient Boosting)	xgboost
3	"lgb.Dataset"	Hafif gradyan artırma makineleri (Light Gradient Boosting Machine)	lightgbm
4	"randomForest"	Rastgele ormanlar (Random Forest)	randomForest
5	"rpart"	Karar ağaçları (Decision Tree)	rpart
6	"svm"	Destek vektör makineleri (Support vector machine)	e1071
7	"glmnet"	Cezalandırılmış, genelleştirilmiş doğrusal modeller (L1 Regularization Models)	glmnet
8	"attrEval"	Bilgi Kazancı (Information Gain)	CORElearn
9	"bayesglm"	Bayesci genelleştirilmiş doğrusal model (Bayesian generalized linear model (glm))	arm

Tüm istatistiksel analizlerde IBM SPSS (Version 22, SPSS Inc., Chicago, IL, USA), R studio (Versiyon 3.5.0) paket programları ve R programlama dili (Versiyon 3.5.0) kullanılmıştır.

Tez çalışmamızda, iki düzeyli sonuç değişkenine sahip klinik tahmin modellerinde risk faktörlerinin skor değerlerini belirlemek amacıyla, benzetim ile türetilen bir veri seti ve ayrıca gerçek bir veri seti kullanılarak lojistik regresyon analizi gerçekleştirilmiştir. Çalışmamızda iki farklı senaryo kullanılmıştır. Hangi senaryoda hangi klinik tahmin modelinin daha başarılı olduğu sonucuna ulaşılmıştır. Farklı senaryolar için 10000 benzetim yapılmıştır.

Benzetim verisi lojistik regresyona uygun olacak şekilde türetilmiştir. Veri setinde bir kategorik sonuç değişkeni (Y), 9 bağımsız değişken $W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9$ bulunmaktadır. Bağımsız değişkenler $W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9$ ise 0 ve 1'lerden oluşan iki düzeyli kategorik değişken olarak türetilmiştir. Bağımsız değişkenler, birbirinden bağımsız olacak şekilde binom dağılımından türetilmiştir. Değişkenler arasında korelasyon bulunmamaktadır.

Sonuç değişkeni, bağımsız değişkenlerin bir fonksiyonu olacak şekilde iki düzeyli (0,1) kategorik bir değişken olarak lojit model $P(Y=1 | W_i) = 1 / (1 + \exp(-\beta \cdot f(W_i)))$ kullanılarak türetilmiştir. Bağımlı ve bağımsız değişkenlerin türetilmesine ilişkin kodlar Ek 1 de verilmiştir.



Şekil 2.1. Benzetim verisindeki bağımlı ve bağımsız değişkenler

Veri türetiminin ilk aşamasında lojistik regresyon katsayılarının yaklaşık olarak tamamının anlamlı olmasına dikkat edilmiştir. Bunun nedeni genellikle en son elde edilen klinik tahmin modellerinde istatistiksel olarak anlamlı olmayan ancak klinik olarak anlamlı olduğu için modele alınan değişkenler dışındaki tüm değişkenler istatistiksel olarak

anlamlıdır. İkinci aşmasında beta katsayılarının tümünün pozitif olduğu durumlar ele alınmıştır. Beta katsayıları negatif olduğu durumda kodlama tersine çevrilerek (referans kategori değiştirilerek) analizlere devam edilebilir. Bu nedenle betaların negatif olduğu durumlar senaryolara dahil edilmemiştir. Üçüncü aşamasında ise değişkenlere ilişkin odds oranlarının dağılımının model başarısı üzerinde etkili olabileceğini düşündüğümüz için odds oranlarındaki dağılımın dengeli ve dengesiz olduğu durumlar senaryolara dahil edilmiştir. Odds oranları arasındaki dengesizlik, modelin bazı değişkenlere daha fazla ağırlık verip diğerlerini ihmal etmesine neden olabilir. Bu durum, modelin genel performansını etkileyebilir. Odds oranlarının dengesiz olması için bir veya iki değişkene ilişkin odds oranları diğer odds oranlarından çok daha yüksek olduğu durumlar odds oranlarının dengesiz olduğu durumlar olarak adlandırılır. Çalışmamızda 9 bağımsız değişkenin bulunduğu lojistik regresyon modelinde örneğin Şekil 2.2’de verilen A senaryosunda değişkenlere ilişkin odds oranları yaklaşık olarak 3,57, 2,08, 1,76, 1,93, 2,71, 2,05, 3,88, 3,03 ve 1,59 bulunmuş ve bu model dengeli dağılıma örnek oluşturmuştur. Şekil 2.3’de verilen B senaryosuna ilişkin örnekte ise odds oranları 15,3, 3,98, 4,42, 3,41, 3,26, 2,24, 2,64, 1,62 ve 1,71 olarak bulunmuş ve bu model odds oranlarının dengesiz olduğu dağılıma örnek oluşturmuştur. Açıklanan durumlara göre A ve B olmak üzere iki farklı senaryo oluşturulmuştur. Senaryolar aşağıda verilmiştir.

A: beta katsayıları pozitif odds oranı dengeli tasarım

B: beta katsayıları pozitif odds oranı dengesiz tasarım

Bu tezde 3 farklı örneklem büyüklüğü kullanılmıştır. Her bir senaryo için örneklem büyüklüğü küçük ($n=200$), orta ($n=500$) ve büyük ($n=1000$) olarak alınmıştır.

Bu çalışmada sadece iki sonuçlu bağımsız değişkenlerden oluşan model kullanılmıştır. Sürekli ve nominal değişkenlerin birlikte bulunduğu modeller kullanılmamıştır. Sürekli değişkenlerin kullanılmama nedeni skorum sisteminde çoğunlukla sürekli değişkenler kategorilere ayrılarak modele alınmaktadır.

Bu tezde her bir senaryoda bağımlı değişkenlerde ilgilenilen olayın görülme/görülme oranı (1-0 oranı) %20-%80 ve %50-%50 olarak ele alınmıştır. Bağımsız değişkenlerde ise etkenin görülme/görülme oranı (1-0 oranı) %20-%80, %50-%50 ve rastgele olarak belirlenmiştir. Rastgele belirlenen bağımsız değişkenlerde 1 lerin oranı 0,1-

0,33; 0,34-0,66 ve 0,67-0,9 aralıklarında olacak şekilde rastgele olarak belirlenmiştir. Her bir aralıkta üçer bağımsız değişken bulunmaktadır.

Senaryolar oluşturulduktan sonra veri türetimine başlanmıştır. Lojistik regresyona uygun benzetim verisi türetiminin detaylarına bakıldığında ilk olarak beta katsayıları belirlenir. Bu katsayılar gerçek klinik tahmin modelleri ile benzerlik gösterecek şekilde belirlenmelidir. Beta katsayıları modeldeki değişkenlerin anlamlılığını, odds oranlarının dağılımını ve bağımlı değişkene ilişkin 0-1 oranlarının dağılımını etkilemektedir. Beta katsayıları belirlenirken bu durumlar gözönünde bulundurulmuştur. Daha sonra binom dağılımına uygun olarak bağımsız değişkenler türetilmiştir. Bağımsız değişkenlere ilişkin 1 oranlarının dağılımının yaklaşık olarak %20, %50 ve rastgele (0,1-0,33 aralığında üç, 0,34-0,66 aralığında üç ve 0,67-0,9 aralığında üç olmak üzere toplam 9 bağımsız değişken rastgele belirlenmiştir) olduğu bu aşamada belirlenmiştir. Sonraki adımda veriler bağımsız değişkenler arasında ilişki olmayacak şekilde lojistik regresyondan türetilmiştir. Sonrasında ise örneklem büyüklüğü belirlenmiştir. Veri türetimine ilişkin R kodları Ek 1’de sunulmuştur.

A ve B senaryolarına ilişkin birer örnek aşağıda verilmiştir.

A: beta katsayıları pozitif odds dengeli tasarım

```
0 1
253 247
> yuzde <- prop.table(table(simdata$z.a)) * 100
>
> # Yüzde dağılımını yazdır
> print(yuzde)

0 1
50.6 49.4

Call:
glm(formula = z.a ~ w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8 + w9,
     family = "binomial", data = simdata)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.0276      0.4142  -9.724 < 2e-16 ***
w1             1.2734      0.2195   5.801 6.61e-09 ***
w2             0.7364      0.2172   3.391 0.000698 ***
w3             0.5657      0.2157   2.623 0.008726 **
w4             0.6598      0.2160   3.055 0.002247 **
w5             1.0003      0.2180   4.589 4.45e-06 ***
w6             0.7183      0.2142   3.354 0.000796 ***
w7             1.3571      0.2202   6.164 7.10e-10 ***
w8             1.1103      0.2170   5.118 3.09e-07 ***
w9             0.4681      0.2141   2.187 0.028763 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> skor
(Intercept)    w1      w2      w3      w4      w5      w6      w7      w8      w9
0.01781646 3.57288999 2.08838720 1.76075929 1.93448754 2.71919181 2.05090684 3.88509986 3.03540374 1.59690933
```

Şekil 2.2. A senaryosuna ilişkin örnek sonuçlar

B: beta katsayıları pozitif odds dengesi tasarım

```
0 1
256 244
> yuzde <- prop.table(table(simdata$z.a)) * 100
>
> # Yüzde dağılımını yazdır
> print(yuzde)

0 1
51.2 48.8

Call:
glm(formula = z.a ~ w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8 + w9,
    family = "binomial", data = simdata)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.4243      0.5317  -10.202  < 2e-16 ***
w1             2.7316      0.2745   9.951  < 2e-16 ***
w2             1.3838      0.2523   5.486  4.12e-08 ***
w3             1.4884      0.2501   5.952  2.65e-09 ***
w4             1.2269      0.2484   4.940  7.82e-07 ***
w5             1.1840      0.2463   4.808  1.53e-06 ***
w6             0.8066      0.2433   3.315  0.000916 ***
w7             0.9741      0.2406   4.048  5.16e-05 ***
w8             0.4828      0.2381   2.027  0.042634 *
w9             0.5422      0.2381   2.277  0.022758 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> skor
(Intercept)  w1      w2      w3      w4      w5      w6      w7      w8
0.004408103 15.357014420 3.989945515 4.429783612 3.410520920 3.267439991 2.240358669 2.648906656 1.620542497
w9
1.719834964
```

Şekil 2.3. B senaryosuna ilişkin örnek sonuçlar

Veri türetimi tamamlandıktan sonra LR, ML ve diğer algoritmalara dayalı skorlar belirlenmiştir ve tüm bireylerin toplam skorları hesaplanmıştır. Toplam skorlar üzerinden ROC analizi yapılarak kesim noktası belirlenmiştir. Bu kesim noktası kullanılarak AUC, duyarlılık, seçicilik, doğruluk, pozitif ve negatif tahmin değerleri, F1 skoru, MCC ölçütleri hesaplanmıştır.

Benzetimle oluşturulan veri seti ve gerçek veri seti üzerinde yapılan lojistik regresyon analizi, β katsayıları, odds oranları ve bu değerlerin tamsayıya dönüştürülmesiyle elde edilen ağırlıklı skorlar ile makine öğrenme algoritmaları ve diğer algoritmalar kullanılarak ağırlıklandırılmış skorlar temelinde toplam 468 farklı klinik tahmin modelinin performansı değerlendirilmiştir. Bu modellere ilişkin model performansları AUC, duyarlılık, seçicilik, doğruluk, pozitif ve negatif tahmin değerleri, F1 skoru, MCC ölçütleri ile değerlendirilmiştir.

2.2. Gerçek Veri Seti Üzerinde Uygulama

Malnütrisyon, vücut kütlesi ve fonksiyonlarında ölçülebilir olumsuz etkilere yol açan beslenme bozukluğu olarak tanımlanmaktadır; bu durum, protein, enerji ve diğer besin öğelerinin eksikliği veya fazlalığına bağlı olarak ortaya çıkmaktadır. Beslenme bozukluğu, üç aydan daha kısa bir süre içinde ortaya çıkarsa, bu duruma “akut malnütrisyon” denir ve genellikle patolojik bir sebep, besin alımında veya günlük beslenme kalitesinde ani bir düşüşle ilişkilidir. Çocukların beslenme bozukluğunun etkisi altında üç aydan daha uzun süre kalması durumu ise “kronik malnütrisyon” olarak adlandırılır ve genellikle kalıcı büyüme-gelişme geriliği ile bağlantılıdır. Orta ve ağır dereceli akut malnütrisyon, özellikle düşük ve orta gelirli ülkelerde önemli bir halk sağlığı sorunudur. Malnütrisyon, dünya genelinde yaklaşık 50 milyon çocuğu etkilemektedir.

Malnütrisyon tanısı antropometrik ölçümler ve hesaplamalara dayalı olarak konulmakla birlikte tanı konusunda kesin bir uzlaşma sağlanamamıştır. Yaşa göre boy, yaşa göre ağırlık, boya göre ağırlık, vücut kütle indeksi ve sol üst orta kol çevresi ölçümü ve z skorları değerlendirilerek malnütrisyonun türü ve ağırlığı değerlendirilebilmektedir. Son yıllarda malnütrisyon tanısında “sol üst orta kol çevresi” ölçümünün daha sık önerildiği gözlenmekle birlikte bu ölçümün farklı gözlemciler tarafından tutarlı bir biçimde ölçülmesi doğru tanı için oldukça önemlidir. Malnütrisyon tanısında kullanılabilecek basit ve pratik bir skorlamanın olmaması ve birden çok değişkenin olması tanıyı zorlaştırmaktadır. Tedavide gecikmenin önlenmesi için en önemli faktör ise ebeveynlerin çocuklarını yaşitlarına göre boy ve kilo açısından karşılaştırarak erken dönemde malnütrisyon şüphesiyle hastaneye başvurmalarıdır.

Çalışmamızda ebeveynlerin çocuklarında malnütrisyon şüphesi ile hastaneye başvuru konusunda karar verebilmesi için çocuğun yaşitlarına göre boy ve kilosu, kendilerine ilişkin (anne/baba) boy ve kilo, sosyo-demografik özellikleri, çocuğun beslenme alışkanlığı ve beslenme davranışı ve diğer etkili faktörleri kullanarak basit bir klinik tahmin modeli oluşturulması amaçlanmıştır.

Bu çalışma için, Hitit Üniversitesi Klinik Araştırmalar Etik Kurulundan onay alınmıştır (Kayıt numarası: 2022-64, 25.08.2022).

Klinik tahmin modeli oluşturmak amacıyla 26 Ekim 2022 ile 19 Nisan 2023 tarihleri arasında Hitit Üniversitesi Çorum Erol Olçok Eğitim ve Araştırma Hastanesi pediatri bölümüne başvuran 2-10 yaş arası toplam 522 çocuk ile çalışılmıştır. 10 yaş üzeri olan, puberte evresinde olan, göçmen statüsünde olan, sendromik hastalıkları olan (down, turner sendromu vb.), nörolojik hastalığı olan, ödem, asit, alçı protez gibi vücut ağırlığını etkileyecek durumları olan çocuklar çalışma dışı bırakılmıştır.

Çalışmaya dahil edilen çocukların malnütrisyon tanısı kilo z-skoru, VKİ z-skoru ve sol üst orta kol çevresi z-skoru kullanılarak konulmuş olup bunlardan en az birinde z skoru -1 ve altında ise “malnütrisyon” olarak değerlendirilmiştir. Boy, kilo ve sol üst orta kol çevresi ölçümleri tek bir klinisyen tarafından ölçülmüştür. Klinik tahmin modelinde kullanılacak açıklayıcı değişkenler ise ebeveynler ile yapılan anket yolu ile elde edilmiştir. Anketler karekod yolu ile ebeveynlere anında ulaştırılmış ve mobil telefonlarını kullanarak doldurmaları sağlanmıştır. Anket sosyo-demografik verileri (anne babanın eğitim durumları, VKİ, boy, kilo, kardeş sayısı, doğum şekli, doğum ağırlığı, doğum haftası, beslenme özellikleri, kronik hastalık öyküsü), çocuğun yaşıtlarına göre boyu ve kilosu konusunda ebeveyn görüşü, çocuğun beslenme alışkanlığı ve beslenme davranışı ve diğer etkili faktörleri (günlük ekran başında kalma süresi, hareketliliği, uyku düzeni v.b.) içeren sorulardan oluşmaktadır.

Çalışmamızda çocuklarda malnütrisyon tanısını öngören skorlama sistemine dayalı klinik tahmin modelleri lojistik regresyondan elde edilen β katsayıları, odds oranları, bu değerlerin tamsayıya dönüştürülmesi ile elde edilen skorlar, makine öğrenme algoritmaları (rastgele orman, destek vektör makineleri vb.) ve diğer algoritmalar (Bayesci GDM, L1 (Lasso), Bilgi Kazancı) elde edilen skorlar kullanılarak geliştirilmiştir.

Güç analizi sonucunda çalışmaya toplam 522 bireyin alınmasına karar verilmiştir. 522 gözlemin bulunduğu bir örneklemden elde edilen ikili bir sonuç değişkeni (Y) üzerindeki ikili bir bağımsız değişken (X) için yapılan lojistik regresyon, gözlemlerin %40'ının X=0 grubunda ve %60'ının X=1 grubunda olduğu bir durumu içermektedir. Bu regresyon, Y=1 olasılığında meydana gelen bir değişikliği tespit etmek için %5 anlamlılık düzeyinde %80 güç elde etmiştir. Bu değişiklik, temel değeri olan 0,280'den 0,401'e doğru bir olasılık değişimine denk gelmektedir. Bu değişiklik, bir odds oranı ile ifade edildiğinde 1,72'ye karşılık gelir. Lojistik regresyondaki ilgili bağımsız değişkenin diğer bağımsız değişkenlere

olan çoklu regresyonu, elde edilen R-Kare değeri 0,035 olduğundan dolayı bir düzeltme yapılmıştır.

Tahmin modelinde yer alacak risk faktörlerini belirlemek amacıyla ilk aşamada tek değişkenli analizler yapılmıştır. Veri tipine ve parametrik varsayımların sağlanma durumuna bağlı olarak malnütrisyon olan ve olmayanlarında farklılıklar Ki-kare testi, Fisher Exact test, Bağımsız Gruplarda t Testi (Student's t Test) Mann Whitney U testi kullanılarak saptanmıştır. Sürekli değişkenlerde en iyi kesim noktasını belirlemek için ROC analizi ve Youden indeksi (maksimum duyarlılık ve özgüllük) kullanılmıştır. Tek değişkenli istatistiksel analizler sonucunda *P* değeri 0,25'den düşük bulunan risk faktörleri çoklu lojistik regresyon modeline, aday değişken olarak alınmıştır. Çoklu lojistik regresyon analizi sonucunda istatistiksel olarak anlamlı bulunan değişkenler için beta odds oranları (OR) ve %95 güven aralıkları sunulmuştur.

2.2.1. Klinik tahmin modeli oluşturma ve skollama sistemi

Çalışmamızda LR'ye ilişkin 4 farklı klinik tahmin modeli kurulmuştur. Tam sayılı skollama yapabilmek için sürekli değişkenler klinik olarak anlamlı kategorilere dönüştürülmüştür. Veri setinde sürekli olan çocuk, anne ve babanın VKİ, boy, kilo ve yaş değerleri, doğum ağırlığı, anne sütü alma süresi vb. değişkenler kategorik değişkenlere dönüştürülmüştür.

Tek değişkenli lojistik regresyon analizinde anlamlı bulunan değişkenler ile çoklu lojistik regresyon modeli oluşturulmuştur. Analiz sonucunda, yaşlarına göre kilosu, yaşlarına göre boyu, ailede yaşayan toplam kişi sayısı ve paketli gıda tüketim sıklığı değişkenleri ile son model oluşturulmuştur. 1. Klinik tahmin modeli için çoklu lojistik regresyon modelinden elde edilen β katsayıları alınarak skorlar elde edildi. 2. klinik tahmin modeli için oluşturulan modelin β katsayıları en yakın tamsayıya yuvarlanarak yeni bir model kurulmuştur. 3. klinik tahmin modelinde çoklu lojistik regresyon modelinden hesaplanan odds oranları alınarak elde edilen skorlar ile model kurulmuştur. 4. klinik tahmin modelinde çoklu lojistik regresyon modelinden hesaplanan odds oranları en yakın tamsayıya yuvarlanarak elde edilen skorlar ile model kurulmuştur.

2.2.2. Makine öğrenme yöntemi

Yapay öğrenme denetimli öğrenme, denetimsiz öğrenme ve kuvvetlendirmeli öğrenme olmak üzere üç kategori altında gruplandırılmaktadır.

Denetimli öğrenme yönteminde, bir modelin etiketli (labelled) veri setlerini kullanarak öğrenmesi hedeflenir. Model, doğru çıktıları içeren bu etiketleri kullanarak giriş verileriyle ilgili ilişkileri öğrenir ve daha sonra yeni giriş verilerine dayanarak doğru çıktılar tahmin edebilir. Sınıflandırma (classification) ve regresyon (regression) gibi görevler, denetimli öğrenme yöntemleriyle çözülebilir. Çalışmamızda hastalardan bilgi toplama yoluyla malnütrisyon tanısını koyabilmek için lojistik regresyon, GBM, XGBoost, LightGBM, rastgele orman, karar ağaçları, bilgi kazancı, L1 düzenlemeli model (Lasso), DVM, bayesci genelleştirilmiş doğrusal model sınıflama algoritmaları kullanılmıştır.

2.2.3. Klinik tahmin modeli

Çocuklardan elde edilen veriler kullanılarak lojistik regresyondan elde edilen β katsayıları, odds oranları, bu değerlerin tamsayıya dönüştürülmesi ile elde edilen skorlar ve makine öğrenme algoritmaları (rastgele orman, destek vektör makineleri, vb.) ile ağırlıklandırılarak elde edilen skorlar aracılığıyla yeni bir klinik tahmin modeli geliştirilmiştir. Klinik tahmin modellerinin performanslarının ölçülmesinde AUC, doğruluk, duyarlılık, seçicilik, PTD, NTD, F1 skoru, ve MCC değerleri kullanılmıştır.

3. BULGULAR

3.1. LR, ML ve diğer algoritmalar ile elde edilen skorlar için benzetim sonuçları

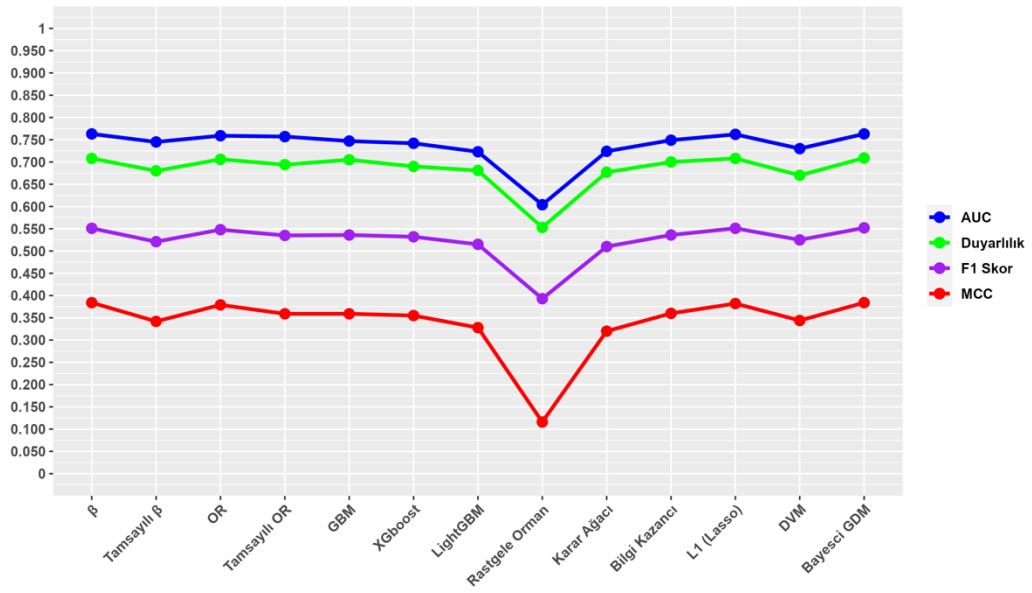
3.1.1. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve %20-%80 (n=200, n=500 ve n=1000)

İki düzeyli sonuç değişkene sahip klinik tahmin modellerinde risk faktörlerinin skor değerlerini belirlemek amacıyla modelde kategorik (nominal) yapıda dokuz adet bağımsız değişkenin yer aldığı A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişkenlerin 1-0 oranının %20-%80 olduğu durumda 3 farklı örneklem büyüklüğünde (n=200, n=500 ve n=1000) türetilen benzetim veri setleri üzerinde lojistik regresyon analizi gerçekleştirilmiştir. Lojistik regresyon analizi kullanılarak benzetim sonucunda üretilen veri seti kullanılarak 10000 benzetim sonucunda skorlara dayalı oluşturulan klinik tahmin modellerinin başarı düzeyleri performans ölçütlerinden, duyarlılık, seçicilik, PTD, NTD, doğruluk, AUC, F1 skorları ve MCC değerleri üzerinden karşılaştırılmıştır.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişken oranının %20-%80 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetilmiştir. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.1'de verilmiştir.

Çizelge 3.1. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,763 (0,685-0,842)	0,708 (0,581-0,835)	0,724 (0,654-0,795)	0,456 (0,345-0,568)	0,885 (0,829-0,941)	0,720 (0,658-0,782)	0,551 (0,439-0,663)	0,384 (0,175-0,592)
Tam sayılı Beta	0,745 (0,667-0,822)	0,680 (0,553-0,808)	0,702 (0,631-0,773)	0,434 (0,323-0,546)	0,874 (0,815-0,933)	0,697 (0,633-0,760)	0,521 (0,409-0,634)	0,342 (0,130-0,555)
OR	0,759 (0,680-0,838)	0,706 (0,579-0,833)	0,720 (0,650-0,791)	0,453 (0,342-0,564)	0,884 (0,827-0,940)	0,717 (0,655-0,779)	0,548 (0,437-0,660)	0,379 (0,171-0,587)
Tam sayılı OR	0,757 (0,679-0,836)	0,694 (0,565-0,822)	0,711 (0,640-0,782)	0,441 (0,330-0,552)	0,878 (0,820-0,936)	0,707 (0,644-0,769)	0,535 (0,424-0,647)	0,359 (0,150-0,568)
GBM	0,747 (0,668-0,826)	0,705 (0,578-0,832)	0,702 (0,630-0,774)	0,436 (0,327-0,546)	0,881 (0,823-0,938)	0,703 (0,640-0,766)	0,536 (0,426-0,646)	0,359 (0,153-0,565)
XGBoost	0,742 (0,660-0,823)	0,690 (0,561-0,819)	0,710 (0,638-0,781)	0,438 (0,328-0,549)	0,877 (0,819-0,935)	0,705 (0,642-0,768)	0,532 (0,421-0,644)	0,355 (0,146-0,564)
LightGBM	0,723 (0,640-0,806)	0,681 (0,551-0,811)	0,690 (0,618-0,763)	0,420 (0,312-0,528)	0,871 (0,811-0,930)	0,688 (0,624-0,752)	0,515 (0,405-0,625)	0,328 (0,120-0,536)
Rastgele Orman	0,604 (0,509-0,700)	0,553 (0,417-0,689)	0,578 (0,503-0,654)	0,312 (0,217-0,406)	0,792 (0,716-0,867)	0,572 (0,506-0,639)	0,393 (0,293-0,493)	0,116 (0,088-0,320)
Karar Ağacı	0,724 (0,642-0,806)	0,677 (0,547-0,807)	0,686 (0,613-0,759)	0,415 (0,307-0,523)	0,868 (0,808-0,929)	0,684 (0,620-0,748)	0,510 (0,401-0,620)	0,320 (0,112-0,529)
Bilgi Kazancı	0,749 (0,669-0,829)	0,700 (0,572-0,828)	0,707 (0,636-0,779)	0,439 (0,329-0,549)	0,880 (0,822-0,937)	0,706 (0,643-0,768)	0,536 (0,425-0,647)	0,360 (0,153-0,568)
L1 (Lasso)	0,762 (0,683-0,840)	0,708 (0,581-0,834)	0,723 (0,652-0,793)	0,456 (0,344-0,568)	0,884 (0,828-0,940)	0,719 (0,657-0,781)	0,551 (0,439-0,663)	0,382 (0,174-0,590)
DVM	0,730 (0,647-0,814)	0,670 (0,539-0,801)	0,715 (0,644-0,786)	0,439 (0,327-0,550)	0,870 (0,811-0,929)	0,704 (0,641-0,767)	0,525 (0,413-0,638)	0,344 (0,132-0,557)
Bayesci GDM	0,763 (0,685-0,842)	0,709 (0,583-0,836)	0,723 (0,652-0,793)	0,457 (0,345-0,569)	0,885 (0,829-0,941)	0,720 (0,658-0,781)	0,552 (0,440-0,664)	0,384 (0,176-0,592)



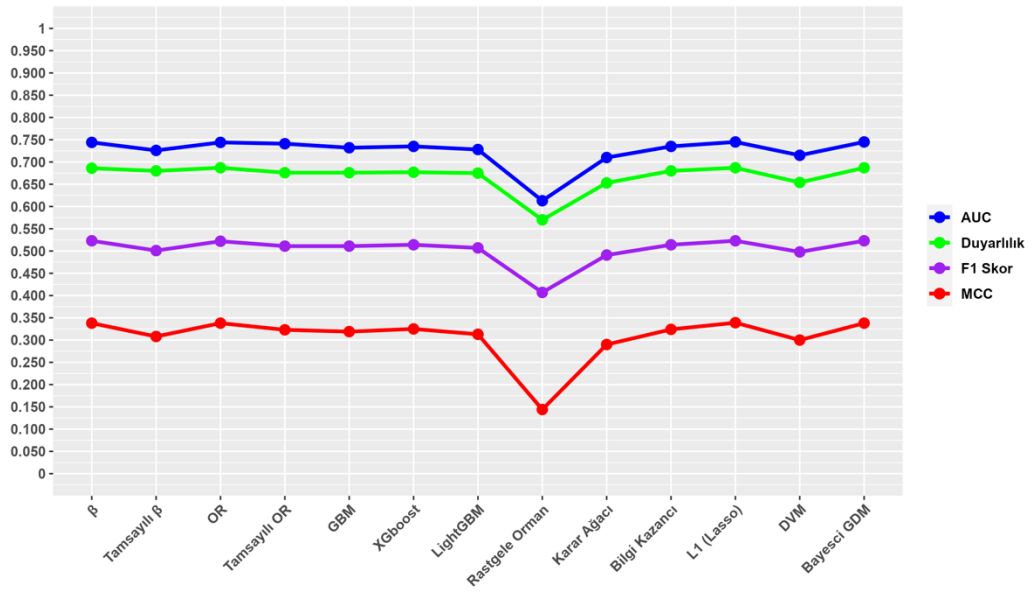
Şekil 3.1. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%20-%80 ve %20-%80)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken oranının %20-%80 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.2’de verilmiştir.

Çizelge 3.2. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,744 (0,694-0,795)	0,686 (0,604-0,768)	0,697 (0,651-0,743)	0,425 (0,356-0,494)	0,873 (0,836-0,911)	0,695 (0,654-0,735)	0,523 (0,453-0,593)	0,338 (0,207-0,469)
Tam sayılı Beta	0,726 (0,677-0,776)	0,680 (0,599-0,761)	0,668 (0,621-0,714)	0,405 (0,338-0,473)	0,869 (0,830-0,908)	0,671 (0,630-0,712)	0,501 (0,432-0,570)	0,308 (0,177-0,440)
OR	0,744 (0,693-0,795)	0,687 (0,605-0,769)	0,696 (0,650-0,742)	0,423 (0,356-0,493)	0,874 (0,836-0,911)	0,694 (0,654-0,734)	0,522 (0,453-0,593)	0,338 (0,206-0,469)
Tam sayılı OR	0,741 (0,690-0,791)	0,676 (0,594-0,759)	0,688 (0,643-0,736)	0,415 (0,347-0,484)	0,869 (0,831-0,907)	0,685 (0,646-0,727)	0,511 (0,442-0,582)	0,323 (0,190-0,455)
GBM	0,732 (0,681-0,783)	0,676 (0,594-0,759)	0,687 (0,640-0,733)	0,413 (0,345-0,481)	0,868 (0,830-0,907)	0,684 (0,644-0,725)	0,511 (0,441-0,580)	0,319 (0,188-0,451)
XGBoost	0,735 (0,683-0,787)	0,677 (0,594-0,760)	0,692 (0,646-0,739)	0,417 (0,348-0,486)	0,870 (0,832-0,908)	0,689 (0,648-0,729)	0,514 (0,445-0,584)	0,325 (0,194-0,457)
LightGBM	0,728 (0,676-0,780)	0,675 (0,592-0,758)	0,681 (0,634-0,728)	0,408 (0,341-0,476)	0,867 (0,828-0,905)	0,679 (0,639-0,720)	0,507 (0,438-0,576)	0,313 (0,182-0,444)
Rastgele Orman	0,613 (0,553-0,672)	0,570 (0,484-0,657)	0,594 (0,545-0,642)	0,320 (0,259-0,381)	0,806 (0,760-0,853)	0,588 (0,546-0,630)	0,407 (0,343-0,472)	0,144 (0,014-0,274)
Karar Ağacı	0,710 (0,657-0,763)	0,653 (0,569-0,738)	0,676 (0,629-0,722)	0,397 (0,329-0,464)	0,858 (0,819-0,898)	0,670 (0,629-0,711)	0,491 (0,422-0,560)	0,290 (0,157-0,422)
Bilgi Kazancı	0,735 (0,684-0,786)	0,680 (0,597-0,762)	0,688 (0,642-0,735)	0,415 (0,347-0,483)	0,870 (0,832-0,908)	0,686 (0,646-0,727)	0,514 (0,444-0,583)	0,324 (0,192-0,455)
L1 (Lasso)	0,745 (0,694-0,796)	0,687 (0,605-0,769)	0,698 (0,652-0,744)	0,425 (0,356-0,494)	0,874 (0,836-0,911)	0,695 (0,655-0,735)	0,523 (0,453-0,593)	0,339 (0,207-0,470)
DVM	0,715 (0,662-0,768)	0,654 (0,570-0,738)	0,684 (0,639-0,732)	0,405 (0,337-0,474)	0,860 (0,821-0,899)	0,678 (0,637-0,719)	0,498 (0,428-0,568)	0,300 (0,167-0,433)
Bayesci GDM	0,745 (0,694-0,796)	0,687 (0,604-0,769)	0,698 (0,651-0,744)	0,425 (0,356-0,494)	0,873 (0,836-0,911)	0,695 (0,655-0,735)	0,523 (0,453-0,593)	0,338 (0,207-0,470)



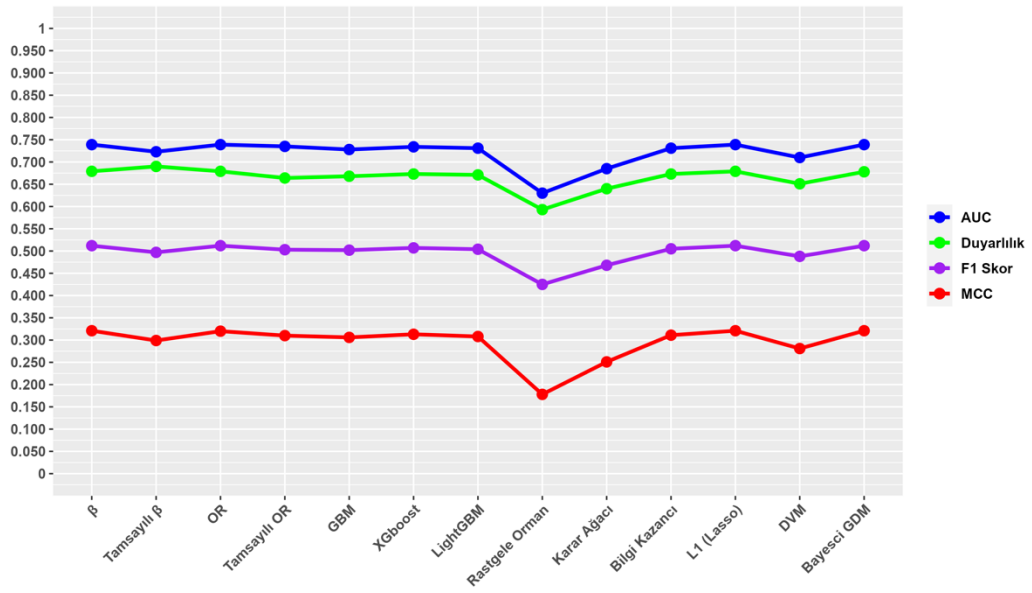
Şekil 3.2 LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%20-%80 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %20-%80 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile üretilen verilere ilişkin klinik tahmin modellerinin performansları 3.3’de verilmiştir.

Çizelge 3.3. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı:%20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,739 (0,703-0,775)	0,679 (0,621-0,738)	0,687 (0,654-0,720)	0,413 (0,360-0,461)	0,869 (0,842-0,896)	0,685 (0,656-0,714)	0,512 (0,463-0,561)	0,321 (0,229-0,414)
Tam sayılı Beta	0,723 (0,688-0,759)	0,690 (0,633-0,746)	0,648 (0,615-0,681)	0,397 (0,350-0,444)	0,869 (0,841-0,897)	0,658 (0,629-0,688)	0,496 (0,448-0,544)	0,299 (0,208-0,391)
OR	0,738 (0,702-0,775)	0,679 (0,620-0,737)	0,686 (0,653-0,719)	0,412 (0,364-0,460)	0,869 (0,842-0,896)	0,684 (0,656-0,713)	0,512 (0,463-0,561)	0,320 (0,228-0,413)
Tam sayılı OR	0,735 (0,699-0,771)	0,664 (0,606-0,723)	0,685 (0,653-0,718)	0,410 (0,361-0,458)	0,865 (0,837-0,892)	0,680 (0,651-0,709)	0,503 (0,454-0,553)	0,310 (0,216-0,404)
GBM	0,727 (0,691-0,764)	0,668 (0,609-0,727)	0,681 (0,648-0,714)	0,404 (0,356-0,452)	0,864 (0,837-0,892)	0,678 (0,649-0,706)	0,502 (0,453-0,551)	0,306 (0,212-0,399)
XGBoost	0,734 (0,697-0,770)	0,673 (0,614-0,731)	0,684 (0,651-0,717)	0,408 (0,360-0,457)	0,867 (0,839-0,894)	0,681 (0,653-0,710)	0,507 (0,458-0,556)	0,313 (0,220-0,406)
LightGBM	0,731 (0,695-0,767)	0,671 (0,612-0,730)	0,680 (0,647-0,713)	0,405 (0,357-0,453)	0,865 (0,838-0,893)	0,678 (0,649-0,707)	0,504 (0,455-0,552)	0,308 (0,215-0,401)
Rastgele Orman	0,630 (0,589-0,672)	0,593 (0,532-0,654)	0,611 (0,577-0,645)	0,334 (0,291-0,379)	0,821 (0,789-0,853)	0,607 (0,577-0,637)	0,426 (0,380-0,472)	0,178 (0,086-0,270)
Karar Ağacı	0,685 (0,647-0,723)	0,640 (0,580-0,700)	0,647 (0,613-0,681)	0,371 (0,325-0,417)	0,848 (0,819-0,878)	0,645 (0,616-0,675)	0,468 (0,420-0,515)	0,251 (0,158-0,343)
Bilgi Kazancı	0,731 (0,695-0,767)	0,673 (0,614-0,732)	0,681 (0,648-0,714)	0,406 (0,358-0,454)	0,866 (0,839-0,894)	0,679 (0,650-0,708)	0,506 (0,457-0,554)	0,311 (0,218-0,403)
L1 (Lasso)	0,739 (0,703-0,775)	0,679 (0,620-0,737)	0,686 (0,654-0,719)	0,413 (0,364-0,461)	0,869 (0,842-0,896)	0,685 (0,656-0,713)	0,512 (0,463-0,561)	0,321 (0,228-0,413)
DVM	0,709 (0,671-0,747)	0,651 (0,591-0,711)	0,670 (0,636-0,703)	0,391 (0,343-0,438)	0,856 (0,828-0,885)	0,665 (0,636-0,694)	0,487 (0,438-0,535)	0,281 (0,188-0,374)
Bayesci GDM	0,739 (0,703-0,775)	0,678 (0,620-0,737)	0,687 (0,655-0,720)	0,413 (0,364-0,461)	0,869 (0,842-0,896)	0,685 (0,656-0,714)	0,512 (0,463-0,561)	0,321 (0,228-0,414)



Şekil 3.3. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%20-%80 ve %20-%80)

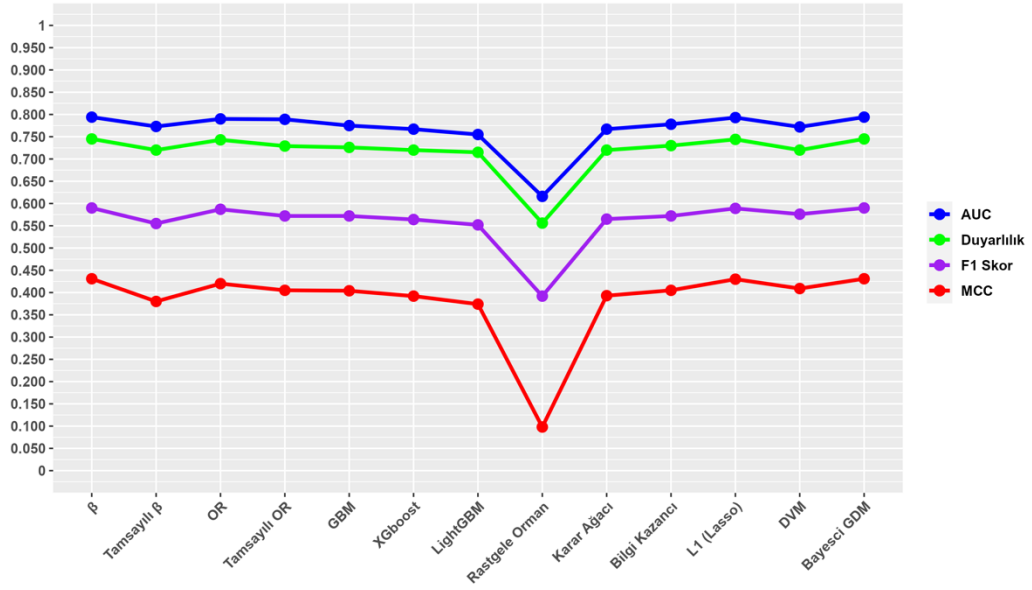
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağacının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.2. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve %50-%50 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişken oranının %50-%50 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetilmiştir. Üretilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara ve ML algoritmalarına dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.4’de verilmiştir.

Çizelge 3.4. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,794 (0,723-0,864)	0,745 (0,625-0,864)	0,736 (0,666-0,806)	0,493 (0,381-0,604)	0,895 (0,841-0,949)	0,738 (0,678-0,799)	0,590 (0,480-0,700)	0,431 (0,229-0,633)
Tam sayılı Beta	0,773 (0,702-0,844)	0,720 (0,600-0,841)	0,703 (0,631-0,774)	0,460 (0,351-0,570)	0,883 (0,826-0,940)	0,707 (0,645-0,770)	0,555 (0,445-0,664)	0,380 (0,176-0,585)
OR	0,790 (0,720-0,861)	0,743 (0,623-0,862)	0,733 (0,663-0,803)	0,489 (0,378-0,601)	0,893 (0,839-0,948)	0,735 (0,675-0,796)	0,587 (0,477-0,696)	0,420 (0,224-0,628)
Tam sayılı OR	0,789 (0,718-0,860)	0,729 (0,608-0,850)	0,723 (0,652-0,794)	0,476 (0,365-0,587)	0,887 (0,832-0,943)	0,724 (0,663-0,786)	0,572 (0,462-0,682)	0,405 (0,202-0,608)
GBM	0,775 (0,701-0,849)	0,726 (0,604-0,848)	0,725 (0,654-0,796)	0,477 (0,366-0,588)	0,886 (0,83-0,942)	0,725 (0,664-0,787)	0,572 (0,462-0,682)	0,404 (0,200-0,608)
XGBoost	0,767 (0,692-0,842)	0,720 (0,597-0,842)	0,718 (0,647-0,789)	0,469 (0,359-0,579)	0,883 (0,826-0,940)	0,719 (0,657-0,780)	0,564 (0,454-0,674)	0,392 (0,188-0,596)
LightGBM	0,755 (0,679-0,831)	0,715 (0,591-0,838)	0,704 (0,632-0,776)	0,455 (0,346-0,564)	0,879 (0,821-0,937)	0,707 (0,644-0,769)	0,552 (0,443-0,661)	0,374 (0,170-0,577)
Rastgele Orman	0,616 (0,528-0,704)	0,556 (0,424-0,689)	0,554 (0,477-0,631)	0,308 (0,216-0,400)	0,780 (0,702-0,858)	0,555 (0,488-0,622)	0,392 (0,295-0,489)	0,098 (0,100-0,297)
Karar Ağacı	0,767 (0,692-0,842)	0,720 (0,597-0,843)	0,719 (0,648-0,79)	0,469 (0,359-0,58)	0,883 (0,827-0,94)	0,719 (0,657-0,781)	0,565 (0,455-0,675)	0,393 (0,189-0,597)
Bilgi Kazancı	0,778 (0,705-0,851)	0,730 (0,608-0,851)	0,723 (0,652-0,794)	0,475 (0,364-0,585)	0,888 (0,832-0,944)	0,725 (0,663-0,786)	0,572 (0,462-0,682)	0,405 (0,202-0,608)
L1 (Lasso)	0,793 (0,722-0,863)	0,744 (0,625-0,864)	0,735 (0,665-0,805)	0,491 (0,380-0,603)	0,894 (0,840-0,948)	0,737 (0,677-0,798)	0,589 (0,479-0,699)	0,430 (0,228-0,632)
DVM	0,772 (0,697-0,847)	0,720 (0,598-0,843)	0,734 (0,664-0,804)	0,484 (0,373-0,596)	0,885 (0,830-0,941)	0,731 (0,670-0,791)	0,576 (0,465-0,686)	0,409 (0,205-0,614)
Bayesci GDM	0,794 (0,723-0,864)	0,745 (0,626-0,864)	0,736 (0,666-0,806)	0,493 (0,381-0,604)	0,895 (0,841-0,949)	0,738 (0,677-0,799)	0,590 (0,480-0,700)	0,431 (0,229-0,633)



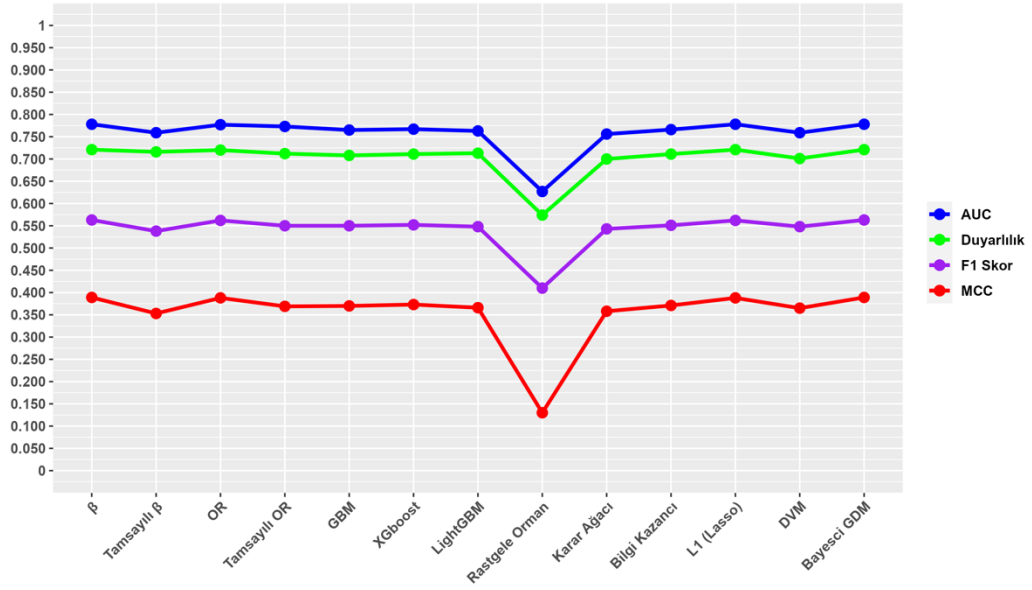
Şekil 3.4. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%20-%80 ve %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken oranının %50-%50 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.5’de verilmiştir.

Çizelge 3.5. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,778 (0,732-0,824)	0,721 (0,643-0,799)	0,715 (0,669-0,760)	0,464 (0,394-0,533)	0,883 (0,847-0,919)	0,716 (0,677-0,756)	0,563 (0,494-0,632)	0,389 (0,260-0,517)
Tam sayılı Beta	0,759 (0,713-0,805)	0,716 (0,638-0,793)	0,679 (0,633-0,726)	0,437 (0,370-0,505)	0,878 (0,840-0,915)	0,689 (0,648-0,729)	0,538 (0,470-0,606)	0,353 (0,225-0,481)
OR	0,777 (0,731-0,823)	0,720 (0,643-0,798)	0,714 (0,669-0,760)	0,463 (0,393-0,532)	0,883 (0,847-0,919)	0,716 (0,677-0,755)	0,562 (0,493-0,631)	0,388 (0,259-0,516)
Tam sayılı OR	0,773 (0,727-0,819)	0,712 (0,633-0,790)	0,703 (0,657-0,749)	0,451 (0,382-0,520)	0,878 (0,841-0,915)	0,705 (0,665-0,745)	0,550 (0,481-0,619)	0,369 (0,240-0,498)
GBM	0,765 (0,717-0,812)	0,708 (0,629-0,787)	0,707 (0,661-0,753)	0,452 (0,383-0,522)	0,877 (0,84-0,914)	0,707 (0,667-0,747)	0,55 (0,481-0,62)	0,37 (0,24-0,499)
XGBoost	0,767 (0,720-0,814)	0,711 (0,632-0,790)	0,708 (0,662-0,754)	0,454 (0,385-0,523)	0,879 (0,842-0,915)	0,709 (0,669-0,748)	0,552 (0,483-0,622)	0,373 (0,244-0,502)
LightGBM	0,763 (0,716-0,810)	0,713 (0,634-0,791)	0,699 (0,653-0,745)	0,447 (0,379-0,516)	0,878 (0,841-0,915)	0,702 (0,663-0,742)	0,548 (0,479-0,616)	0,366 (0,238-0,495)
Rastgele Orman	0,627 (0,572-0,682)	0,574 (0,490-0,658)	0,573 (0,524-0,622)	0,322 (0,262-0,381)	0,794 (0,746-0,842)	0,574 (0,531-0,616)	0,410 (0,347-0,473)	0,130 (0,004-0,257)
Karar Ağacı	0,756 (0,708-0,805)	0,700 (0,62-0,779)	0,703 (0,657-0,749)	0,446 (0,377-0,515)	0,874 (0,836-0,911)	0,702 (0,662-0,742)	0,543 (0,473-0,612)	0,358 (0,228-0,488)
Bilgi Kazancı	0,766 (0,719-0,813)	0,711 (0,633-0,790)	0,706 (0,660-0,752)	0,452 (0,383-0,521)	0,878 (0,842-0,915)	0,707 (0,667-0,747)	0,551 (0,482-0,620)	0,371 (0,242-0,500)
L1 (Lasso)	0,778 (0,732-0,824)	0,721 (0,644-0,799)	0,714 (0,669-0,760)	0,463 (0,394-0,533)	0,883 (0,847-0,919)	0,716 (0,677-0,756)	0,562 (0,493-0,632)	0,388 (0,260-0,517)
DVM	0,759 (0,711-0,807)	0,701 (0,621-0,780)	0,708 (0,662-0,754)	0,452 (0,383-0,521)	0,875 (0,837-0,912)	0,706 (0,667-0,746)	0,548 (0,479-0,617)	0,365 (0,236-0,495)
Bayesci GDM	0,778 (0,732-0,824)	0,721 (0,643-0,799)	0,715 (0,67-0,761)	0,464 (0,394-0,533)	0,883 (0,847-0,919)	0,717 (0,677-0,756)	0,563 (0,494-0,632)	0,389 (0,26-0,518)



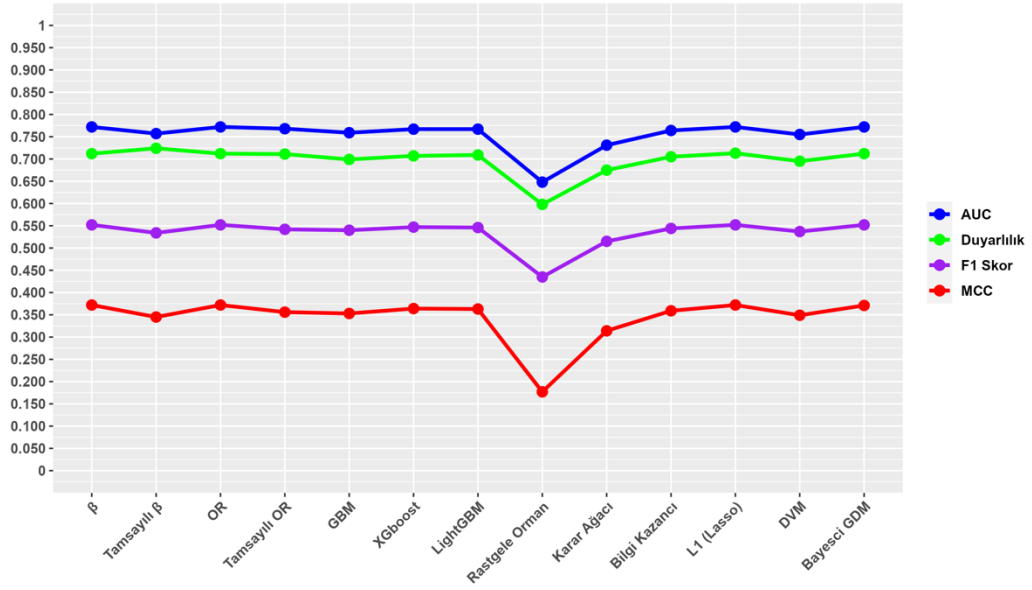
Şekil 3.5. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%20-%80 ve %50-%50)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %50-%50 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları 3.6'da verilmiştir.

Çizelge 3.6. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Senaryo A (n=1000)								
Pozitif Beta katsayıları, dengeli OR dağılımı, bağımlı değişken oranı: %20-%80, bağımsız değişkenlerin oranı: %50-%50								
Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,772 (0,740-0,805)	0,712 (0,657-0,768)	0,706 (0,674-0,739)	0,452 (0,403-0,501)	0,879 (0,853-0,905)	0,708 (0,680-0,736)	0,552 (0,503-0,601)	0,372 (0,281-0,463)
Tam sayılı Beta	0,757 (0,724-0,790)	0,724 (0,670-0,779)	0,664 (0,631-0,698)	0,428 (0,381-0,475)	0,879 (0,852-0,906)	0,680 (0,651-0,708)	0,534 (0,486-0,581)	0,345 (0,256-0,434)
OR	0,772 (0,739-0,805)	0,712 (0,656-0,767)	0,707 (0,674-0,739)	0,452 (0,404-0,501)	0,879 (0,853-0,905)	0,708 (0,680-0,736)	0,552 (0,503-0,601)	0,372 (0,281-0,463)
Tam sayılı OR	0,768 (0,735-0,801)	0,711 (0,656-0,767)	0,690 (0,657-0,723)	0,440 (0,392-0,489)	0,876 (0,850-0,903)	0,696 (0,667-0,724)	0,542 (0,494-0,590)	0,356 (0,266-0,447)
GBM	0,759 (0,725-0,793)	0,699 (0,642-0,755)	0,698 (0,665-0,731)	0,441 (0,392-0,489)	0,873 (0,846-0,899)	0,698 (0,67-0,727)	0,540 (0,491-0,588)	0,353 (0,261-0,444)
XGBoost	0,767 (0,734-0,800)	0,707 (0,652-0,763)	0,702 (0,669-0,735)	0,447 (0,399-0,496)	0,876 (0,850-0,903)	0,703 (0,675-0,732)	0,547 (0,498-0,596)	0,364 (0,273-0,455)
LightGBM	0,767 (0,734-0,800)	0,709 (0,653-0,765)	0,700 (0,667-0,733)	0,446 (0,397-0,494)	0,877 (0,850-0,903)	0,702 (0,674-0,730)	0,546 (0,498-0,595)	0,363 (0,272-0,454)
Rastgele Orman	0,648 (0,609-0,686)	0,598 (0,538-0,657)	0,602 (0,568-0,637)	0,344 (0,300-0,388)	0,812 (0,780-0,844)	0,601 (0,571-0,631)	0,435 (0,389-0,481)	0,177 (0,086-0,267)
Karar Ağacı	0,731 (0,696-0,767)	0,675 (0,617-0,732)	0,679 (0,646-0,713)	0,418 (0,37-0,465)	0,860 (0,833-0,888)	0,678 (0,649-0,707)	0,515 (0,466-0,563)	0,314 (0,222-0,406)
Bilgi Kazancı	0,764 (0,731-0,797)	0,705 (0,649-0,761)	0,699 (0,667-0,732)	0,444 (0,396-0,493)	0,875 (0,849-0,902)	0,701 (0,673-0,729)	0,544 (0,495-0,593)	0,359 (0,268-0,451)
L1 (Lasso)	0,772 (0,739-0,805)	0,713 (0,657-0,768)	0,706 (0,673-0,738)	0,451 (0,403-0,500)	0,879 (0,853-0,905)	0,707 (0,679-0,736)	0,552 (0,503-0,601)	0,372 (0,281-0,463)
DVM	0,755 (0,721-0,789)	0,695 (0,638-0,751)	0,698 (0,665-0,731)	0,440 (0,391-0,488)	0,871 (0,844-0,898)	0,697 (0,669-0,725)	0,537 (0,489-0,586)	0,349 (0,257-0,441)
Bayesici GDM	0,772 (0,739-0,805)	0,712 (0,657-0,768)	0,706 (0,673-0,738)	0,452 (0,403-0,5)	0,879 (0,853-0,905)	0,707 (0,679-0,735)	0,552 (0,503-0,601)	0,371 (0,28-0,463)



Şekil 3.6. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%20-%80 ve %50-%50)

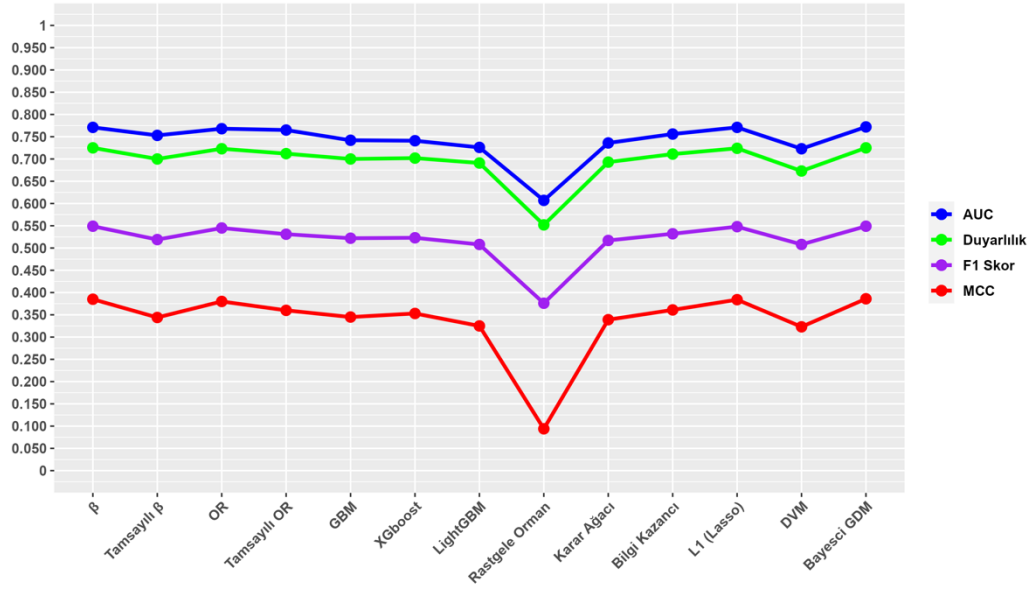
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağacının sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları 3.7’de verilmiştir.

3.1.3. Senaryo A, bağımlı ve bağımsız değişken oranı %20-%80 ve rastgele (n=200, n=500 ve n=1000)

Çizelge 3.7. Senaryo A; n=200; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,771 (0,696-0,846)	0,725 (0,599-0,851)	0,714 (0,643-0,785)	0,447 (0,337-0,557)	0,891 (0,837-0,946)	0,717 (0,655-0,779)	0,549 (0,438-0,659)	0,385 (0,180-0,591)
Tam sayılı Beta	0,753 (0,678-0,827)	0,700 (0,574-0,827)	0,689 (0,618-0,761)	0,424 (0,315-0,533)	0,882 (0,825-0,939)	0,692 (0,628-0,755)	0,519 (0,408-0,629)	0,344 (0,135-0,554)
OR	0,768 (0,692-0,844)	0,723 (0,596-0,850)	0,711 (0,640-0,783)	0,442 (0,333-0,552)	0,891 (0,836-0,946)	0,714 (0,652-0,776)	0,545 (0,434-0,655)	0,380 (0,174-0,585)
Tam sayılı OR	0,765 (0,689-0,841)	0,712 (0,584-0,840)	0,700 (0,628-0,772)	0,429 (0,320-0,538)	0,886 (0,830-0,942)	0,703 (0,640-0,766)	0,531 (0,421-0,641)	0,360 (0,153-0,566)
GBM	0,742 (0,661-0,822)	0,700 (0,570-0,829)	0,696 (0,624-0,769)	0,422 (0,313-0,530)	0,881 (0,823-0,938)	0,697 (0,634-0,760)	0,522 (0,412-0,632)	0,345 (0,138-0,553)
XGBoost	0,741 (0,660-0,823)	0,702 (0,571-0,832)	0,702 (0,631-0,773)	0,425 (0,316-0,534)	0,885 (0,829-0,941)	0,702 (0,639-0,765)	0,523 (0,412-0,634)	0,353 (0,145-0,561)
LightGBM	0,726 (0,644-0,808)	0,691 (0,561-0,822)	0,682 (0,609-0,755)	0,408 (0,301-0,515)	0,876 (0,817-0,935)	0,684 (0,620-0,748)	0,508 (0,399-0,617)	0,325 (0,118-0,532)
Rastgele Orman	0,607 (0,515-0,698)	0,552 (0,415-0,689)	0,556 (0,479-0,632)	0,291 (0,201-0,382)	0,792 (0,715-0,868)	0,555 (0,488-0,622)	0,376 (0,279-0,473)	0,094 (-0,106-0,294)
Karar Ağacı	0,736 (0,654-0,817)	0,693 (0,562-0,823)	0,695 (0,622-0,767)	0,419 (0,311-0,528)	0,878 (0,820-0,936)	0,694 (0,631-0,758)	0,517 (0,407-0,628)	0,339 (0,130-0,547)
Bilgi Kazancı	0,756 (0,678-0,834)	0,711 (0,583-0,840)	0,703 (0,631-0,775)	0,431 (0,322-0,540)	0,886 (0,829-0,942)	0,705 (0,642-0,768)	0,532 (0,422-0,642)	0,361 (0,155-0,568)
L1 (Lasso)	0,771 (0,695-0,846)	0,724 (0,598-0,850)	0,715 (0,644-0,786)	0,446 (0,336-0,556)	0,891 (0,837-0,946)	0,717 (0,655-0,779)	0,548 (0,437-0,659)	0,384 (0,178-0,590)
DVM	0,723 (0,640-0,805)	0,673 (0,541-0,805)	0,693 (0,621-0,765)	0,416 (0,308-0,524)	0,871 (0,811-0,930)	0,688 (0,625-0,752)	0,508 (0,398-0,618)	0,323 (0,114-0,533)
Bayesci GLM	0,772 (0,697-0,848)	0,725 (0,599-0,852)	0,716 (0,645-0,787)	0,447 (0,337-0,557)	0,892 (0,838-0,947)	0,718 (0,656-0,780)	0,549 (0,438-0,660)	0,386 (0,180-0,592)



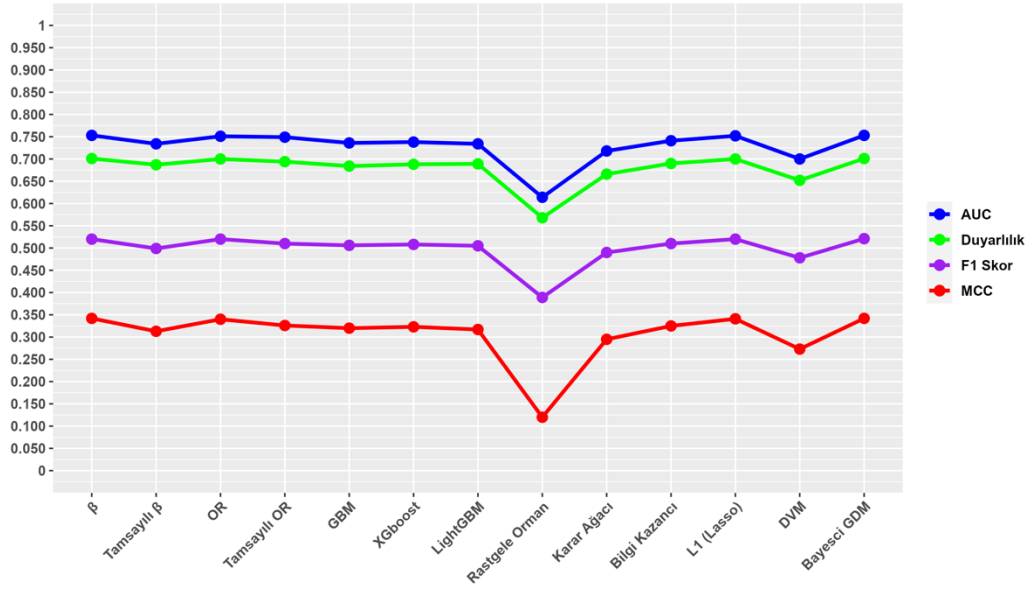
Şekil 3.7. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%20-%80 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları 3.8’de verilmiştir.

Çizelge 3.8. Senaryo A; n=500; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,753 (0,703-0,802)	0,701 (0,619-0,783)	0,692 (0,646-0,739)	0,417 (0,349-0,485)	0,881 (0,844-0,917)	0,695 (0,654-0,735)	0,520 (0,451-0,590)	0,342 (0,212-0,473)
Tam sayılı Beta	0,734 (0,685-0,783)	0,687 (0,606-0,768)	0,668 (0,622-0,715)	0,402 (0,334-0,470)	0,874 (0,837-0,912)	0,673 (0,632-0,714)	0,499 (0,430-0,568)	0,313 (0,181-0,445)
OR	0,751 (0,702-0,801)	0,700 (0,618-0,782)	0,691 (0,645-0,737)	0,417 (0,349-0,485)	0,880 (0,843-0,916)	0,693 (0,653-0,734)	0,520 (0,451-0,589)	0,340 (0,210-0,470)
Tam sayılı OR	0,749 (0,700-0,798)	0,694 (0,612-0,777)	0,680 (0,634-0,727)	0,408 (0,340-0,475)	0,876 (0,839-0,914)	0,684 (0,643-0,724)	0,510 (0,442-0,579)	0,326 (0,196-0,456)
GBM	0,736 (0,685-0,787)	0,684 (0,601-0,767)	0,684 (0,637-0,73)	0,405 (0,337-0,473)	0,873 (0,836-0,911)	0,684 (0,643-0,724)	0,506 (0,437-0,575)	0,320 (0,188-0,451)
XGBoost	0,738-0,789 (0,687-0,771)	0,688 (0,605-0,771)	0,683 (0,637-0,730)	0,406 (0,339-0,474)	0,875 (0,837-0,912)	0,684 (0,644-0,725)	0,508 (0,439-0,577)	0,323 (0,192-0,453)
LightGBM	0,734 (0,683-0,785)	0,689 (0,606-0,772)	0,677 (0,630-0,723)	0,402 (0,335-0,469)	0,874 (0,836-0,911)	0,679 (0,639-0,720)	0,505 (0,436-0,573)	0,317 (0,187-0,447)
Rastgele Orman	0,614 (0,556-0,671)	0,568 (0,481-0,655)	0,570 (0,522-0,619)	0,301 (0,242-0,359)	0,804 (0,757-0,851)	0,570 (0,527-0,612)	0,389 (0,327-0,452)	0,120 (0,007-0,248)
Rastgele Orman	0,718 (0,665-0,770)	0,666 (0,582-0,751)	0,673 (0,626-0,720)	0,392 (0,325-0,459)	0,865 (0,827-0,904)	0,672 (0,631-0,713)	0,490 (0,422-0,559)	0,295 (0,163-0,427)
Bilgi Kazancı	0,741 (0,691-0,791)	0,690 (0,607-0,773)	0,684 (0,638-0,731)	0,408 (0,340-0,475)	0,875 (0,838-0,913)	0,686 (0,645-0,726)	0,510 (0,441-0,579)	0,325 (0,195-0,456)
L1 (Lasso)	0,752 (0,703-0,802)	0,700 (0,618-0,782)	0,692 (0,646-0,738)	0,418 (0,349-0,486)	0,880 (0,843-0,917)	0,694 (0,654-0,734)	0,520 (0,451-0,590)	0,341 (0,211-0,472)
DVM	0,700 (0,647-0,753)	0,652 (0,567-0,737)	0,660 (0,613-0,707)	0,382 (0,316-0,448)	0,857 (0,817-0,897)	0,658 (0,617-0,699)	0,478 (0,410-0,545)	0,273 (0,142-0,404)
Bayesci GDM	0,753 (0,704-0,802)	0,701 (0,619-0,783)	0,692 (0,646-0,738)	0,418 (0,349-0,486)	0,881 (0,844-0,917)	0,694 (0,654-0,735)	0,521 (0,452-0,590)	0,342 (0,212-0,472)



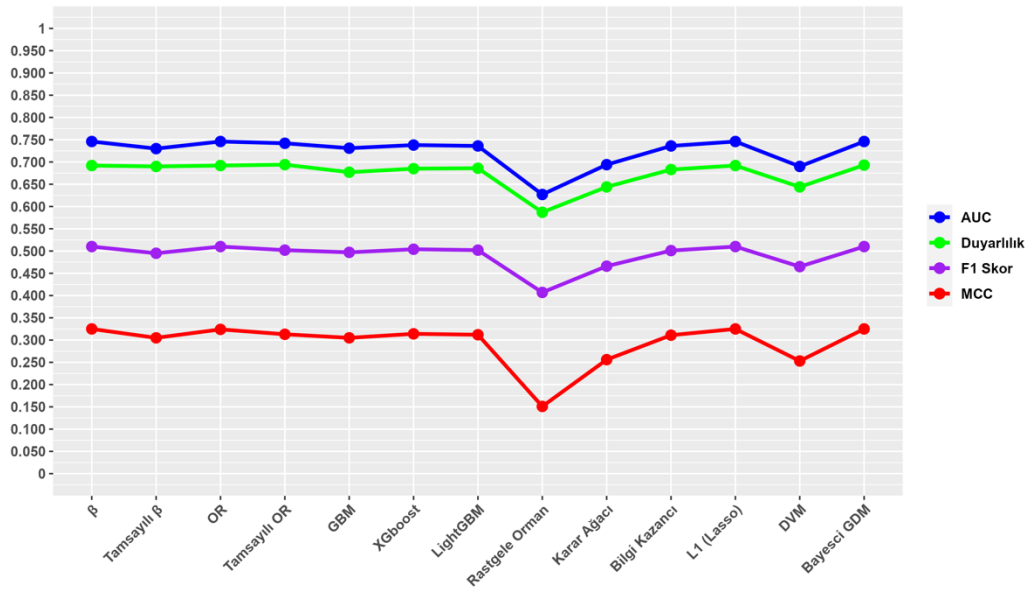
Şekil 3.8. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%20-%80 ve rastgele)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Karar ağacı, DVM ve rastgele ormanın sınıflama başarısının düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları 3.9’da verilmiştir.

Çizelge 3.9. Senaryo A; n=1000; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,746 (0,711-0,782)	0,692 (0,633-0,751)	0,683 (0,650-0,716)	0,406 (0,359-0,454)	0,876 (0,850-0,902)	0,685 (0,656-0,714)	0,510 (0,462-0,559)	0,325 (0,233-0,417)
Tam sayılı Beta	0,730 (0,696-0,765)	0,690 (0,633-0,748)	0,658 (0,625-0,691)	0,395 (0,348-0,442)	0,873 (0,846-0,901)	0,666 (0,637-0,695)	0,495 (0,447-0,544)	0,305 (0,213-0,397)
OR	0,746 (0,710-0,781)	0,692 (0,633-0,750)	0,682 (0,650-0,715)	0,406 (0,358-0,453)	0,876 (0,849-0,902)	0,685 (0,656-0,713)	0,510 (0,461-0,558)	0,324 (0,232-0,416)
Tam sayılı OR	0,742 (0,707-0,777)	0,694 (0,635-0,752)	0,667 (0,634-0,700)	0,398 (0,351-0,445)	0,874 (0,848-0,901)	0,674 (0,645-0,703)	0,502 (0,454-0,550)	0,313 (0,221-0,404)
GBM	0,731 (0,695-0,767)	0,677 (0,618-0,737)	0,675 (0,642-0,708)	0,395 (0,347-0,442)	0,870 (0,847-0,897)	0,676 (0,643-0,697)	0,497 (0,447-0,545)	0,305 (0,213-0,397)
XGBoost	0,738 (0,702-0,774)	0,685 (0,626-0,744)	0,678 (0,645-0,711)	0,401 (0,354-0,448)	0,872 (0,845-0,899)	0,680 (0,651-0,708)	0,504 (0,455-0,552)	0,314 (0,222-0,407)
LightGBM	0,736 (0,700-0,772)	0,686 (0,627-0,745)	0,675 (0,642-0,708)	0,398 (0,351-0,445)	0,873 (0,846-0,900)	0,678 (0,649-0,707)	0,502 (0,453-0,550)	0,312 (0,220-0,405)
Rastgele Orman	0,627 (0,587-0,667)	0,587 (0,525-0,649)	0,588 (0,553-0,622)	0,315 (0,272-0,357)	0,817 (0,784-0,849)	0,587 (0,557-0,618)	0,407 (0,362-0,452)	0,151 (0,061-0,242)
Karar Ağacı	0,694 (0,656-0,732)	0,644 (0,584-0,705)	0,651 (0,618-0,685)	0,369 (0,323-0,415)	0,854 (0,825-0,883)	0,650 (0,620-0,679)	0,466 (0,418-0,514)	0,256 (0,163-0,349)
Bilgi Kazancı	0,736 (0,700-0,771)	0,683 (0,624-0,742)	0,676 (0,643-0,709)	0,399 (0,351-0,446)	0,871 (0,844-0,898)	0,677 (0,648-0,706)	0,501 (0,453-0,550)	0,311 (0,218-0,403)
L1 (Lasso)	0,746 (0,711-0,781)	0,692 (0,634-0,751)	0,682 (0,649-0,715)	0,406 (0,359-0,454)	0,876 (0,849-0,902)	0,685 (0,656-0,714)	0,510 (0,462-0,559)	0,325 (0,233-0,417)
DVM	0,690 (0,653-0,728)	0,644 (0,584-0,705)	0,647 (0,613-0,680)	0,368 (0,322-0,413)	0,853 (0,824-0,882)	0,646 (0,616-0,675)	0,465 (0,418-0,512)	0,253 (0,161-0,345)
Bayesci GLM	0,746 (0,711-0,781)	0,693 (0,634-0,751)	0,682 (0,649-0,715)	0,406 (0,358-0,453)	0,876 (0,850-0,903)	0,685 (0,656-0,713)	0,510 (0,461-0,558)	0,325 (0,233-0,417)



Şekil 3.9. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%20-%80 ve rastgele)

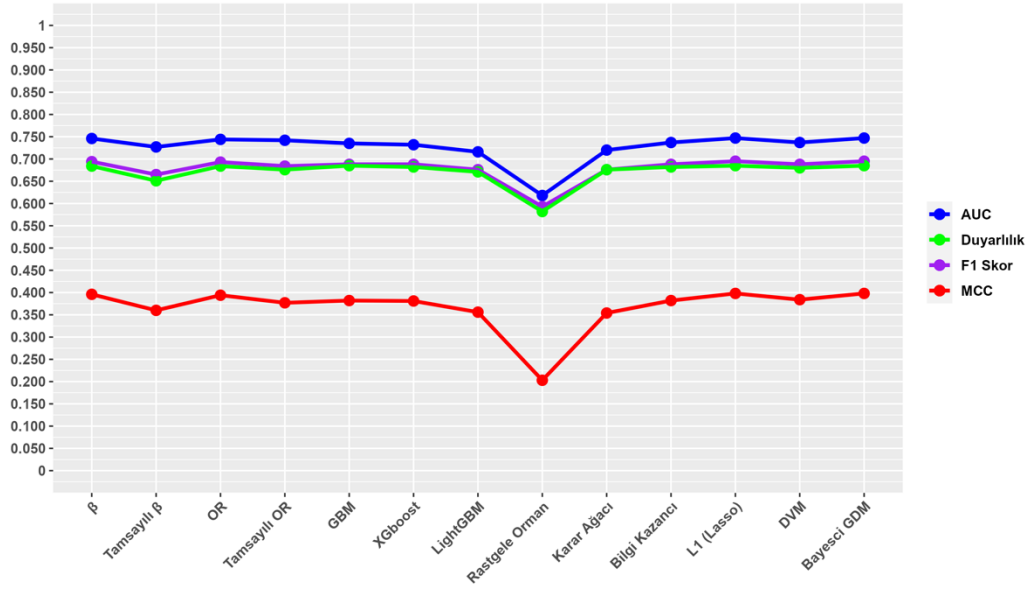
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Karar ağacı, DVM ve rastgele ormanın sınıflama başarısının düşük olduğu belirlenmiştir.

3.1.4. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve %20-%80 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken oranının %20-%80 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetilmiştir. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.10'da verilmiştir.

Çizelge 3.10. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,746 (0,679-0,814)	0,684 (0,594-0,775)	0,711 (0,622-0,800)	0,708 (0,618-0,798)	0,690 (0,600-0,779)	0,698 (0,634-0,761)	0,694 (0,603-0,785)	0,396 (0,214-0,579)
Tam sayılı Beta	0,727 (0,661-0,793)	0,651 (0,560-0,741)	0,703 (0,616-0,790)	0,697 (0,604-0,790)	0,671 (0,580-0,761)	0,676 (0,612-0,741)	0,665 (0,569-0,761)	0,360 (0,170-0,551)
OR	0,744 (0,676-0,812)	0,684 (0,594-0,774)	0,710 (0,621-0,798)	0,706 (0,616-0,796)	0,689 (0,600-0,778)	0,697 (0,633-0,760)	0,693 (0,601-0,784)	0,394 (0,212-0,577)
Tam sayılı OR	0,742 (0,674-0,809)	0,676 (0,585-0,767)	0,699 (0,610-0,789)	0,697 (0,606-0,788)	0,682 (0,592-0,772)	0,688 (0,624-0,752)	0,684 (0,591-0,776)	0,377 (0,193-0,561)
GBM	0,735 (0,666-0,804)	0,685 (0,595-0,775)	0,696 (0,606-0,786)	0,696 (0,606-0,787)	0,686 (0,596-0,777)	0,690 (0,626-0,754)	0,688 (0,597-0,780)	0,382 (0,199-0,564)
XGBoost	0,732 (0,663-0,802)	0,682 (0,592-0,772)	0,698 (0,608-0,788)	0,697 (0,607-0,788)	0,685 (0,595-0,775)	0,690 (0,626-0,754)	0,688 (0,596-0,779)	0,381 (0,199-0,564)
LightGBM	0,716 (0,645-0,787)	0,671 (0,580-0,762)	0,684 (0,593-0,775)	0,684 (0,593-0,775)	0,673 (0,582-0,764)	0,678 (0,613-0,742)	0,676 (0,583-0,768)	0,356 (0,172-0,540)
Rastgele Orman	0,618 (0,539-0,696)	0,582 (0,488-0,677)	0,619 (0,526-0,713)	0,610 (0,514-0,707)	0,595 (0,501-0,688)	0,601 (0,534-0,668)	0,593 (0,495-0,690)	0,203 (0,009-0,397)
Karar Ağacı	0,720 (0,650-0,789)	0,676 (0,586-0,767)	0,677 (0,585-0,768)	0,681 (0,590-0,772)	0,674 (0,582-0,766)	0,676 (0,612-0,741)	0,676 (0,585-0,768)	0,354 (0,171-0,538)
Bilgi Kazancı	0,737 (0,668-0,805)	0,682 (0,591-0,772)	0,699 (0,609-0,789)	0,698 (0,608-0,789)	0,684 (0,594-0,775)	0,690 (0,626-0,754)	0,688 (0,596-0,779)	0,382 (0,199-0,564)
L1 (Lasso)	0,747 (0,679-0,814)	0,685 (0,595-0,775)	0,712 (0,623-0,800)	0,708 (0,619-0,798)	0,691 (0,601-0,780)	0,698 (0,635-0,762)	0,695 (0,603-0,786)	0,398 (0,216-0,580)
DVM	0,737 (0,668-0,805)	0,680 (0,591-0,771)	0,702 (0,613-0,792)	0,700 (0,610-0,791)	0,685 (0,595-0,775)	0,691 (0,628-0,755)	0,688 (0,597-0,780)	0,384 (0,201-0,567)
Bayesci GDM	0,747 (0,679-0,814)	0,685 (0,595-0,775)	0,712 (0,624-0,800)	0,708 (0,619-0,798)	0,690 (0,601-0,780)	0,698 (0,635-0,762)	0,695 (0,603-0,786)	0,398 (0,216-0,580)



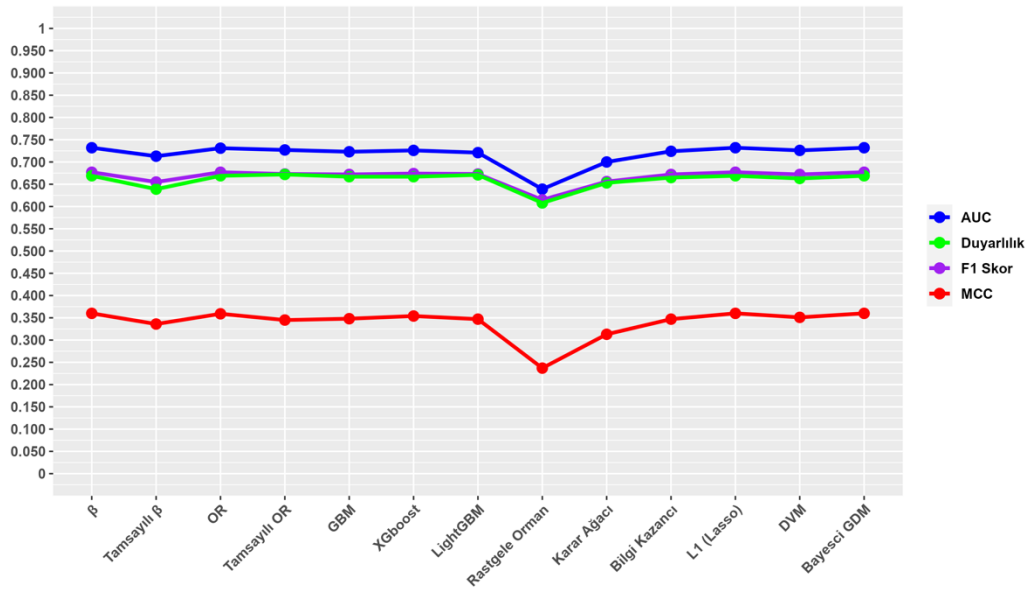
Şekil 3.10. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%50-%50 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının %20-%80 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan 10000 benzetim ile üretilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.11’de verilmiştir

Çizelge 3.11. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,732 (0,688-0,775)	0,669 (0,611-0,727)	0,690 (0,633-0,747)	0,688 (0,630-0,746)	0,673 (0,615-0,730)	0,680 (0,639-0,720)	0,677 (0,619-0,736)	0,360 (0,243-0,477)
Tam sayılı Beta	0,713 (0,670-0,756)	0,639 (0,581-0,697)	0,692 (0,635-0,748)	0,682 (0,623-0,742)	0,658 (0,600-0,715)	0,665 (0,624-0,707)	0,655 (0,594-0,716)	0,336 (0,214-0,457)
OR	0,731 (0,687-0,774)	0,669 (0,611-0,727)	0,689 (0,632-0,746)	0,687 (0,629-0,745)	0,672 (0,615-0,730)	0,679 (0,638-0,720)	0,677 (0,618-0,735)	0,359 (0,242-0,476)
Tam sayılı OR	0,727 (0,684-0,771)	0,672 (0,614-0,730)	0,671 (0,613-0,729)	0,677 (0,619-0,735)	0,669 (0,611-0,728)	0,672 (0,631-0,713)	0,673 (0,614-0,731)	0,345 (0,228-0,461)
GBM	0,723 (0,679-0,767)	0,667 (0,610-0,725)	0,679 (0,622-0,737)	0,680 (0,622-0,738)	0,668 (0,611-0,727)	0,673 (0,632-0,715)	0,672 (0,614-0,731)	0,348 (0,231-0,465)
XGBoost	0,726 (0,683-0,770)	0,667 (0,609-0,725)	0,686 (0,628-0,744)	0,684 (0,626-0,742)	0,670 (0,613-0,728)	0,677 (0,636-0,718)	0,674 (0,616-0,733)	0,354 (0,237-0,471)
LightGBM	0,721 (0,677-0,766)	0,671 (0,613-0,728)	0,675 (0,618-0,733)	0,678 (0,620-0,736)	0,669 (0,611-0,728)	0,673 (0,632-0,714)	0,673 (0,615-0,731)	0,347 (0,230-0,463)
Rastgele Orman	0,639 (0,591-0,687)	0,608 (0,549-0,668)	0,628 (0,569-0,688)	0,626 (0,565-0,686)	0,613 (0,553-0,672)	0,618 (0,576-0,660)	0,615 (0,554-0,676)	0,237 (0,116-0,359)
Karar Ağacı	0,700 (0,655-0,745)	0,653 (0,595-0,712)	0,659 (0,600-0,718)	0,662 (0,603-0,720)	0,652 (0,593-0,711)	0,656 (0,615-0,698)	0,656 (0,597-0,715)	0,313 (0,195-0,431)
Bilgi Kazancı	0,724 (0,680-0,768)	0,665 (0,607-0,723)	0,681 (0,623-0,739)	0,681 (0,623-0,739)	0,667 (0,609-0,725)	0,673 (0,632-0,714)	0,672 (0,613-0,730)	0,347 (0,230-0,464)
L1 (Lasso)	0,732 (0,688-0,775)	0,669 (0,611-0,727)	0,690 (0,633-0,747)	0,687 (0,630-0,745)	0,673 (0,615-0,731)	0,679 (0,639-0,720)	0,677 (0,619-0,736)	0,360 (0,243-0,477)
DVM	0,726 (0,682-0,770)	0,663 (0,605-0,721)	0,687 (0,629-0,744)	0,684 (0,626-0,742)	0,668 (0,610-0,726)	0,675 (0,634-0,716)	0,672 (0,613-0,731)	0,351 (0,233-0,468)
Bayesci GDM	0,732 (0,688-0,775)	0,669 (0,611-0,727)	0,690 (0,633-0,748)	0,688 (0,630-0,746)	0,673 (0,615-0,730)	0,679 (0,639-0,720)	0,677 (0,619-0,736)	0,360 (0,243-0,477)



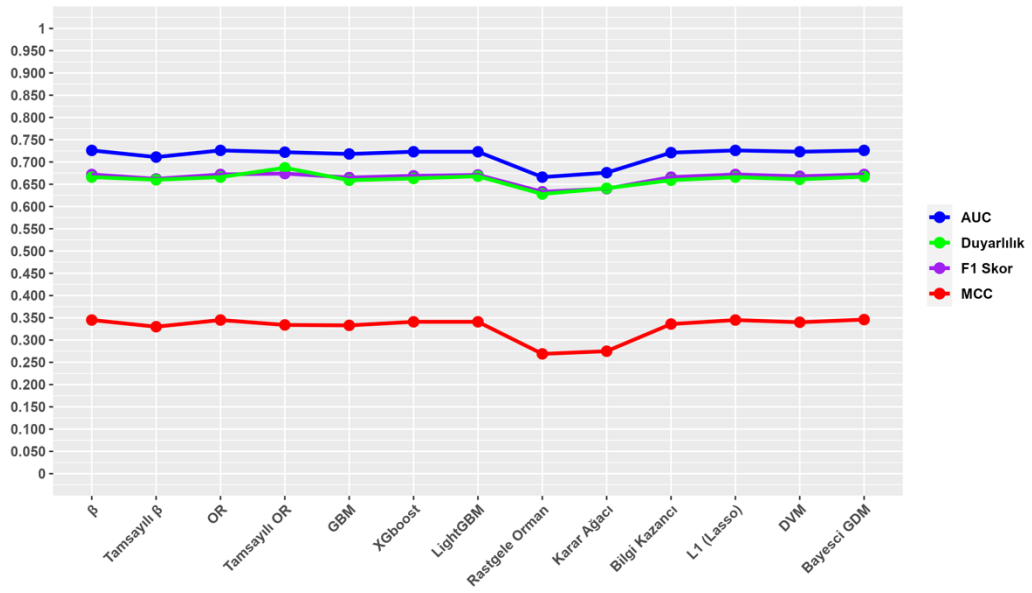
Şekil 3.11. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%50-%50 ve %20-%80)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağaçlarının sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının %20-%80 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.12’de verilmişti

Çizelge 3.12. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı:%50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,726 (0,695-0,757)	0,666 (0,625-0,707)	0,678 (0,637-0,719)	0,678 (0,637-0,720)	0,667 (0,626-0,708)	0,672 (0,643-0,701)	0,672 (0,630-0,713)	0,345 (0,262-0,428)
Tam sayılı Beta	0,711 (0,681-0,741)	0,660 (0,619-0,701)	0,667 (0,626-0,708)	0,670 (0,629-0,712)	0,662 (0,621-0,703)	0,663 (0,634-0,693)	0,662 (0,620-0,704)	0,330 (0,246-0,413)
OR	0,726 (0,690-0,757)	0,666 (0,625-0,707)	0,678 (0,637-0,719)	0,679 (0,637-0,72)	0,667 (0,626-0,708)	0,672 (0,643-0,701)	0,672 (0,630-0,713)	0,345 (0,262-0,428)
Tam sayılı OR	0,722 (0,691-0,752)	0,687 (0,647-0,727)	0,645 (0,603-0,687)	0,665 (0,620-0,705)	0,671 (0,629-0,713)	0,666 (0,637-0,695)	0,674 (0,634-0,714)	0,334 (0,253-0,415)
GBM	0,718 (0,687-0,750)	0,659 (0,618-0,700)	0,673 (0,632-0,715)	0,673 (0,631-0,714)	0,660 (0,619-0,702)	0,666 (0,637-0,695)	0,665 (0,623-0,707)	0,333 (0,249-0,416)
XGBoost	0,723 (0,692-0,754)	0,663 (0,622-0,704)	0,678 (0,636-0,719)	0,677 (0,636-0,718)	0,665 (0,624-0,706)	0,670 (0,641-0,699)	0,669 (0,628-0,711)	0,341 (0,258-0,424)
LightGBM	0,723 (0,692-0,754)	0,668 (0,627-0,709)	0,672 (0,631-0,714)	0,675 (0,634-0,716)	0,666 (0,625-0,707)	0,670 (0,641-0,699)	0,671 (0,630-0,712)	0,341 (0,258-0,423)
Rastgele Orman	0,666 (0,633-0,700)	0,628 (0,586-0,670)	0,641 (0,599-0,683)	0,641 (0,598-0,683)	0,629 (0,587-0,671)	0,634 (0,604-0,664)	0,633 (0,590-0,675)	0,269 (0,184-0,354)
Karar Ağacı	0,676 (0,644-0,708)	0,641 (0,599-0,682)	0,634 (0,592-0,676)	0,640 (0,599-0,682)	0,635 (0,593-0,678)	0,637 (0,608-0,667)	0,640 (0,598-0,682)	0,275 (0,191-0,359)
Bilgi Kazancı	0,721 (0,689-0,752)	0,659 (0,618-0,701)	0,676 (0,635-0,717)	0,675 (0,633-0,716)	0,662 (0,621-0,703)	0,667 (0,638-0,697)	0,666 (0,624-0,708)	0,336 (0,253-0,419)
L1 (Lasso)	0,726 (0,695-0,757)	0,666 (0,625-0,707)	0,679 (0,638-0,720)	0,678 (0,637-0,719)	0,667 (0,626-0,708)	0,672 (0,643-0,701)	0,672 (0,630-0,713)	0,345 (0,263-0,428)
DVM	0,723 (0,691-0,754)	0,661 (0,619-0,702)	0,679 (0,638-0,720)	0,677 (0,636-0,719)	0,664 (0,623-0,705)	0,670 (0,641-0,699)	0,668 (0,627-0,710)	0,340 (0,257-0,424)
Bayesci GDM	0,726 (0,695-0,757)	0,667 (0,626-0,708)	0,678 (0,637-0,719)	0,679 (0,638-0,720)	0,667 (0,626-0,708)	0,673 (0,643-0,702)	0,672 (0,631-0,713)	0,346 (0,263-0,428)



Şekil 3.12. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%50-%50 ve %20-%80)

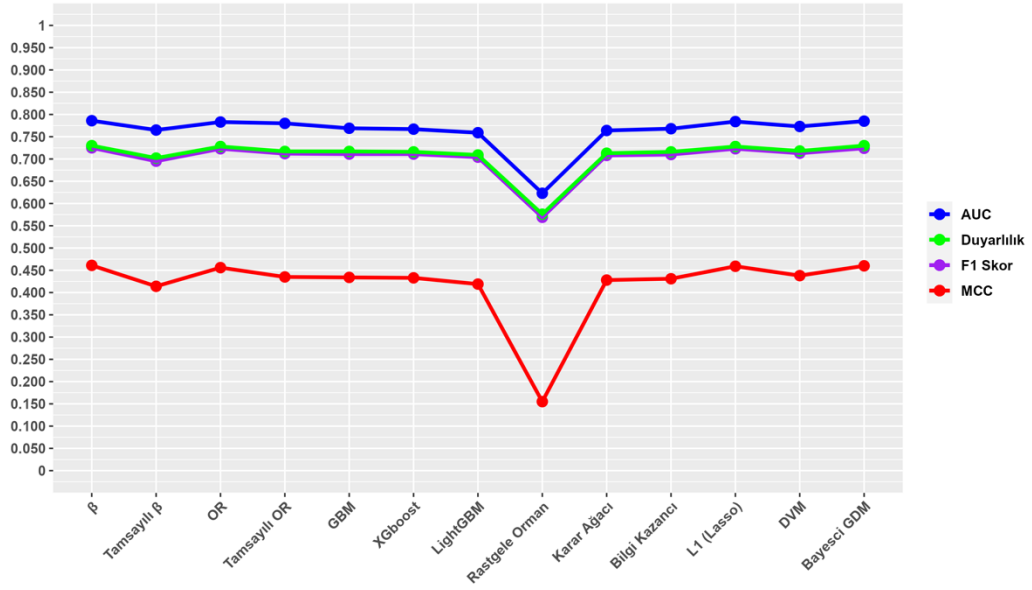
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağaçlarının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.5. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve %50-%50 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken oranının %50-%50 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan 10000 benzetim ile veriler türetildi. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.13' de verilmiştir.

Çizelge 3.13. Senaryo A; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,786 (0,723-0,848)	0,730 (0,643-0,818)	0,730 (0,644-0,815)	0,723 (0,635-0,811)	0,739 (0,653-0,824)	0,730 (0,669-0,791)	0,725 (0,637-0,813)	0,461 (0,286-0,636)
Tam sayılı Beta	0,765 (0,702-0,828)	0,702 (0,614-0,790)	0,705 (0,619-0,791)	0,702 (0,612-0,793)	0,718 (0,630-0,806)	0,704 (0,640-0,767)	0,695 (0,604-0,787)	0,414 (0,232-0,595)
OR	0,783 (0,720-0,846)	0,728 (0,641-0,816)	0,727 (0,641-0,813)	0,721 (0,630-0,809)	0,736 (0,650-0,822)	0,728 (0,666-0,789)	0,723 (0,635-0,811)	0,456 (0,281-0,632)
Tam sayılı OR	0,780 (0,717-0,843)	0,717 (0,629-0,806)	0,716 (0,629-0,803)	0,710 (0,621-0,799)	0,726 (0,639-0,813)	0,717 (0,655-0,779)	0,712 (0,622-0,801)	0,435 (0,258-0,612)
GBM	0,769 (0,704-0,834)	0,717 (0,629-0,806)	0,716 (0,629-0,803)	0,709 (0,620-0,798)	0,726 (0,639-0,813)	0,717 (0,654-0,779)	0,711 (0,622-0,801)	0,434 (0,257-0,612)
XGBoost	0,767 (0,702-0,832)	0,716 (0,627-0,804)	0,717 (0,630-0,803)	0,710 (0,621-0,799)	0,725 (0,638-0,811)	0,716 (0,654-0,778)	0,711 (0,622-0,800)	0,433 (0,256-0,611)
LightGBM	0,759 (0,693-0,825)	0,709 (0,620-0,798)	0,710 (0,622-0,797)	0,702 (0,612-0,792)	0,718 (0,631-0,806)	0,709 (0,647-0,772)	0,704 (0,614-0,794)	0,419 (0,241-0,598)
Rastgele Orman	0,623 (0,547-0,700)	0,576 (0,481-0,671)	0,578 (0,485-0,671)	0,568 (0,473-0,664)	0,587 (0,493-0,682)	0,577 (0,510-0,644)	0,569 (0,474-0,665)	0,155 (-0,037-0,346)
Karar Ağacı	0,764 (0,699-0,830)	0,713 (0,624-0,802)	0,714 (0,627-0,801)	0,706 (0,617-0,796)	0,723 (0,636-0,81)	0,714 (0,651-0,776)	0,708 (0,618-0,797)	0,428 (0,25-0,606)
Bilgi Kazancı	0,768 (0,703-0,833)	0,716 (0,627-0,804)	0,714 (0,627-0,802)	0,708 (0,619-0,797)	0,724 (0,637-0,811)	0,715 (0,653-0,777)	0,710 (0,621-0,799)	0,431 (0,253-0,609)
L1 (Lasso)	0,784 (0,721-0,847)	0,728 (0,641-0,816)	0,730 (0,644-0,815)	0,722 (0,634-0,810)	0,738 (0,652-0,823)	0,729 (0,668-0,790)	0,723 (0,635-0,812)	0,459 (0,284-0,634)
DVM	0,773 (0,709-0,838)	0,718 (0,630-0,807)	0,718 (0,632-0,805)	0,712 (0,623-0,801)	0,727 (0,641-0,814)	0,718 (0,656-0,780)	0,713 (0,624-0,802)	0,438 (0,261-0,615)
Bayesci GDM	0,785 (0,723-0,848)	0,730 (0,643-0,817)	0,730 (0,644-0,815)	0,722 (0,635-0,81)	0,739 (0,653-0,824)	0,730 (0,668-0,791)	0,724 (0,636-0,812)	0,46 (0,285-0,635)



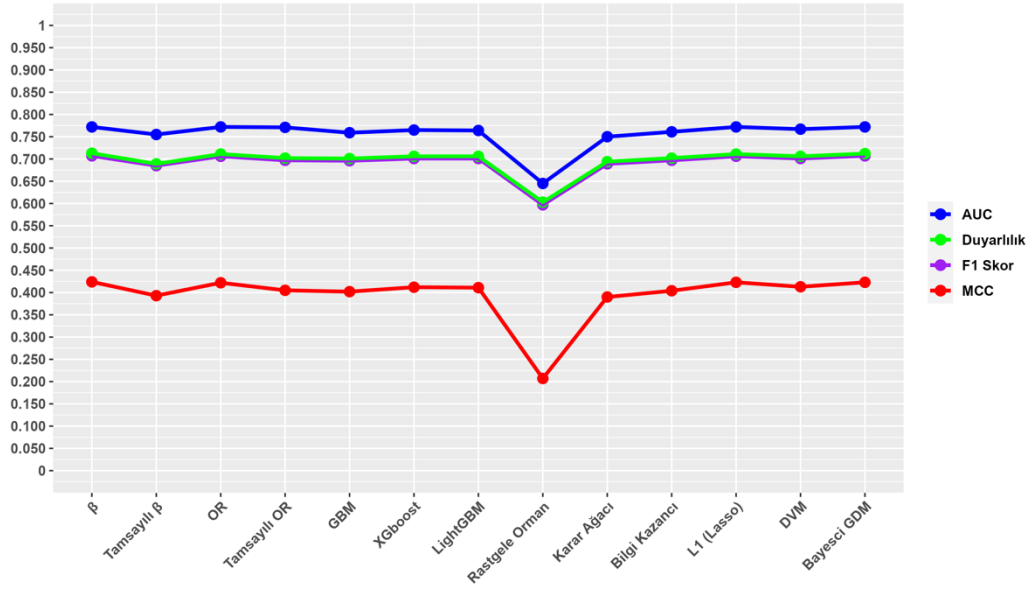
Şekil 3.13. LR ve ML algoritmaları ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%50-%50 ve %50-%50)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının %50-%50 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.14’de verilmiştir.

Çizelge 3.14. Senaryo A; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,772 (0,732-0,813)	0,713 (0,656-0,769)	0,711 (0,655-0,766)	0,704 (0,647-0,760)	0,721 (0,665-0,776)	0,712 (0,672-0,751)	0,707 (0,651-0,764)	0,424 (0,311-0,536)
Tam sayılı Beta	0,755 (0,714-0,795)	0,689 (0,633-0,746)	0,697 (0,642-0,752)	0,693 (0,635-0,751)	0,706 (0,65-0,763)	0,693 (0,653-0,734)	0,685 (0,650-0,763)	0,393 (0,276-0,509)
OR	0,772 (0,731-0,812)	0,711 (0,655-0,768)	0,711 (0,655-0,766)	0,703 (0,646-0,760)	0,720 (0,664-0,775)	0,711 (0,671-0,751)	0,706 (0,626-0,744)	0,422 (0,310-0,535)
Tam sayılı OR	0,771 (0,731-0,812)	0,702 (0,645-0,759)	0,702 (0,646-0,757)	0,695 (0,637-0,752)	0,711 (0,655-0,767)	0,702 (0,662-0,742)	0,697 (0,640-0,754)	0,405 (0,291-0,519)
GBM	0,759 (0,717-0,800)	0,701 (0,643-0,758)	0,701 (0,645-0,757)	0,693 (0,636-0,751)	0,709 (0,653-0,765)	0,701 (0,661-0,741)	0,696 (0,639-0,753)	0,402 (0,288-0,516)
XGBoost	0,765 (0,724-0,806)	0,706 (0,649-0,762)	0,706 (0,650-0,762)	0,698 (0,641-0,755)	0,715 (0,659-0,770)	0,706 (0,666-0,746)	0,701 (0,644-0,758)	0,412 (0,299-0,526)
LightGBM	0,764 (0,720-0,805)	0,706 (0,649-0,763)	0,705 (0,649-0,761)	0,698 (0,641-0,755)	0,714 (0,659-0,770)	0,705 (0,666-0,745)	0,701 (0,644-0,758)	0,411 (0,298-0,525)
Rastgele Orman	0,645 (0,598-0,693)	0,603 (0,543-0,663)	0,604 (0,545-0,663)	0,595 (0,534-0,655)	0,613 (0,554-0,673)	0,603 (0,561-0,646)	0,597 (0,537-0,658)	0,207 (0,087-0,328)
Karar Ağacı	0,750 (0,708-0,792)	0,694 (0,637-0,752)	0,695 (0,638-0,751)	0,686 (0,629-0,744)	0,704 (0,647-0,76)	0,694 (0,654-0,735)	0,689 (0,632-0,747)	0,39 (0,275-0,504)
Bilgi Kazancı	0,761 (0,719-0,802)	0,702 (0,645-0,759)	0,701 (0,645-0,757)	0,694 (0,637-0,751)	0,711 (0,655-0,767)	0,702 (0,662-0,742)	0,697 (0,640-0,754)	0,404 (0,290-0,518)
L1 (Lasso)	0,772 (0,731-0,813)	0,711 (0,655-0,768)	0,711 (0,655-0,766)	0,703 (0,646-0,760)	0,720 (0,665-0,776)	0,711 (0,671-0,751)	0,706 (0,650-0,763)	0,423 (0,310-0,536)
DVM	0,767 (0,726-0,808)	0,706 (0,649-0,763)	0,706 (0,651-0,762)	0,699 (0,642-0,756)	0,715 (0,659-0,771)	0,706 (0,666-0,746)	0,701 (0,644-0,758)	0,413 (0,300-0,527)
Bayesci GDM	0,772 (0,732-0,813)	0,712 (0,656-0,769)	0,711 (0,655-0,766)	0,704 (0,647-0,76)	0,720 (0,665-0,776)	0,711 (0,672-0,751)	0,707 (0,65-0,763)	0,423 (0,310-0,536)



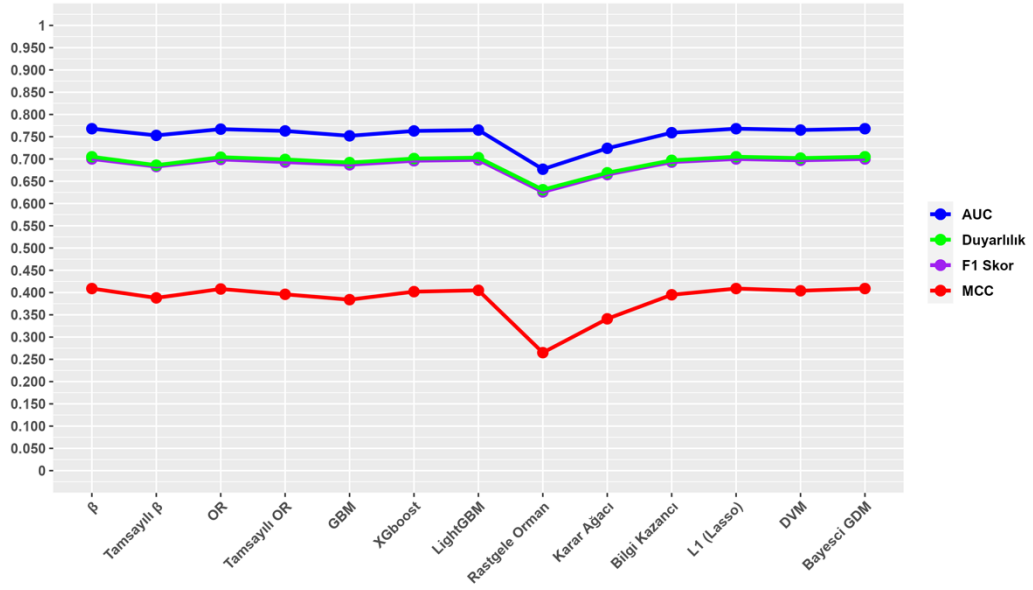
Şekil 3.14. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%50-%50 ve %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının %50-%50 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.15’de verilmiştir.

Çizelge 3.15. Senaryo A; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,768 (0,739-0,797)	0,705 (0,664-0,745)	0,704 (0,665-0,744)	0,696 (0,656-0,737)	0,723 (0,674-0,753)	0,704 (0,676-0,733)	0,700 (0,659-0,740)	0,409 (0,329-0,489)
Tam sayılı Beta	0,753 (0,725-0,782)	0,686 (0,646-0,726)	0,697 (0,658-0,736)	0,691 (0,65-0,732)	0,708 (0,663-0,743)	0,692 (0,663-0,720)	0,683 (0,641-0,725)	0,388 (0,306-0,471)
OR	0,767 (0,738-0,796)	0,704 (0,664-0,745)	0,703 (0,664-0,743)	0,695 (0,655-0,736)	0,730 (0,673-0,752)	0,704 (0,675-0,732)	0,699 (0,659-0,740)	0,408 (0,327-0,488)
Tam sayılı OR	0,763 (0,734-0,792)	0,699 (0,658-0,739)	0,697 (0,657-0,737)	0,690 (0,649-0,730)	0,717 (0,667-0,747)	0,698 (0,669-0,726)	0,693 (0,653-0,734)	0,396 (0,315-0,477)
GBM	0,752 (0,723-0,782)	0,692 (0,651-0,733)	0,691 (0,651-0,731)	0,684 (0,643-0,725)	0,701 (0,661-0,741)	0,692 (0,663-0,72)	0,687 (0,646-0,728)	0,384 (0,303-0,465)
XGBoost	0,763 (0,734-0,792)	0,701 (0,660-0,741)	0,701 (0,661-0,740)	0,693 (0,652-0,733)	0,710 (0,670-0,749)	0,701 (0,672-0,729)	0,696 (0,655-0,736)	0,402 (0,321-0,482)
LightGBM	0,765 (0,736-0,794)	0,703 (0,663-0,743)	0,702 (0,662-0,742)	0,694 (0,654-0,735)	0,711 (0,672-0,751)	0,703 (0,674-0,731)	0,698 (0,658-0,739)	0,405 (0,325-0,486)
Rastgele Orman	0,677 (0,645-0,710)	0,631 (0,589-0,674)	0,633 (0,591-0,674)	0,624 (0,581-0,666)	0,641 (0,600-0,683)	0,632 (0,602-0,662)	0,626 (0,584-0,669)	0,265 (0,180-0,349)
Karar Ağacı	0,724 (0,693-0,755)	0,669 (0,627-0,71)	0,672 (0,631-0,712)	0,663 (0,621-0,704)	0,679 (0,638-0,720)	0,670 (0,641-0,699)	0,665 (0,623-0,707)	0,341 (0,258-0,424)
Bilgi Kazancı	0,759 (0,730-0,789)	0,697 (0,657-0,738)	0,697 (0,657-0,737)	0,689 (0,649-0,730)	0,706 (0,666-0,745)	0,697 (0,669-0,726)	0,693 (0,652-0,733)	0,395 (0,314-0,475)
L1 (Lasso)	0,768 (0,739-0,797)	0,705 (0,665-0,745)	0,704 (0,664-0,744)	0,697 (0,656-0,737)	0,713 (0,673-0,752)	0,704 (0,676-0,733)	0,700 (0,660-0,740)	0,409 (0,329-0,490)
DVM	0,765 (0,736-0,794)	0,702 (0,662-0,743)	0,702 (0,662-0,741)	0,694 (0,654-0,735)	0,711 (0,671-0,750)	0,702 (0,674-0,730)	0,697 (0,657-0,738)	0,404 (0,324-0,485)
Bayesci GDM	0,768 (0,739-0,797)	0,705 (0,665-0,745)	0,704 (0,665-0,744)	0,697 (0,656-0,737)	0,713 (0,674-0,753)	0,704 (0,676-0,733)	0,700 (0,66-0,74)	0,409 (0,329-0,490)



Şekil 3.15. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%50-%50 ve %50-%50)

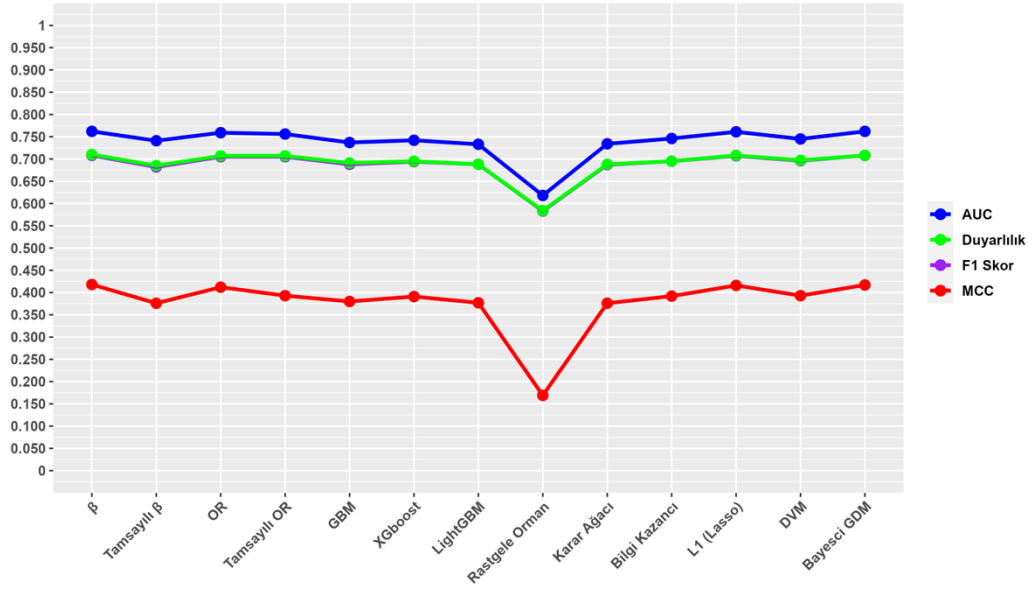
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağacının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.6. Senaryo A, bağımlı ve bağımsız değişken oranı %50-%50 ve rastgele (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.16'de verilmiştir.

Çizelge 3.16. Senaryo A; n=200; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,762 (0,696-0,828)	0,710 (0,622-0,798)	0,708 (0,619-0,797)	0,711 (0,623-0,799)	0,707 (0,618-0,796)	0,709 (0,646-0,772)	0,708 (0,619-0,797)	0,418 (0,240-0,596)
Tam sayılı Beta	0,741 (0,676-0,806)	0,685 (0,597-0,773)	0,685 (0,596-0,774)	0,694 (0,604-0,784)	0,688 (0,192-0,560)	0,685 (0,597-0,779)	0,682 (0,621-0,749)	0,376 (0,192-0,560)
OR	0,759 (0,693-0,825)	0,707 (0,618-0,795)	0,706 (0,617-0,795)	0,709 (0,620-0,797)	0,704 (0,615-0,793)	0,706 (0,643-0,769)	0,705 (0,616-0,795)	0,412 (0,234-0,591)
Tam sayılı OR	0,756 (0,690-0,822)	0,707 (0,606-0,785)	0,706 (0,607-0,786)	0,709 (0,609-0,788)	0,704 (0,605-0,785)	0,706 (0,633-0,760)	0,705 (0,604-0,785)	0,393 (0,212-0,573)
GBM	0,737 (0,668-0,806)	0,691 (0,601-0,781)	0,689 (0,598-0,779)	0,691 (0,601-0,781)	0,689 (0,598-0,779)	0,690 (0,626-0,754)	0,688 (0,598-0,779)	0,380 (0,198-0,561)
XGBoost	0,742 (0,674-0,811)	0,695 (0,606-0,785)	0,695 (0,605-0,785)	0,698 (0,608-0,787)	0,693 (0,603-0,783)	0,695 (0,632-0,759)	0,694 (0,604-0,784)	0,391 (0,210-0,571)
LightGBM	0,733 (0,664-0,802)	0,688 (0,598-0,778)	0,689 (0,599-0,780)	0,692 (0,602-0,782)	0,686 (0,595-0,776)	0,689 (0,625-0,753)	0,688 (0,597-0,778)	0,377 (0,196-0,559)
Rastgele Orman	0,618 (0,541-0,695)	0,584 (0,490-0,678)	0,584 (0,489-0,679)	0,588 (0,493-0,683)	0,582 (0,486-0,677)	0,584 (0,517-0,651)	0,583 (0,487-0,678)	0,169 (0,022-0,360)
Karar Ağacı	0,734 (0,665-0,803)	0,688 (0,598-0,778)	0,687 (0,597-0,778)	0,691 (0,601-0,781)	0,686 (0,595-0,777)	0,688 (0,624-0,752)	0,687 (0,596-0,778)	0,376 (0,194-0,558)
Bilgi Kazancı	0,746 (0,678-0,814)	0,695 (0,606-0,785)	0,696 (0,606-0,786)	0,699 (0,610-0,788)	0,693 (0,603-0,783)	0,696 (0,632-0,759)	0,695 (0,605-0,785)	0,392 (0,211-0,572)
L1 (Lasso)	0,761 (0,695-0,827)	0,708 (0,620-0,797)	0,708 (0,619-0,796)	0,710 (0,621-0,798)	0,706 (0,618-0,795)	0,708 (0,645-0,771)	0,707 (0,618-0,796)	0,416 (0,238-0,594)
DVM	0,745 (0,677-0,813)	0,697 (0,608-0,787)	0,696 (0,606-0,786)	0,699 (0,610-0,788)	0,694 (0,604-0,784)	0,696 (0,633-0,760)	0,696 (0,606-0,786)	0,393 (0,213-0,573)
Bayesci GLM	0,762 (0,696-0,828)	0,708 (0,620-0,797)	0,709 (0,620-0,798)	0,711 (0,623-0,799)	0,706 (0,617-0,795)	0,709 (0,646-0,771)	0,708 (0,618-0,797)	0,417 (0,239-0,595)



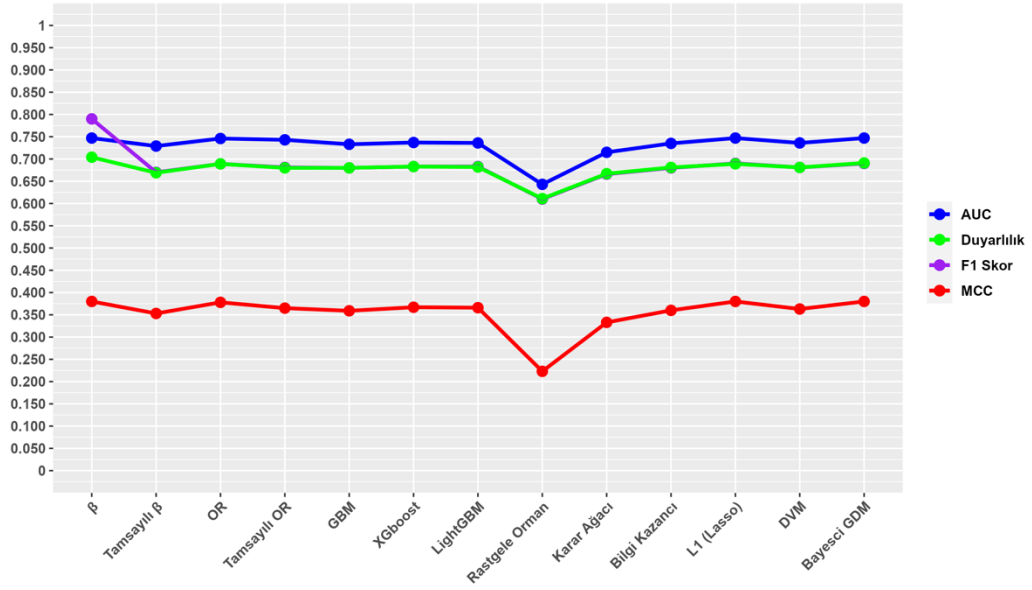
Şekil 3.16. LR ve ML algoritmaları ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=200; 1-0 oranı:%50-%50 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.17’de verilmiştir.

Çizelge 3.17. Senaryo A; n=500; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,747 (0,704-0,790)	0,704 (0,632-0,746)	0,790 (0,633-0,748)	0,704 (0,635-0,749)	0,790 (0,630-0,745)	0,704 (0,649-0,730)	0,790 (0,632-0,747)	0,380 (0,265-0,494)
Tam sayılı Beta	0,729 (0,687-0,771)	0,669 (0,612-0,725)	0,680 (0,623-0,736)	0,684 (0,626-0,743)	0,674 (0,616-0,732)	0,675 (0,634-0,716)	0,670 (0,611-0,729)	0,353 (0,235-0,471)
OR	0,746 (0,703-0,789)	0,689 (0,632-0,746)	0,689 (0,632-0,747)	0,691 (0,634-0,749)	0,686 (0,629-0,744)	0,689 (0,649-0,730)	0,689 (0,631-0,746)	0,378 (0,263-0,493)
Tam sayılı OR	0,743 (0,700-0,785)	0,680 (0,623-0,737)	0,684 (0,627-0,742)	0,687 (0,629-0,744)	0,679 (0,621-0,737)	0,683 (0,642-0,723)	0,681 (0,623-0,739)	0,365 (0,249-0,481)
GBM	0,733 (0,689-0,777)	0,680 (0,622-0,737)	0,679 (0,621-0,737)	0,683 (0,625-0,740)	0,676 (0,618-0,734)	0,679 (0,639-0,72)	0,680 (0,622-0,737)	0,359 (0,243-0,475)
XGBoost	0,737 (0,694-0,781)	0,683 (0,626-0,741)	0,684 (0,626-0,741)	0,686 (0,629-0,743)	0,681 (0,623-0,738)	0,683 (0,643-0,724)	0,683 (0,626-0,741)	0,367 (0,251-0,482)
LightGBM	0,736 (0,693-0,780)	0,682 (0,625-0,740)	0,684 (0,627-0,742)	0,686 (0,629-0,744)	0,680 (0,622-0,738)	0,684 (0,643-0,724)	0,683 (0,625-0,741)	0,366 (0,251-0,482)
Rastgele Orman	0,643 (0,596-0,691)	0,611 (0,551-0,670)	0,612 (0,552-0,672)	0,614 (0,554-0,674)	0,609 (0,549-0,669)	0,611 (0,569-0,654)	0,610 (0,550-0,670)	0,223 (0,103-0,343)
Karar Ağacı	0,715 (0,671-0,760)	0,667 (0,609-0,725)	0,666 (0,608-0,724)	0,669 (0,611-0,727)	0,664 (0,606-0,723)	0,667 (0,625-0,708)	0,666 (0,608-0,725)	0,333 (0,216-0,450)
Bilgi Kazancı	0,735 (0,691-0,778)	0,681 (0,623-0,738)	0,680 (0,622-0,738)	0,683 (0,625-0,740)	0,678 (0,620-0,736)	0,680 (0,639-0,721)	0,680 (0,622-0,738)	0,360 (0,245-0,476)
L1 (Lasso)	0,747 (0,704-0,789)	0,689 (0,632-0,747)	0,690 (0,633-0,748)	0,692 (0,635-0,749)	0,687 (0,630-0,744)	0,690 (0,649-0,730)	0,690 (0,632-0,747)	0,380 (0,265-0,494)
DVM	0,736 (0,692-0,779)	0,681 (0,624-0,739)	0,682 (0,624-0,740)	0,684 (0,627-0,741)	0,679 (0,622-0,737)	0,682 (0,641-0,722)	0,681 (0,623-0,739)	0,363 (0,248-0,479)
Bayesci GDM	0,747 (0,705-0,790)	0,691 (0,633-0,748)	0,690 (0,632-0,747)	0,692 (0,635-0,749)	0,688 (0,630-0,745)	0,690 (0,650-0,731)	0,690 (0,633-0,747)	0,380 (0,265-0,495)



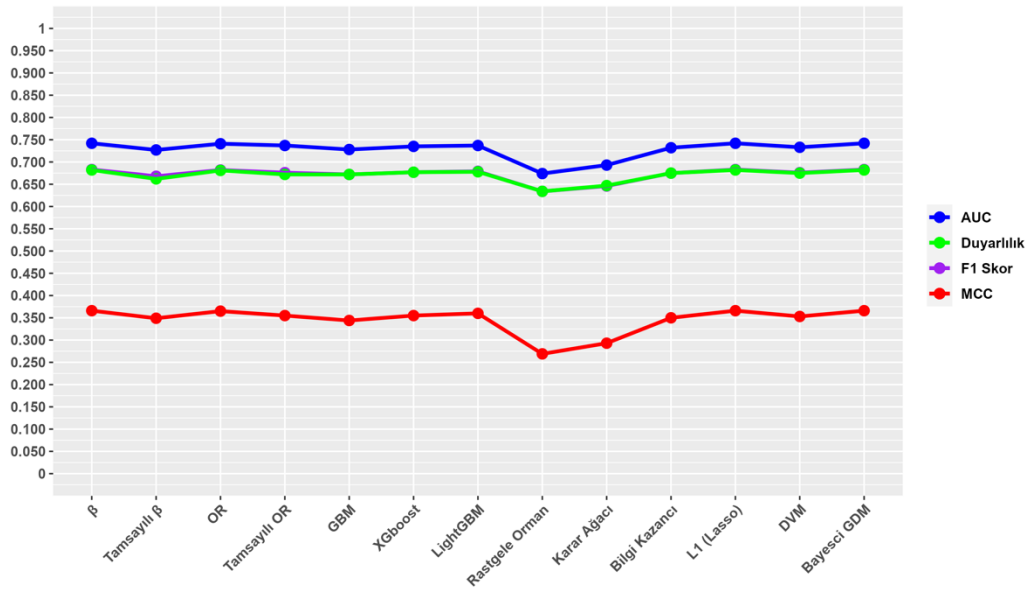
Şekil 3.17. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=500; 1-0 oranı:%50-%50 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu A senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişken 1-0 oranının rastgele olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.18’de verilmiştir.

Çizelge 3.18. Senaryo A; n=1000; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,742 (0,712-0,772)	0,682 (0,641-0,722)	0,684 (0,644-0,725)	0,686 (0,645-0,726)	0,680 (0,639-0,720)	0,683 (0,654-0,712)	0,683 (0,642-0,724)	0,366 (0,284-0,447)
Tam sayılı Beta	0,727 (0,698-0,757)	0,662 (0,622-0,703)	0,682 (0,642-0,722)	0,683 (0,642-0,724)	0,671 (0,630-0,712)	0,674 (0,645-0,703)	0,668 (0,626-0,710)	0,349 (0,265-0,433)
OR	0,741 (0,711-0,772)	0,681 (0,640-0,722)	0,684 (0,643-0,725)	0,685 (0,644-0,725)	0,680 (0,639-0,721)	0,683 (0,654-0,712)	0,682 (0,641-0,723)	0,365 (0,283-0,447)
Tam sayılı OR	0,737 (0,707-0,767)	0,672 (0,631-0,713)	0,683 (0,642-0,723)	0,682 (0,641-0,723)	0,674 (0,633-0,715)	0,678 (0,649-0,707)	0,676 (0,634-0,717)	0,355 (0,273-0,438)
GBM	0,728 (0,697-0,759)	0,672 (0,631-0,713)	0,672 (0,633-0,715)	0,674 (0,633-0,715)	0,669 (0,628-0,711)	0,672 (0,643-0,701)	0,672 (0,631-0,713)	0,344 (0,261-0,426)
XGBoost	0,735 (0,704-0,766)	0,695 (0,606-0,785)	0,695 (0,605-0,785)	0,698 (0,608-0,787)	0,693 (0,603-0,783)	0,695 (0,632-0,759)	0,694 (0,604-0,784)	0,391 (0,210-0,571)
LightGBM	0,737 (0,706-0,767)	0,678 (0,637-0,719)	0,682 (0,641-0,723)	0,682 (0,642-0,723)	0,677 (0,636-0,718)	0,680 (0,651-0,709)	0,679 (0,638-0,720)	0,360 (0,278-0,442)
Rastgele Orman	0,674 (0,641-0,707)	0,634 (0,592-0,676)	0,634 (0,592-0,676)	0,637 (0,595-0,679)	0,631 (0,589-0,673)	0,634 (0,604-0,664)	0,634 (0,592-0,676)	0,269 (0,184-0,353)
Karar Ağacı	0,693 (0,660-0,725)	0,647 (0,605-0,689)	0,645 (0,603-0,687)	0,649 (0,607-0,690)	0,645 (0,603-0,687)	0,646 (0,617-0,676)	0,646 (0,604-0,688)	0,293 (0,209-0,377)
Bilgi Kazancı	0,732 (0,701-0,763)	0,675 (0,634-0,716)	0,675 (0,634-0,716)	0,677 (0,636-0,718)	0,672 (0,631-0,713)	0,675 (0,646-0,704)	0,675 (0,634-0,716)	0,350 (0,268-0,432)
L1 (Lasso)	0,742 (0,712-0,772)	0,682 (0,642-0,723)	0,684 (0,643-0,725)	0,686 (0,646-0,727)	0,679 (0,638-0,720)	0,683 (0,654-0,712)	0,683 (0,643-0,724)	0,366 (0,284-0,447)
DVM	0,733 (0,702-0,764)	0,675 (0,634-0,716)	0,678 (0,637-0,719)	0,680 (0,639-0,720)	0,673 (0,632-0,714)	0,677 (0,648-0,706)	0,676 (0,635-0,717)	0,353 (0,271-0,435)
Bayesci GDM	0,742 (0,712-0,772)	0,682 (0,641-0,723)	0,684 (0,643-0,725)	0,686 (0,645-0,727)	0,680 (0,639-0,721)	0,683 (0,655-0,712)	0,683 (0,642-0,724)	0,366 (0,284-0,448)



Şekil 3.18. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo A; n=1000; 1-0 oranı:%50-%50 ve rastgele)

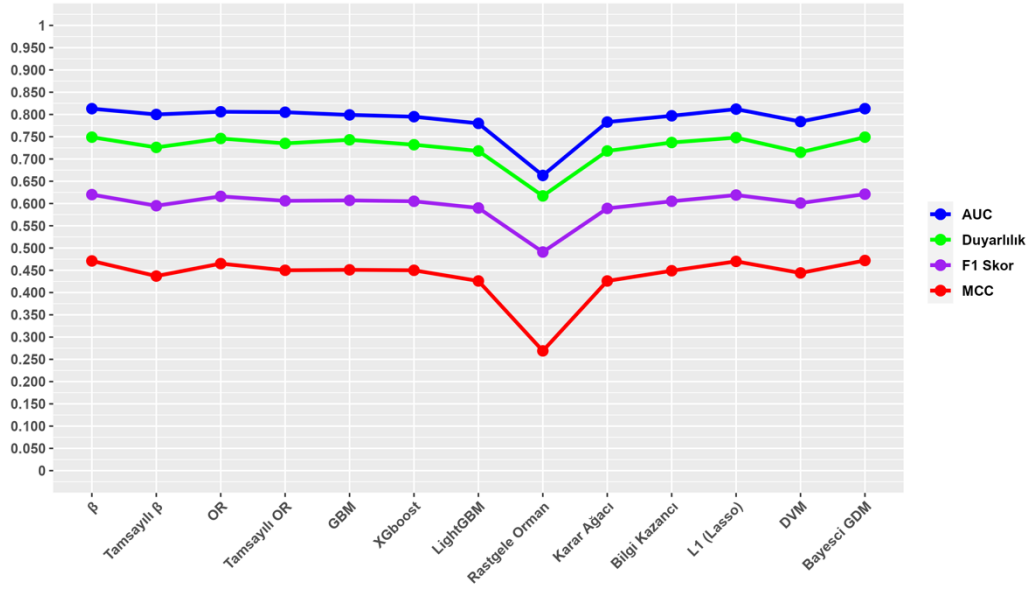
AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağacı sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.7. Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve %20-%80 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %20-%80 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetilmiştir. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalarla dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.1.19'da verilmiştir.

Çizelge 3.19. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,813 (0,743-0,883)	0,749 (0,631-0,866)	0,767 (0,700-0,835)	0,534 (0,420-0,647)	0,897 (0,845-0,950)	0,763 (0,704-0,821)	0,620 (0,509-0,731)	0,471 (0,269-0,673)
Tam sayılı Beta	0,800 (0,730-0,869)	0,726 (0,607-0,846)	0,750 (0,681-0,818)	0,514 (0,400-0,627)	0,888 (0,833-0,943)	0,744 (0,684-0,804)	0,595 (0,482-0,707)	0,437 (0,230-0,643)
OR	0,806 (0,735-0,877)	0,746 (0,629-0,864)	0,763 (0,696-0,831)	0,529 (0,415-0,642)	0,896 (0,843-0,948)	0,759 (0,700-0,818)	0,616 (0,505-0,727)	0,465 (0,262-0,667)
Tam sayılı OR	0,805 (0,734-0,876)	0,735 (0,615-0,854)	0,758 (0,690-0,827)	0,521 (0,407-0,634)	0,891 (0,836-0,945)	0,752 (0,693-0,812)	0,606 (0,494-0,718)	0,450 (0,246-0,654)
GBM	0,799 (0,728-0,870)	0,743 (0,625-0,861)	0,753 (0,684-0,821)	0,517 (0,405-0,630)	0,893 (0,839-0,947)	0,750 (0,690-0,810)	0,607 (0,496-0,717)	0,451 (0,248-0,653)
XGBoost	0,795 (0,722-0,868)	0,732 (0,612-0,852)	0,761 (0,693-0,829)	0,522 (0,407-0,636)	0,890 (0,836-0,944)	0,754 (0,694-0,813)	0,605 (0,493-0,717)	0,450 (0,245-0,655)
LightGBM	0,780 (0,705-0,856)	0,718 (0,597-0,840)	0,749 (0,680-0,817)	0,506 (0,393-0,620)	0,883 (0,827-0,939)	0,741 (0,680-0,801)	0,590 (0,478-0,702)	0,426 (0,220-0,632)
Rastgele Orman	0,663 (0,573-0,754)	0,617 (0,488-0,747)	0,676 (0,603-0,748)	0,417 (0,310-0,524)	0,830 (0,763-0,897)	0,660 (0,596-0,724)	0,491 (0,382-0,600)	0,269 (0,058-0,480)
Karar Ağacı	0,783 (0,709-0,857)	0,718 (0,596-0,839)	0,749 (0,680-0,817)	0,506 (0,393-0,620)	0,883 (0,828-0,939)	0,741 (0,680-0,801)	0,589 (0,477-0,701)	0,426 (0,219-0,632)
Bilgi Kazancı	0,797 (0,725-0,870)	0,737 (0,618-0,856)	0,756 (0,688-0,825)	0,518 (0,405-0,632)	0,891 (0,837-0,945)	0,751 (0,692-0,811)	0,605 (0,494-0,716)	0,449 (0,245-0,653)
L1 (Lasso)	0,812 (0,742-0,883)	0,748 (0,630-0,865)	0,767 (0,700-0,834)	0,533 (0,419-0,647)	0,897 (0,844-0,949)	0,762 (0,703-0,821)	0,619 (0,508-0,730)	0,470 (0,267-0,672)
DVM	0,784 (0,708-0,860)	0,715 (0,593-0,837)	0,768 (0,701-0,835)	0,524 (0,408-0,639)	0,885 (0,830-0,940)	0,754 (0,695-0,813)	0,601 (0,488-0,715)	0,444 (0,236-0,652)
Bayesci GDM	0,813 (0,743-0,883)	0,749 (0,632-0,866)	0,768 (0,701-0,835)	0,535 (0,421-0,649)	0,897 (0,844-0,949)	0,763 (0,704-0,821)	0,621 (0,510-0,732)	0,472 (0,270-0,674)



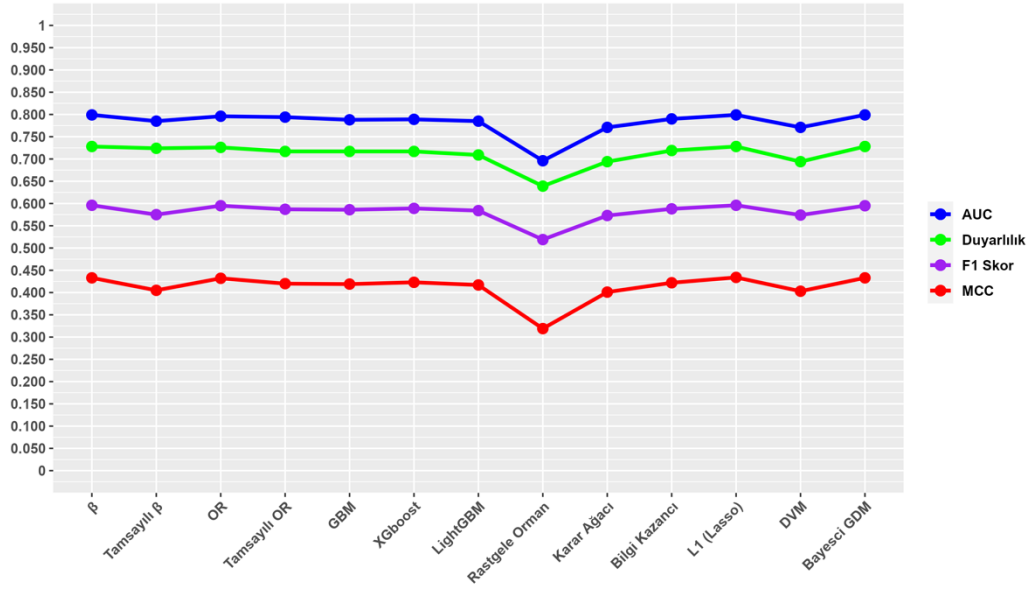
Şekil 3.19. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%20-%80 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %20-%80 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.20’de verilmiştir.

Çizelge 3.20. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,799 (0,753-0,844)	0,728 (0,651-0,804)	0,750 (0,706-0,794)	0,507 (0,435-0,578)	0,887 (0,852-0,922)	0,744 (0,706-0,782)	0,596 (0,525-0,666)	0,433 (0,304-0,563)
Tam sayılı Beta	0,785 (0,739-0,830)	0,724 (0,648-0,800)	0,721 (0,677-0,766)	0,486 (0,416-0,556)	0,884 (0,847-0,920)	0,722 (0,683-0,761)	0,575 (0,506-0,645)	0,405 (0,276-0,534)
OR	0,796 (0,750-0,842)	0,726 (0,649-0,802)	0,749 (0,706-0,793)	0,506 (0,434-0,578)	0,886 (0,851-0,922)	0,743 (0,705-0,782)	0,595 (0,524-0,665)	0,432 (0,302-0,561)
Tam sayılı OR	0,794 (0,748-0,840)	0,717 (0,640-0,794)	0,746 (0,702-0,790)	0,500 (0,428-0,571)	0,883 (0,847-0,918)	0,738 (0,700-0,777)	0,587 (0,516-0,657)	0,420 (0,290-0,551)
GBM	0,788 (0,742-0,834)	0,717 (0,640-0,794)	0,745 (0,701-0,789)	0,499 (0,427-0,570)	0,883 (0,847-0,918)	0,738 (0,699-0,776)	0,586 (0,516-0,657)	0,419 (0,289-0,550)
XGBoost	0,789 (0,742-0,835)	0,717 (0,640-0,794)	0,748 (0,704-0,792)	0,502 (0,430-0,573)	0,883 (0,848-0,919)	0,740 (0,702-0,779)	0,589 (0,518-0,659)	0,423 (0,293-0,553)
LightGBM	0,785 (0,738-0,833)	0,709 (0,632-0,787)	0,749 (0,705-0,793)	0,499 (0,427-0,571)	0,881 (0,845-0,916)	0,738 (0,700-0,777)	0,584 (0,513-0,655)	0,417 (0,286-0,548)
Rastgele Orman	0,696 (0,641-0,752)	0,639 (0,557-0,720)	0,711 (0,665-0,756)	0,443 (0,373-0,514)	0,848 (0,809-0,888)	0,692 (0,652-0,732)	0,519 (0,448-0,591)	0,319 (0,183-0,454)
Karar Ağacı	0,771 (0,723-0,820)	0,694 (0,615-0,773)	0,745 (0,701-0,789)	0,492 (0,420-0,564)	0,875 (0,838-0,911)	0,732 (0,693-0,771)	0,573 (0,502-0,645)	0,401 (0,268-0,534)
Bilgi Kazancı	0,790 (0,744-0,836)	0,719 (0,642-0,796)	0,746 (0,702-0,790)	0,500 (0,428-0,571)	0,884 (0,848-0,919)	0,739 (0,701-0,777)	0,588 (0,517-0,659)	0,422 (0,292-0,552)
L1 (Lasso)	0,799 (0,753-0,845)	0,728 (0,651-0,804)	0,750 (0,706-0,794)	0,507 (0,435-0,578)	0,887 (0,852-0,922)	0,744 (0,706-0,782)	0,596 (0,525-0,666)	0,434 (0,304-0,563)
DVM	0,771 (0,722-0,820)	0,694 (0,615-0,773)	0,747 (0,703-0,791)	0,493 (0,421-0,565)	0,875 (0,839-0,911)	0,734 (0,695-0,772)	0,574 (0,503-0,646)	0,403 (0,270-0,536)
Bayesian GDM	0,799 (0,753-0,845)	0,728 (0,652-0,804)	0,749 (0,705-0,793)	0,506 (0,434-0,577)	0,887 (0,852-0,922)	0,743 (0,705-0,782)	0,595 (0,525-0,666)	0,433 (0,303-0,562)



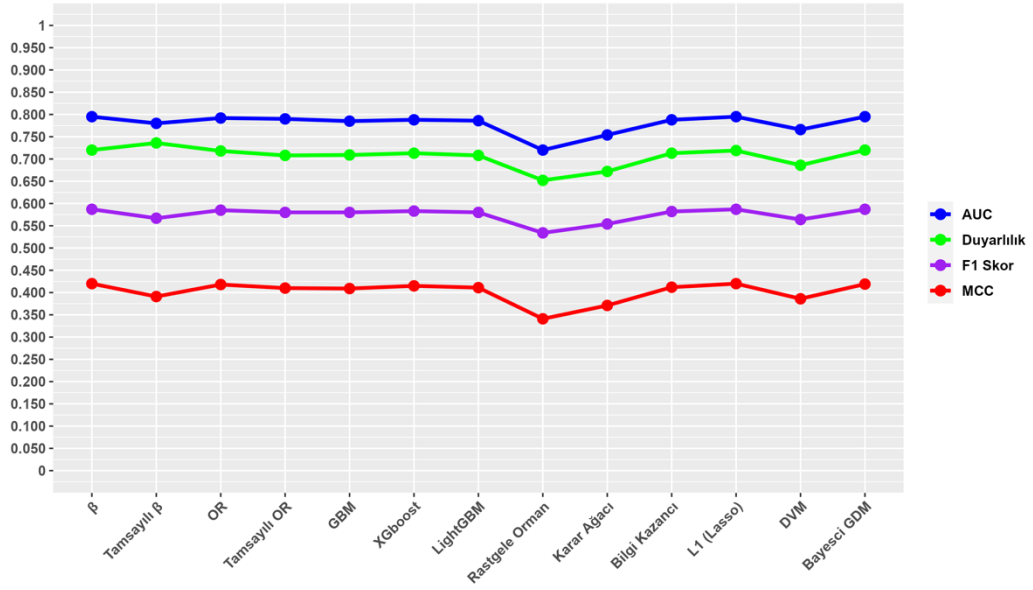
Şekil 3.20. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%20-%80 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişkenlerin 1-0 oranı: %20-%80 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.21’de verilmiştir.

Çizelge 3.21. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,795 (0,762-0,827)	0,720 (0,665-0,774)	0,744 (0,712-0,775)	0,497 (0,447-0,548)	0,884 (0,858-0,909)	0,737 (0,710-0,765)	0,587 (0,537-0,637)	0,420 (0,328-0,512)
Tam sayılı Beta	0,780 (0,747-0,812)	0,736 (0,683-0,789)	0,697 (0,665-0,730)	0,469 (0,420-0,517)	0,885 (0,859-0,910)	0,707 (0,679-0,735)	0,567 (0,518-0,615)	0,391 (0,301-0,480)
OR	0,792 (0,760-0,825)	0,718 (0,663-0,773)	0,743 (0,712-0,774)	0,496 (0,445-0,546)	0,883 (0,858-0,908)	0,736 (0,709-0,764)	0,585 (0,536-0,635)	0,418 (0,326-0,510)
Tam sayılı OR	0,790 (0,758-0,823)	0,708 (0,653-0,763)	0,743 (0,712-0,774)	0,493 (0,442-0,544)	0,879 (0,854-0,905)	0,734 (0,706-0,761)	0,580 (0,530-0,630)	0,410 (0,317-0,502)
GBM	0,785 (0,752-0,818)	0,709 (0,654-0,765)	0,742 (0,710-0,773)	0,492 (0,441-0,542)	0,879 (0,854-0,905)	0,733 (0,706-0,761)	0,580 (0,530-0,630)	0,409 (0,317-0,501)
XGBoost	0,788 (0,755-0,821)	0,713 (0,658-0,768)	0,744 (0,713-0,775)	0,495 (0,445-0,546)	0,881 (0,856-0,906)	0,736 (0,709-0,763)	0,583 (0,534-0,633)	0,415 (0,322-0,507)
LightGBM	0,786 (0,753-0,819)	0,708 (0,652-0,763)	0,744 (0,713-0,776)	0,493 (0,443-0,545)	0,879 (0,854-0,905)	0,735 (0,708-0,762)	0,580 (0,531-0,631)	0,411 (0,318-0,503)
Rastgele Orman	0,720 (0,682-0,759)	0,652 (0,595-0,710)	0,722 (0,690-0,754)	0,455 (0,404-0,505)	0,856 (0,828-0,883)	0,704 (0,676-0,732)	0,534 (0,483-0,584)	0,341 (0,245-0,436)
Karar Ağacı	0,754 (0,720-0,790)	0,672 (0,615-0,729)	0,735 (0,704-0,767)	0,473 (0,423-0,524)	0,865 (0,839-0,892)	0,719 (0,691-0,747)	0,554 (0,503-0,605)	0,371 (0,276-0,466)
Bilgi Kazancı	0,788 (0,755-0,821)	0,713 (0,658-0,768)	0,742 (0,710-0,773)	0,493 (0,442-0,543)	0,881 (0,855-0,906)	0,734 (0,707-0,761)	0,582 (0,532-0,632)	0,412 (0,320-0,504)
L1 (Lasso)	0,795 (0,762-0,827)	0,719 (0,665-0,774)	0,744 (0,713-0,775)	0,497 (0,447-0,548)	0,883 (0,858-0,909)	0,738 (0,710-0,765)	0,587 (0,537-0,637)	0,420 (0,328-0,512)
DVM	0,766 (0,731-0,801)	0,686 (0,630-0,742)	0,738 (0,707-0,770)	0,481 (0,430-0,531)	0,870 (0,844-0,897)	0,725 (0,697-0,752)	0,564 (0,513-0,614)	0,386 (0,292-0,480)
Bayesci GDM	0,795 (0,762-0,827)	0,720 (0,665-0,774)	0,743 (0,712-0,775)	0,497 (0,446-0,547)	0,884 (0,858-0,909)	0,737 (0,710-0,764)	0,587 (0,537-0,637)	0,419 (0,328-0,511)



Şekil 3.21. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=1000; 1-0 oranı:%20-%80 ve %20-%80)

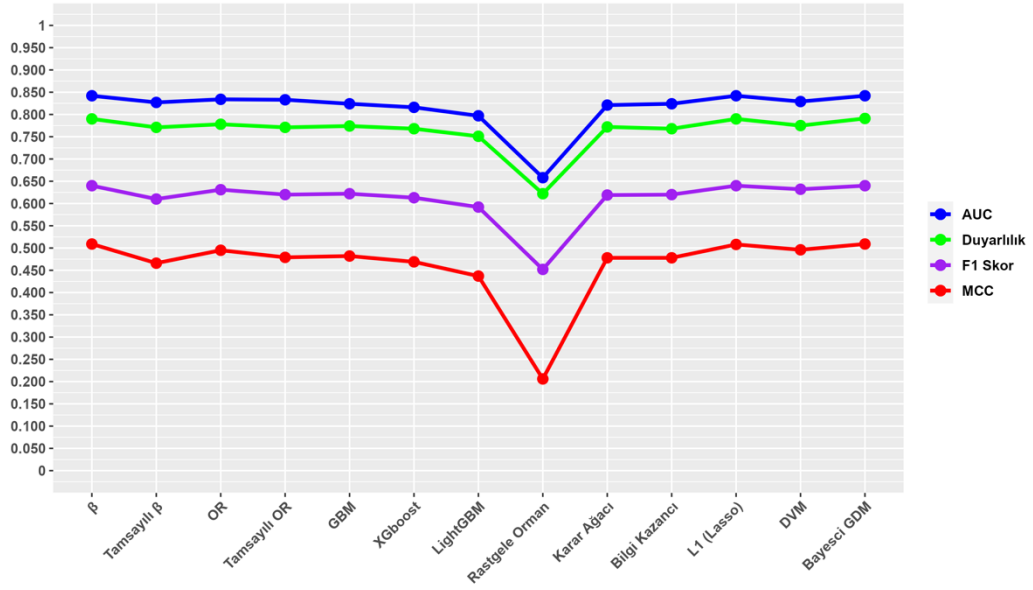
AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazanci, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanınve karar ağacının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.8 Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve %50-%50 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %50-%50 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetildi. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.22’de verilmiştir.

Çizelge 3.22. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,842 (0,781-0,903)	0,790 (0,678-0,902)	0,774 (0,708-0,841)	0,542 (0,428-0,656)	0,917 (0,869-0,965)	0,778 (0,721-0,836)	0,640 (0,530-0,750)	0,509 (0,311-0,706)
Tam sayılı Beta	0,827 (0,765-0,889)	0,771 (0,656-0,885)	0,747 (0,679-0,815)	0,512 (0,400-0,625)	0,908 (0,858-0,958)	0,753 (0,694-0,812)	0,610 (0,500-0,720)	0,466 (0,266-0,666)
OR	0,834 (0,772-0,897)	0,778 (0,663-0,892)	0,771 (0,705-0,838)	0,535 (0,421-0,649)	0,912 (0,863-0,961)	0,773 (0,715-0,831)	0,631 (0,520-0,742)	0,495 (0,296-0,695)
Tam sayılı OR	0,833 (0,770-0,896)	0,771 (0,655-0,887)	0,761 (0,694-0,829)	0,523 (0,410-0,636)	0,909 (0,859-0,959)	0,764 (0,705-0,822)	0,620 (0,510-0,730)	0,479 (0,279-0,679)
GBM	0,824 (0,759-0,890)	0,774 (0,659-0,889)	0,763 (0,696-0,830)	0,524 (0,411-0,637)	0,910 (0,861-0,960)	0,766 (0,707-0,824)	0,622 (0,511-0,732)	0,482 (0,283-0,682)
XGBoost	0,816 (0,750-0,882)	0,768 (0,653-0,884)	0,754 (0,686-0,822)	0,515 (0,402-0,627)	0,907 (0,856-0,957)	0,758 (0,699-0,817)	0,613 (0,503-0,723)	0,469 (0,270-0,669)
LightGBM	0,797 (0,728-0,866)	0,751 (0,632-0,870)	0,738 (0,668-0,807)	0,493 (0,382-0,605)	0,899 (0,846-0,952)	0,741 (0,681-0,801)	0,592 (0,482-0,702)	0,437 (0,236-0,639)
Rastgele Orman	0,658 (0,573-0,742)	0,622 (0,493-0,751)	0,610 (0,534-0,685)	0,360 (0,262-0,458)	0,823 (0,753-0,893)	0,613 (0,547-0,679)	0,452 (0,350-0,553)	0,206 (0,005-0,406)
Karar Ağacı	0,821 (0,755-0,887)	0,772 (0,657-0,887)	0,760 (0,692-0,828)	0,522 (0,409-0,634)	0,909 (0,859-0,959)	0,763 (0,704-0,822)	0,619 (0,509-0,729)	0,478 (0,279-0,678)
Bilgi Kazancı	0,824 (0,760-0,889)	0,768 (0,652-0,884)	0,763 (0,696-0,831)	0,524 (0,410-0,637)	0,908 (0,857-0,958)	0,764 (0,706-0,823)	0,620 (0,509-0,731)	0,478 (0,278-0,679)
L1 (Lasso)	0,842 (0,781-0,903)	0,790 (0,678-0,902)	0,774 (0,708-0,840)	0,542 (0,428-0,656)	0,917 (0,869-0,964)	0,778 (0,721-0,835)	0,640 (0,530-0,750)	0,508 (0,311-0,706)
DVM	0,829 (0,764-0,894)	0,775 (0,660-0,889)	0,774 (0,708-0,841)	0,538 (0,424-0,652)	0,911 (0,863-0,960)	0,775 (0,717-0,832)	0,632 (0,521-0,743)	0,496 (0,296-0,696)
Bayesci GDM	0,842 (0,781-0,903)	0,791 (0,680-0,903)	0,773 (0,707-0,84)	0,542 (0,429-0,656)	0,917 (0,869-0,965)	0,778 (0,721-0,835)	0,640 (0,531-0,75)	0,509 (0,312-0,706)



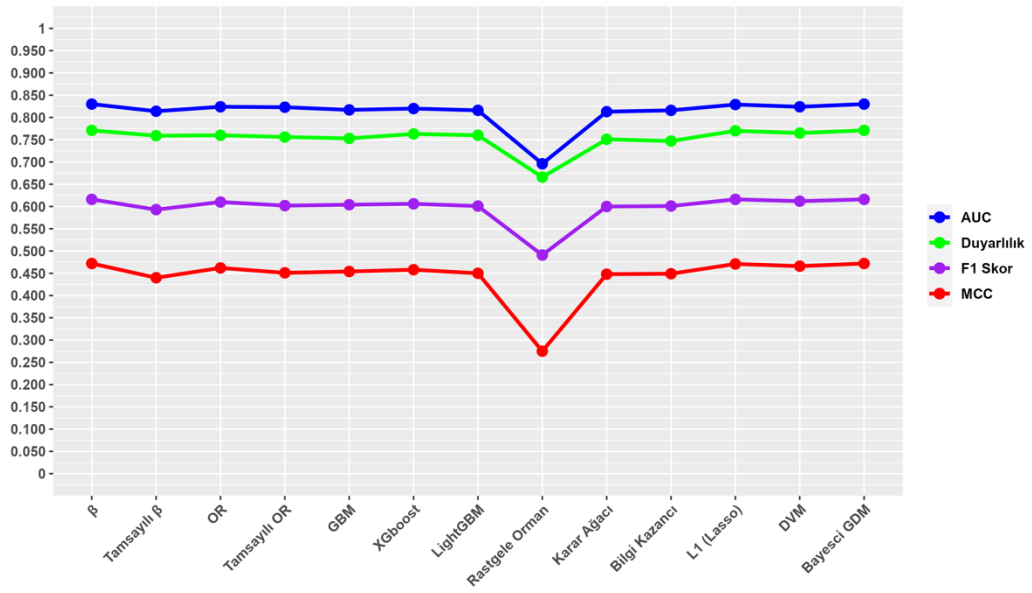
Şekil 3.22. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%20-%80 ve : %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %20-%80 ve bağımsız değişken 1-0 oranının %50-%50 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.1.23’de verilmiştir.

Çizelge 3.23. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
beta	0,830 (0,790-0,870)	0,771 (0,698-0,845)	0,755 (0,712-0,799)	0,515 (0,444-0,586)	0,908 (0,876-0,940)	0,759 (0,722-0,797)	0,616 (0,547-0,686)	0,472 (0,346-0,598)
Tam sayılı Beta	0,814 (0,774-0,855)	0,759 (0,686-0,833)	0,731 (0,687-0,775)	0,494 (0,423-0,564)	0,902 (0,869-0,935)	0,738 (0,700-0,777)	0,593 (0,524-0,663)	0,440 (0,313-0,567)
OR	0,824 (0,784-0,865)	0,760 (0,686-0,835)	0,755 (0,712-0,798)	0,511 (0,439-0,582)	0,904 (0,872-0,937)	0,756 (0,719-0,794)	0,610 (0,540-0,680)	0,462 (0,335-0,590)
Tam sayılı OR	0,823 (0,782-0,864)	0,756 (0,681-0,831)	0,748 (0,704-0,791)	0,502 (0,431-0,574)	0,902 (0,869-0,935)	0,750 (0,712-0,787)	0,602 (0,532-0,672)	0,451 (0,323-0,578)
GBM	0,817 (0,775-0,859)	0,753 (0,678-0,828)	0,753 (0,709-0,796)	0,506 (0,434-0,577)	0,901 (0,868-0,934)	0,753 (0,715-0,790)	0,604 (0,533-0,674)	0,454 (0,326-0,582)
XGBoost	0,820 (0,779-0,861)	0,763 (0,689-0,838)	0,748 (0,705-0,792)	0,505 (0,434-0,576)	0,905 (0,872-0,937)	0,752 (0,714-0,790)	0,606 (0,537-0,676)	0,458 (0,331-0,584)
LightGBM	0,816 (0,774-0,857)	0,760 (0,685-0,834)	0,744 (0,700-0,788)	0,500 (0,429-0,571)	0,903 (0,870-0,936)	0,748 (0,710-0,786)	0,601 (0,532-0,671)	0,450 (0,324-0,577)
Rastgele Orman	0,696 (0,645-0,746)	0,666 (0,586-0,747)	0,645 (0,598-0,693)	0,392 (0,328-0,457)	0,851 (0,810-0,892)	0,650 (0,609-0,692)	0,491 (0,425-0,558)	0,275 (0,148-0,402)
Karar Ağacı	0,813 (0,771-0,856)	0,751 (0,676-0,827)	0,749 (0,705-0,793)	0,502 (0,43-0,573)	0,900 (0,867-0,933)	0,749 (0,712-0,787)	0,600 (0,53-0,67)	0,448 (0,32-0,576)
Bilgi Kazancı	0,816 (0,774-0,857)	0,747 (0,671-0,823)	0,753 (0,710-0,797)	0,505 (0,433-0,576)	0,899 (0,866-0,932)	0,752 (0,714-0,789)	0,601 (0,531-0,671)	0,449 (0,321-0,578)
L1 (Lasso)	0,829 (0,789-0,869)	0,770 (0,697-0,844)	0,756 (0,712-0,799)	0,515 (0,443-0,586)	0,908 (0,876-0,940)	0,759 (0,722-0,797)	0,616 (0,546-0,685)	0,471 (0,345-0,598)
DVM	0,824 (0,782-0,865)	0,765 (0,691-0,839)	0,755 (0,712-0,798)	0,512 (0,440-0,583)	0,906 (0,874-0,938)	0,758 (0,720-0,795)	0,612 (0,542-0,682)	0,466 (0,339-0,593)
Bayesci GDM	0,830 (0,790-0,870)	0,771 (0,698-0,845)	0,756 (0,712-0,799)	0,515 (0,444-0,587)	0,908 (0,876-0,94)	0,760 (0,722-0,797)	0,616 (0,547-0,686)	0,472 (0,346-0,598)



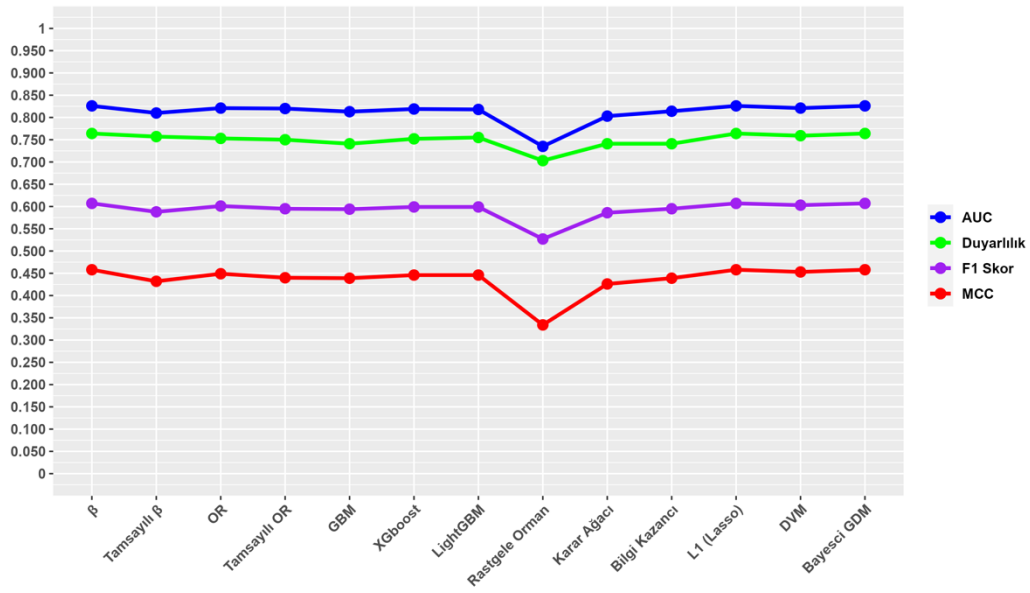
Şekil 3.23. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%20-%80 ve %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişkenlerin 1-0 oranı: %50-%50 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.24'te verilmiştir.

Çizelge 3.24. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %20-%80 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,826 (0,797-0,854)	0,764 (0,711-0,816)	0,748 (0,717-0,779)	0,504 (0,454-0,555)	0,905 (0,882-0,928)	0,752 (0,726-0,779)	0,607 (0,558-0,656)	0,458 (0,368-0,547)
Tam sayılı Beta	0,810 (0,781-0,839)	0,757 (0,705-0,810)	0,725 (0,694-0,757)	0,486 (0,436-0,535)	0,901 (0,877-0,925)	0,733 (0,706-0,761)	0,588 (0,539-0,637)	0,432 (0,342-0,522)
OR	0,821 (0,792-0,850)	0,753 (0,700-0,806)	0,749 (0,718-0,780)	0,502 (0,451-0,552)	0,901 (0,877-0,924)	0,750 (0,723-0,777)	0,601 (0,552-0,651)	0,449 (0,359-0,540)
Tam sayılı OR	0,820 (0,791-0,849)	0,750 (0,696-0,803)	0,743 (0,712-0,774)	0,495 (0,445-0,545)	0,899 (0,875-0,923)	0,745 (0,718-0,772)	0,595 (0,546-0,645)	0,440 (0,350-0,531)
GBM	0,813 (0,783-0,843)	0,741 (0,687-0,796)	0,749 (0,718-0,780)	0,497 (0,446-0,548)	0,897 (0,873-0,921)	0,747 (0,720-0,774)	0,594 (0,544-0,644)	0,439 (0,348-0,530)
XGBoost	0,819 (0,789-0,848)	0,752 (0,699-0,806)	0,746 (0,715-0,777)	0,498 (0,448-0,549)	0,900 (0,877-0,924)	0,748 (0,721-0,775)	0,599 (0,549-0,648)	0,446 (0,356-0,536)
LightGBM	0,818 (0,789-0,848)	0,755 (0,701-0,808)	0,745 (0,714-0,776)	0,498 (0,448-0,548)	0,901 (0,878-0,924)	0,747 (0,720-0,774)	0,599 (0,550-0,648)	0,446 (0,356-0,536)
Rastgele Orman	0,735 (0,701-0,770)	0,703 (0,647-0,759)	0,676 (0,643-0,709)	0,424 (0,377-0,471)	0,871 (0,844-0,898)	0,683 (0,654-0,711)	0,527 (0,480-0,575)	0,334 (0,244-0,424)
Karar Ağacı	0,803 (0,772-0,834)	0,741 (0,687-0,795)	0,736 (0,705-0,768)	0,486 (0,436-0,536)	0,895 (0,871-0,919)	0,738 (0,71-0,765)	0,586 (0,536-0,635)	0,426 (0,336-0,517)
Bilgi Kazancı	0,814 (0,784-0,843)	0,741 (0,687-0,795)	0,749 (0,718-0,780)	0,498 (0,447-0,548)	0,897 (0,873-0,920)	0,747 (0,720-0,774)	0,595 (0,545-0,644)	0,439 (0,348-0,531)
L1 (Lasso)	0,826 (0,797-0,854)	0,764 (0,711-0,816)	0,749 (0,718-0,780)	0,505 (0,454-0,555)	0,905 (0,882-0,928)	0,753 (0,726-0,779)	0,607 (0,558-0,656)	0,458 (0,369-0,548)
DVM	0,821 (0,792-0,850)	0,759 (0,706-0,812)	0,748 (0,716-0,779)	0,502 (0,452-0,552)	0,903 (0,880-0,926)	0,750 (0,724-0,777)	0,603 (0,554-0,653)	0,453 (0,363-0,543)
Bayesci GDM	0,826 (0,797-0,854)	0,764 (0,711-0,816)	0,748 (0,717-0,779)	0,504 (0,454-0,555)	0,905 (0,882-0,928)	0,752 (0,725-0,779)	0,607 (0,558-0,656)	0,458 (0,368-0,547)



Şekil 3.24. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=1000; 1-0 oranı:%20-%80 ve : %50-%50)

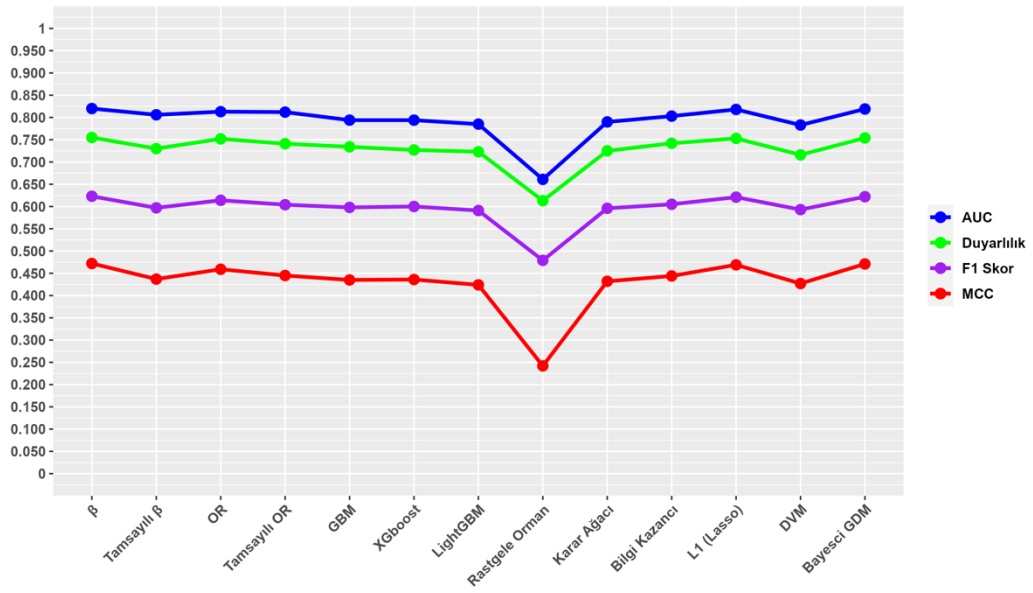
AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.9. Senaryo B, bağımlı ve bağımsız değişken oranı %20-%80 ve rastgele (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.25'te verilmiştir.

Çizelge 3.25. Senaryo B; n=200; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,820 (0,753-0,887)	0,755 (0,640-0,871)	0,762 (0,694-0,830)	0,536 (0,424-0,648)	0,896 (0,842-0,949)	0,760 (0,701-0,819)	0,623 (0,513-0,732)	0,472 (0,272-0,672)
Tam sayılı Beta	0,806 (0,739-0,872)	0,730 (0,613-0,848)	0,744 (0,675-0,813)	0,519 (0,407-0,631)	0,885 (0,829-0,941)	0,740 (0,680-0,801)	0,597 (0,487-0,708)	0,437 (0,232-0,641)
OR	0,813 (0,746-0,881)	0,752 (0,636-0,868)	0,753 (0,684-0,822)	0,526 (0,414-0,637)	0,893 (0,839-0,947)	0,753 (0,693-0,812)	0,614 (0,505-0,723)	0,459 (0,259-0,659)
Tam sayılı OR	0,812 (0,745-0,880)	0,741 (0,623-0,858)	0,749 (0,680-0,818)	0,517 (0,406-0,629)	0,889 (0,834-0,944)	0,747 (0,687-0,807)	0,604 (0,495-0,714)	0,445 (0,243-0,647)
GBM	0,794 (0,722-0,865)	0,734 (0,615-0,853)	0,745 (0,676-0,815)	0,512 (0,400-0,623)	0,885 (0,830-0,941)	0,742 (0,682-0,802)	0,598 (0,488-0,708)	0,435 (0,233-0,638)
XGBoost	0,794 (0,722-0,866)	0,727 (0,608-0,847)	0,748 (0,678-0,817)	0,518 (0,406-0,629)	0,883 (0,826-0,939)	0,743 (0,683-0,803)	0,600 (0,489-0,710)	0,436 (0,232-0,639)
LightGBM	0,785 (0,711-0,858)	0,723 (0,603-0,843)	0,741 (0,672-0,811)	0,508 (0,397-0,619)	0,880 (0,824-0,937)	0,736 (0,676-0,797)	0,591 (0,481-0,701)	0,424 (0,221-0,627)
Rastgele Orman	0,661 (0,574-0,748)	0,613 (0,485-0,742)	0,650 (0,575-0,724)	0,404 (0,301-0,507)	0,819 (0,749-0,888)	0,640 (0,575-0,704)	0,479 (0,373-0,585)	0,242 (0,035-0,448)
Karar Ağacı	0,790 (0,717-0,862)	0,725 (0,606-0,845)	0,747 (0,678-0,817)	0,514 (0,402-0,626)	0,883 (0,826-0,939)	0,742 (0,681-0,802)	0,596 (0,486-0,707)	0,432 (0,228-0,636)
Bilgi Kazancı	0,803 (0,734-0,873)	0,742 (0,624-0,860)	0,746 (0,677-0,816)	0,516 (0,405-0,627)	0,888 (0,833-0,943)	0,745 (0,685-0,805)	0,605 (0,495-0,714)	0,444 (0,243-0,645)
L1 (Lasso)	0,818 (0,751-0,885)	0,753 (0,637-0,869)	0,761 (0,693-0,829)	0,536 (0,424-0,648)	0,894 (0,841-0,947)	0,759 (0,700-0,818)	0,621 (0,512-0,731)	0,469 (0,269-0,669)
DVM	0,783 (0,709-0,857)	0,716 (0,596-0,837)	0,749 (0,680-0,818)	0,513 (0,401-0,625)	0,879 (0,823-0,936)	0,741 (0,680-0,801)	0,593 (0,482-0,704)	0,427 (0,221-0,632)
Bayesci GDM	0,819 (0,752-0,885)	0,754 (0,639-0,870)	0,761 (0,693-0,829)	0,536 (0,424-0,648)	0,894 (0,841-0,948)	0,759 (0,700-0,818)	0,622 (0,513-0,731)	0,471 (0,271-0,670)



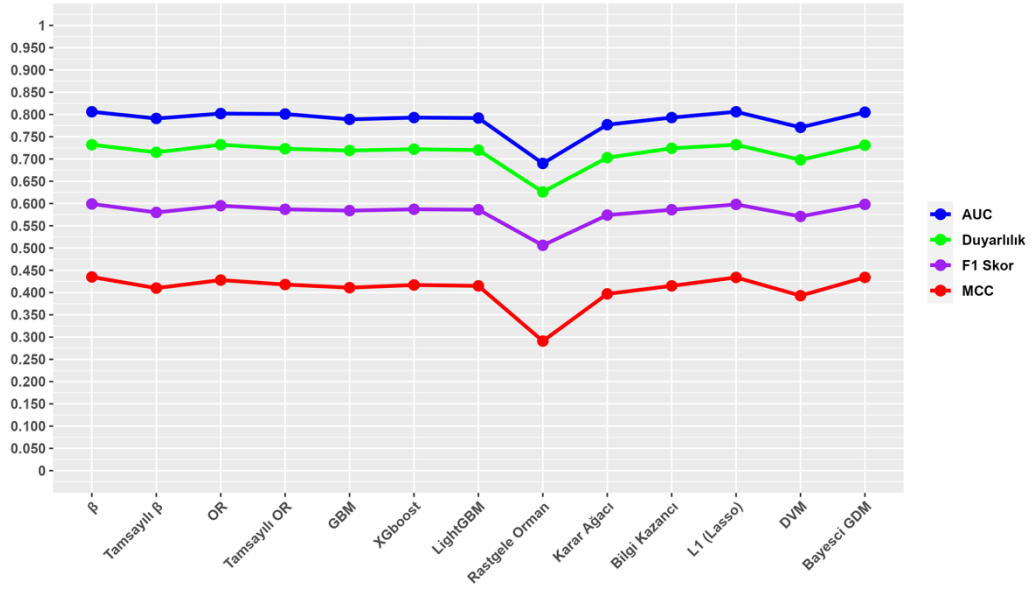
Şekil 3.25. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%20-%80 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.26'da verilmiştir.

Çizelge 3.26. Senaryo B; n=500; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,806 (0,762-0,850)	0,732 (0,657-0,807)	0,746 (0,702-0,790)	0,512 (0,442-0,583)	0,884 (0,849-0,920)	0,742 (0,704-0,780)	0,599 (0,530-0,669)	0,435 (0,307-0,563)
Tam sayılı Beta	0,791 (0,747-0,835)	0,715 (0,639-0,790)	0,732 (0,688-0,776)	0,500 (0,429-0,570)	0,878 (0,842-0,914)	0,727 (0,688-0,766)	0,580 (0,509-0,650)	0,410 (0,280-0,540)
OR	0,802 (0,759-0,846)	0,732 (0,656-0,807)	0,740 (0,695-0,784)	0,505 (0,435-0,575)	0,883 (0,848-0,919)	0,737 (0,699-0,776)	0,595 (0,526-0,664)	0,428 (0,300-0,555)
Tam sayılı OR	0,801 (0,757-0,844)	0,723 (0,647-0,799)	0,737 (0,693-0,782)	0,500 (0,430-0,571)	0,880 (0,844-0,916)	0,733 (0,695-0,772)	0,587 (0,518-0,657)	0,418 (0,289-0,547)
GBM	0,789 (0,743-0,835)	0,719 (0,643-0,796)	0,734 (0,689-0,779)	0,496 (0,426-0,566)	0,878 (0,841-0,914)	0,730 (0,691-0,769)	0,584 (0,515-0,654)	0,411 (0,283-0,540)
XGBoost	0,793 (0,748-0,839)	0,722 (0,646-0,798)	0,737 (0,692-0,782)	0,499 (0,429-0,569)	0,880 (0,843-0,916)	0,733 (0,694-0,772)	0,587 (0,518-0,657)	0,417 (0,288-0,545)
LightGBM	0,792 (0,747-0,838)	0,720 (0,643-0,796)	0,738 (0,693-0,782)	0,499 (0,429-0,569)	0,879 (0,843-0,915)	0,733 (0,694-0,772)	0,586 (0,516-0,656)	0,415 (0,286-0,544)
Rastgele Orman	0,690 (0,636-0,744)	0,626 (0,545-0,708)	0,690 (0,644-0,737)	0,433 (0,365-0,502)	0,835 (0,794-0,877)	0,673 (0,632-0,714)	0,506 (0,436-0,575)	0,291 (0,158-0,425)
Karar Ağacı	0,777 (0,730-0,824)	0,703 (0,625-0,780)	0,733 (0,688-0,778)	0,491 (0,420-0,561)	0,872 (0,835-0,909)	0,725 (0,686-0,764)	0,574 (0,503-0,644)	0,397 (0,267-0,528)
Bilgi Kazancı	0,793 (0,748-0,838)	0,724 (0,648-0,800)	0,734 (0,689-0,779)	0,496 (0,426-0,566)	0,880 (0,844-0,916)	0,731 (0,693-0,770)	0,586 (0,517-0,655)	0,415 (0,286-0,543)
L1 (Lasso)	0,806 (0,762-0,850)	0,732 (0,657-0,808)	0,745 (0,701-0,790)	0,510 (0,440-0,581)	0,885 (0,849-0,920)	0,742 (0,704-0,780)	0,598 (0,529-0,668)	0,434 (0,306-0,562)
DVM	0,771 (0,724-0,819)	0,698 (0,620-0,775)	0,733 (0,688-0,778)	0,489 (0,418-0,559)	0,870 (0,833-0,907)	0,724 (0,685-0,763)	0,571 (0,501-0,641)	0,393 (0,262-0,523)
Bayesci GDM	0,805 (0,762-0,849)	0,731 (0,656-0,807)	0,745 (0,701-0,790)	0,511 (0,440-0,581)	0,884 (0,849-0,919)	0,741 (0,703-0,780)	0,598 (0,529-0,668)	0,434 (0,306-0,562)



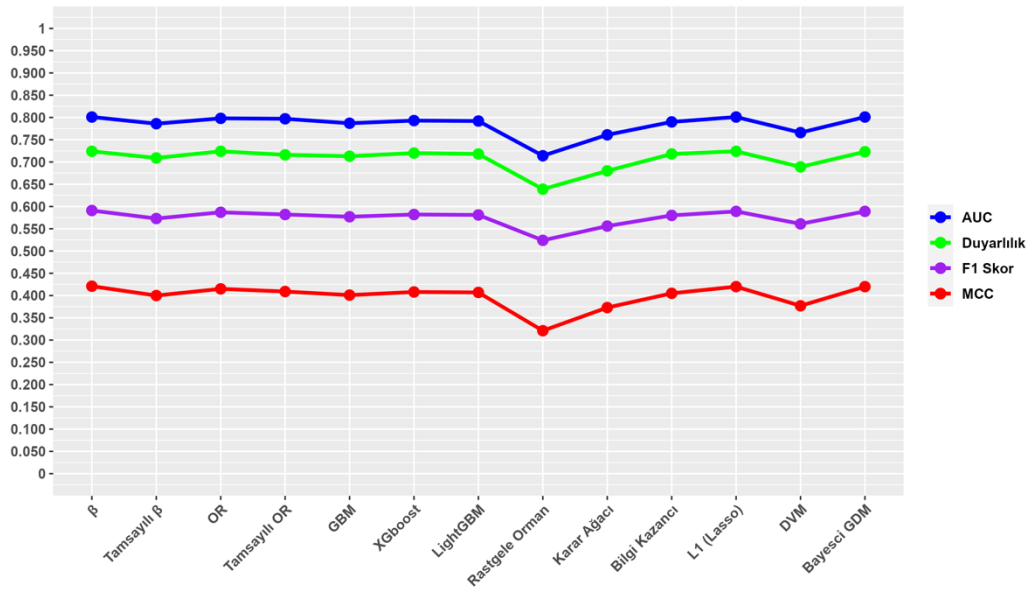
Şekil 3.26. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%20-%80 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, L1 (lasso), DVM ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.27’de verilmiştir.

Çizelge 3.27. Senaryo B; n=1000; bağımlı değişken 1-0 oranı: %20-%80 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,801 (0,770-0,833)	0,724 (0,670-0,778)	0,740 (0,708-0,771)	0,503 (0,453-0,552)	0,881 (0,855-0,906)	0,735 (0,708-0,763)	0,591 (0,542-0,640)	0,421 (0,330-0,512)
Tam sayılı Beta	0,786 (0,755-0,817)	0,709 (0,655-0,763)	0,726 (0,695-0,758)	0,493 (0,444-0,543)	0,87 (0,849-0,901)	0,721 (0,693-0,749)	0,573 (0,523-0,623)	0,400 (0,307-0,492)
OR	0,798 (0,767-0,830)	0,724 (0,670-0,778)	0,735 (0,703-0,766)	0,497 (0,447-0,546)	0,880 (0,854-0,906)	0,732 (0,704-0,759)	0,587 (0,538-0,636)	0,415 (0,325-0,506)
Tam sayılı OR	0,797 (0,765-0,828)	0,716 (0,662-0,770)	0,734 (0,702-0,766)	0,496 (0,446-0,546)	0,877 (0,851-0,903)	0,729 (0,702-0,756)	0,582 (0,533-0,631)	0,409 (0,318-0,501)
GBM	0,787 (0,754-0,819)	0,713 (0,659-0,768)	0,729 (0,697-0,761)	0,488 (0,439-0,538)	0,875 (0,849-0,901)	0,725 (0,698-0,753)	0,577 (0,528-0,626)	0,401 (0,310-0,492)
XGBoost	0,793 (0,761-0,825)	0,720 (0,666-0,774)	0,731 (0,699-0,763)	0,492 (0,443-0,542)	0,878 (0,852-0,904)	0,728 (0,701-0,756)	0,582 (0,533-0,631)	0,408 (0,317-0,499)
LightGBM	0,792 (0,760-0,824)	0,718 (0,664-0,772)	0,731 (0,699-0,763)	0,492 (0,443-0,542)	0,877 (0,851-0,903)	0,728 (0,700-0,755)	0,581 (0,533-0,630)	0,407 (0,316-0,498)
Rastgele Orman	0,714 (0,677-0,751)	0,639 (0,581-0,696)	0,711 (0,679-0,743)	0,452 (0,402-0,501)	0,844 (0,816-0,873)	0,691 (0,663-0,720)	0,524 (0,474-0,574)	0,321 (0,226-0,416)
Karar Ağacı	0,761 (0,727-0,795)	0,680 (0,625-0,736)	0,726 (0,694-0,757)	0,479 (0,429-0,530)	0,864 (0,837-0,891)	0,714 (0,686-0,742)	0,556 (0,506-0,606)	0,373 (0,279-0,467)
Bilgi Kazancı	0,790 (0,759-0,822)	0,718 (0,664-0,773)	0,729 (0,697-0,761)	0,490 (0,441-0,539)	0,877 (0,851-0,903)	0,726 (0,699-0,754)	0,580 (0,531-0,629)	0,405 (0,314-0,495)
L1 (Lasso)	0,801 (0,770-0,832)	0,724 (0,670-0,778)	0,739 (0,708-0,771)	0,501 (0,451-0,550)	0,881 (0,855-0,906)	0,735 (0,708-0,762)	0,589 (0,540-0,638)	0,420 (0,329-0,511)
DVM	0,766 (0,732-0,800)	0,689 (0,633-0,744)	0,726 (0,694-0,758)	0,477 (0,428-0,527)	0,866 (0,839-0,893)	0,716 (0,688-0,744)	0,561 (0,511-0,610)	0,377 (0,285-0,470)
Bayesci GDM	0,801 (0,770-0,832)	0,723 (0,669-0,777)	0,740 (0,708-0,771)	0,501 (0,452-0,551)	0,880 (0,855-0,906)	0,735 (0,708-0,762)	0,589 (0,5407-0,639)	0,420 (0,329-0,511)



Şekil 3.27. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%20-%80 ve rastgele)

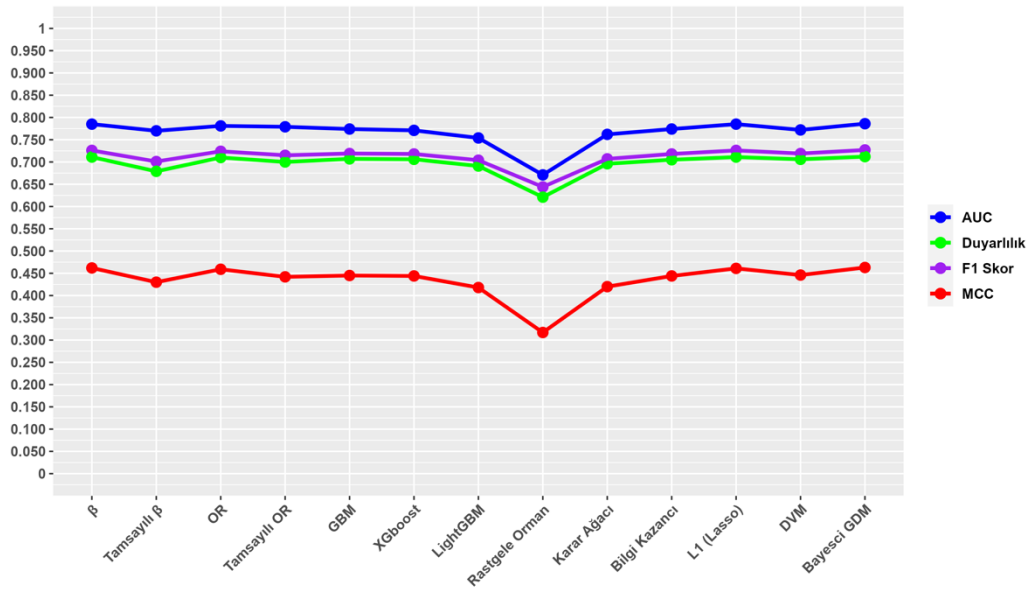
AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın, DVM ve karar ağacının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.10. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve %20-%80 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişkenin 1-0 oranının %20-%80 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetildi. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.28’de verilmiştir.

Çizelge 3.28. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,785 (0,723-0,849)	0,711 (0,623-0,799)	0,750 (0,665-0,834)	0,745 (0,659-0,831)	0,718 (0,632-0,804)	0,730 (0,669-0,792)	0,726 (0,637-0,815)	0,462 (0,285-0,639)
Tam sayılı Beta	0,770 (0,707-0,832)	0,679 (0,590-0,769)	0,745 (0,661-0,829)	0,738 (0,649-0,826)	0,699 (0,611-0,786)	0,712 (0,649-0,774)	0,701 (0,608-0,795)	0,430 (0,246-0,615)
OR	0,781 (0,718-0,845)	0,710 (0,622-0,798)	0,748 (0,663-0,833)	0,743 (0,657-0,830)	0,716 (0,630-0,803)	0,729 (0,667-0,790)	0,724 (0,635-0,814)	0,459 (0,281-0,636)
Tam sayılı OR	0,779 (0,716-0,843)	0,700 (0,611-0,789)	0,741 (0,655-0,826)	0,736 (0,649-0,824)	0,708 (0,621-0,795)	0,720 (0,658-0,782)	0,715 (0,625-0,806)	0,442 (0,263-0,622)
GBM	0,774 (0,709-0,838)	0,707 (0,619-0,795)	0,736 (0,650-0,823)	0,735 (0,648-0,822)	0,711 (0,623-0,798)	0,722 (0,660-0,784)	0,719 (0,630-0,808)	0,445 (0,267-0,622)
XGBoost	0,771 (0,706-0,836)	0,706 (0,617-0,794)	0,737 (0,651-0,823)	0,735 (0,648-0,822)	0,710 (0,622-0,797)	0,721 (0,659-0,783)	0,718 (0,629-0,808)	0,444 (0,266-0,622)
LightGBM	0,754 (0,688-0,821)	0,691 (0,601-0,780)	0,726 (0,639-0,813)	0,723 (0,634-0,811)	0,697 (0,608-0,785)	0,708 (0,645-0,771)	0,704 (0,613-0,795)	0,418 (0,237-0,599)
Rastgele Orman	0,671 (0,597-0,746)	0,621 (0,528-0,714)	0,693 (0,604-0,782)	0,677 (0,583-0,772)	0,642 (0,552-0,733)	0,657 (0,591-0,722)	0,644 (0,547-0,741)	0,317 (0,124-0,509)
Karar Ağacı	0,762 (0,698-0,827)	0,696 (0,607-0,785)	0,723 (0,635-0,810)	0,722 (0,633-0,810)	0,700 (0,611-0,789)	0,709 (0,647-0,772)	0,707 (0,616-0,797)	0,420 (0,240-0,600)
Bilgi Kazancı	0,774 (0,710-0,838)	0,705 (0,616-0,793)	0,739 (0,653-0,825)	0,735 (0,648-0,822)	0,710 (0,623-0,798)	0,721 (0,659-0,783)	0,718 (0,628-0,807)	0,444 (0,266-0,623)
L1 (Lasso)	0,785 (0,722-0,848)	0,711 (0,623-0,799)	0,749 (0,664-0,833)	0,744 (0,658-0,831)	0,717 (0,631-0,804)	0,730 (0,668-0,791)	0,726 (0,637-0,815)	0,461 (0,284-0,638)
SVM	0,772 (0,707-0,837)	0,706 (0,618-0,794)	0,739 (0,654-0,825)	0,736 (0,649-0,823)	0,711 (0,624-0,798)	0,722 (0,660-0,784)	0,719 (0,630-0,808)	0,446 (0,268-0,624)
Bayesci GLM	0,786 (0,723-0,849)	0,712 (0,624-0,800)	0,750 (0,665-0,834)	0,746 (0,659-0,832)	0,718 (0,631-0,805)	0,731 (0,669-0,792)	0,727 (0,638-0,815)	0,463 (0,286-0,639)



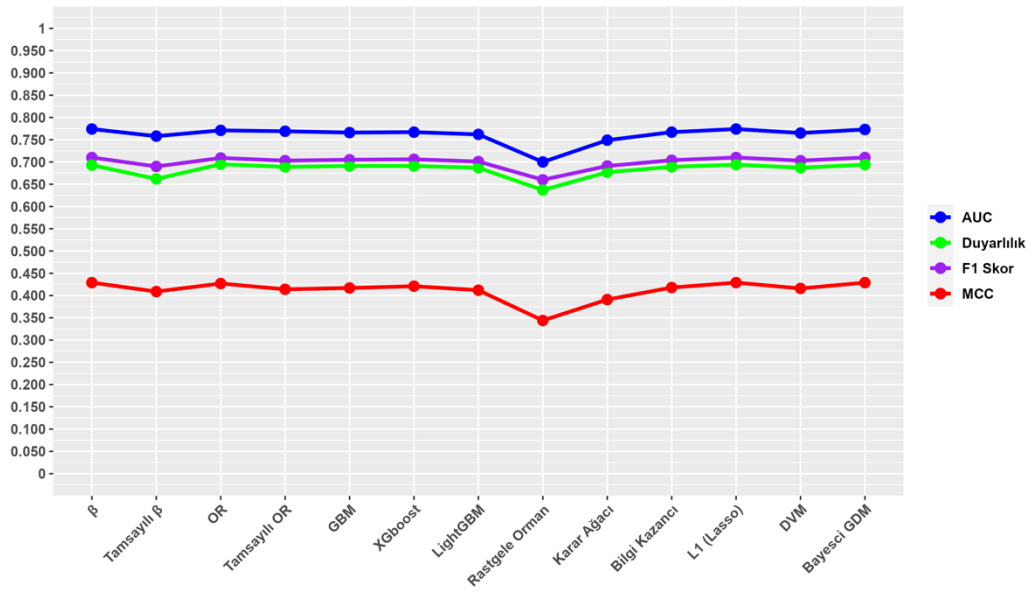
Şekil 3.28. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%50-%50 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazanci, DVM, karar ağacı, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: %20-%80 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.29’da verilmiştir.

Çizelge 3.29. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,774 (0,733-0,814)	0,693 (0,637-0,750)	0,735 (0,680-0,790)	0,729 (0,673-0,785)	0,701 (0,645-0,757)	0,714 (0,674-0,754)	0,710 (0,653-0,767)	0,429 (0,315-0,544)
Tam sayılı Beta	0,758 (0,717-0,798)	0,662 (0,605-0,720)	0,742 (0,689-0,796)	0,729 (0,671-0,786)	0,685 (0,63-0,741)	0,702 (0,662-0,742)	0,690 (0,630-0,750)	0,409 (0,291-0,528)
OR	0,771 (0,730-0,812)	0,695 (0,638-0,751)	0,731 (0,676-0,786)	0,727 (0,671-0,783)	0,701 (0,645-0,757)	0,713 (0,673-0,752)	0,709 (0,652-0,767)	0,427 (0,313-0,541)
Tam sayılı OR	0,769 (0,728-0,809)	0,689 (0,632-0,745)	0,724 (0,669-0,780)	0,721 (0,664-0,777)	0,695 (0,639-0,751)	0,706 (0,666-0,746)	0,703 (0,645-0,760)	0,414 (0,300-0,529)
GBM	0,766 (0,725-0,808)	0,691 (0,635-0,748)	0,725 (0,670-0,781)	0,721 (0,665-0,778)	0,697 (0,641-0,753)	0,708 (0,668-0,748)	0,705 (0,648-0,762)	0,417 (0,303-0,532)
XGBoost	0,767 (0,726-0,809)	0,691 (0,634-0,748)	0,730 (0,674-0,785)	0,724 (0,668-0,781)	0,698 (0,642-0,754)	0,710 (0,670-0,750)	0,706 (0,649-0,764)	0,421 (0,307-0,536)
LightGBM	0,762 (0,721-0,804)	0,687 (0,630-0,744)	0,724 (0,669-0,780)	0,719 (0,663-0,776)	0,694 (0,637-0,750)	0,705 (0,665-0,745)	0,701 (0,644-0,759)	0,412 (0,297-0,527)
Rastgele Orman	0,700 (0,654-0,746)	0,637 (0,578-0,696)	0,705 (0,648-0,761)	0,690 (0,631-0,749)	0,656 (0,599-0,713)	0,670 (0,629-0,712)	0,660 (0,599-0,721)	0,344 (0,223-0,464)
Karar Ağacı	0,749 (0,707-0,790)	0,677 (0,619-0,734)	0,713 (0,657-0,769)	0,708 (0,651-0,765)	0,684 (0,627-0,741)	0,695 (0,654-0,735)	0,691 (0,632-0,749)	0,391 (0,275-0,507)
Bilgi Kazancı	0,767 (0,726-0,808)	0,689 (0,632-0,746)	0,728 (0,673-0,783)	0,723 (0,666-0,779)	0,696 (0,640-0,752)	0,708 (0,668-0,748)	0,704 (0,647-0,762)	0,418 (0,303-0,533)
L1 (Lasso)	0,774 (0,733-0,814)	0,694 (0,637-0,750)	0,735 (0,680-0,790)	0,729 (0,673-0,785)	0,701 (0,645-0,757)	0,714 (0,674-0,754)	0,710 (0,653-0,767)	0,429 (0,315-0,543)
DVM	0,765 (0,723-0,806)	0,687 (0,630-0,744)	0,727 (0,672-0,783)	0,722 (0,665-0,778)	0,695 (0,639-0,751)	0,707 (0,667-0,747)	0,703 (0,645-0,761)	0,416 (0,301-0,530)
Bayesci GDM	0,773 (0,733-0,814)	0,694 (0,637-0,750)	0,734 (0,679-0,789)	0,729 (0,673-0,785)	0,701 (0,645-0,756)	0,714 (0,674-0,753)	0,710 (0,653-0,767)	0,429 (0,314-0,543)



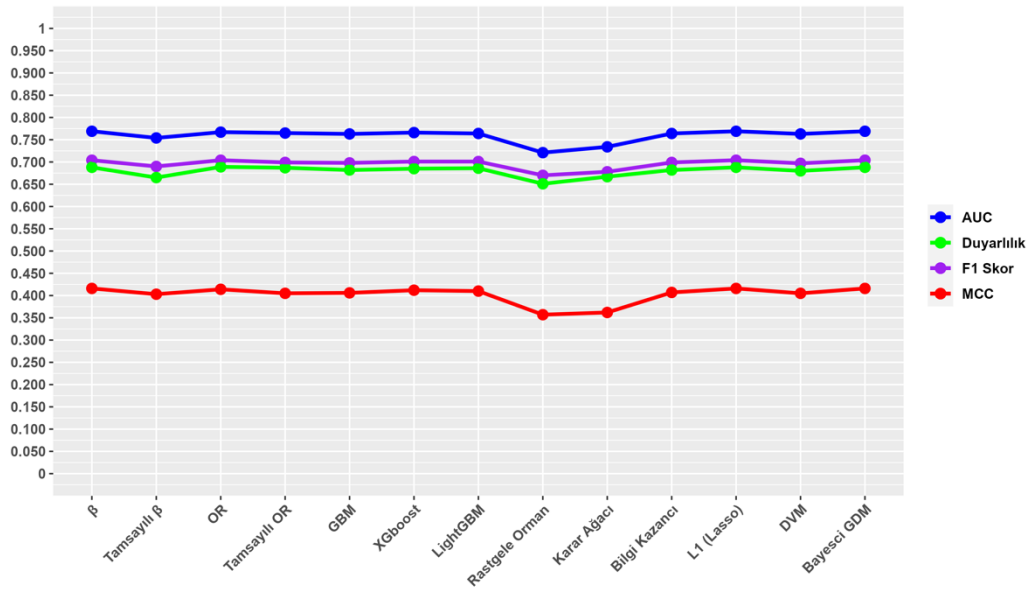
Şekil 3.29. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%50-%50 ve %20-%80)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, bilgi kazanci, DVM, karar ağacı, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: %20-%80 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.30'da verilmiştir.

Çizelge 3.30. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %20-%80 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,769 (0,741-0,798)	0,688 (0,648-0,729)	0,727 (0,688-0,767)	0,721 (0,681-0,761)	0,695 (0,656-0,735)	0,708 (0,679-0,736)	0,704 (0,663-0,745)	0,416 (0,335-0,497)
Tam sayılı Beta	0,754 (0,725-0,783)	0,665 (0,624-0,705)	0,736 (0,697-0,774)	0,723 (0,682-0,763)	0,684 (0,644-0,723)	0,700 (0,671-0,728)	0,690 (0,648-0,732)	0,403 (0,320-0,487)
OR	0,767 (0,738-0,796)	0,689 (0,649-0,729)	0,725 (0,686-0,764)	0,720 (0,680-0,760)	0,695 (0,655-0,735)	0,707 (0,679-0,735)	0,704 (0,663-0,744)	0,414 (0,333-0,495)
Tam sayılı OR	0,765 (0,736-0,794)	0,687 (0,640-0,728)	0,716 (0,677-0,756)	0,715 (0,674-0,755)	0,692 (0,652-0,732)	0,702 (0,673-0,730)	0,699 (0,659-0,740)	0,405 (0,324-0,486)
GBM	0,763 (0,734-0,792)	0,682 (0,641-0,722)	0,723 (0,684-0,763)	0,717 (0,677-0,757)	0,689 (0,650-0,729)	0,702 (0,674-0,731)	0,698 (0,657-0,739)	0,406 (0,324-0,487)
XGBoost	0,766 (0,736-0,795)	0,685 (0,645-0,726)	0,726 (0,687-0,765)	0,720 (0,680-0,760)	0,693 (0,653-0,732)	0,705 (0,677-0,733)	0,701 (0,661-0,742)	0,412 (0,330-0,493)
LightGBM	0,764 (0,735-0,794)	0,686 (0,645-0,726)	0,724 (0,684-0,763)	0,718 (0,678-0,758)	0,692 (0,653-0,732)	0,704 (0,676-0,733)	0,701 (0,660-0,742)	0,410 (0,329-0,491)
Rastgele Orman	0,721 (0,689-0,752)	0,651 (0,609-0,692)	0,705 (0,665-0,745)	0,694 (0,653-0,735)	0,664 (0,624-0,705)	0,677 (0,648-0,706)	0,670 (0,628-0,713)	0,357 (0,273-0,441)
Karar Ağacı	0,734 (0,705-0,764)	0,667 (0,626-0,708)	0,695 (0,654-0,735)	0,692 (0,651-0,733)	0,671 (0,630-0,711)	0,681 (0,652-0,709)	0,678 (0,637-0,720)	0,362 (0,279-0,445)
Bilgi Kazancı	0,764 (0,735-0,793)	0,682 (0,642-0,722)	0,724 (0,685-0,763)	0,718 (0,677-0,758)	0,690 (0,650-0,730)	0,703 (0,674-0,731)	0,699 (0,658-0,740)	0,407 (0,325-0,488)
L1 (Lasso)	0,769 (0,740-0,798)	0,688 (0,648-0,728)	0,728 (0,689-0,767)	0,722 (0,682-0,762)	0,695 (0,655-0,735)	0,708 (0,679-0,736)	0,704 (0,663-0,745)	0,416 (0,335-0,498)
DVM	0,763 (0,733-0,792)	0,680 (0,640-0,721)	0,724 (0,685-0,763)	0,717 (0,677-0,757)	0,689 (0,649-0,729)	0,702 (0,673-0,730)	0,697 (0,656-0,738)	0,405 (0,323-0,487)
Bayesci GDM	0,769 (0,740-0,798)	0,688 (0,648-0,728)	0,728 (0,688-0,767)	0,722 (0,682-0,762)	0,695 (0,655-0,735)	0,708 (0,679-0,736)	0,704 (0,663-0,745)	0,416 (0,335-0,497)



Şekil 3.30. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=1000; 1-0 oranı:%50-%50 ve %20-%80)

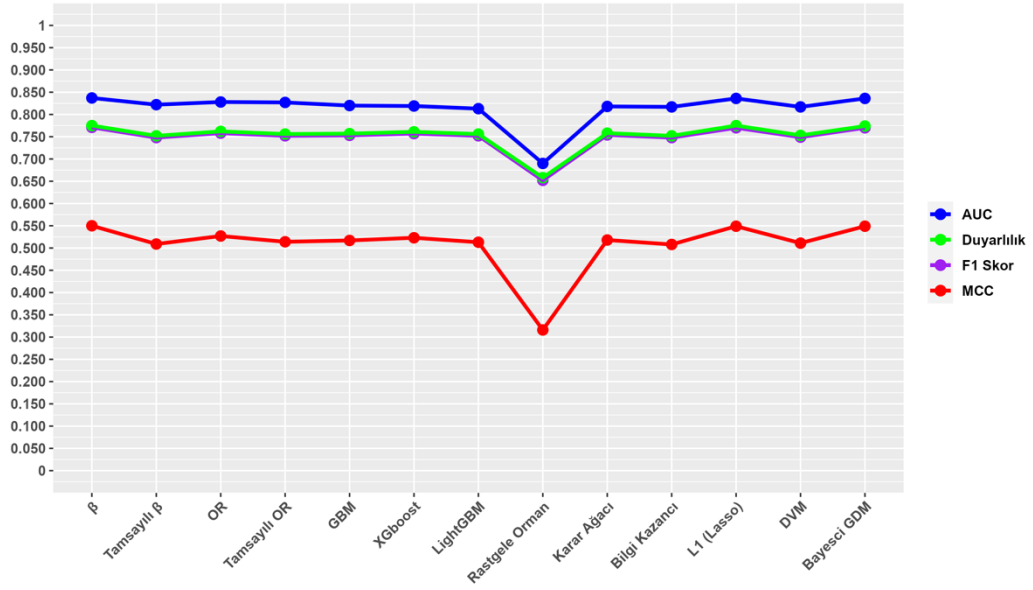
AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağaçlarının sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.1.11. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve %50-%50 (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranının %50-%50 ve bağımsız değişkenin 1-0 oranının %50-%50 olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile veriler türetildi. Türetilen bu veri setinde lojistik regresyon analizinde anlamlı bulunan β katsayılarına, odds oranlarına ve bu değerlerin tamsayıya dönüştürülmesi ile elde edilen ağırlıklı skorlara, ML algoritmalarına ve diğer algoritmalara dayalı oluşturulan klinik tahmin modellerinin performansları Çizelge 3.31’de verilmiştir.

Çizelge 3.31. Senaryo B; n=200; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,837 (0,782-0,892)	0,775 (0,693-0,857)	0,775 (0,694-0,855)	0,769 (0,687-0,852)	0,781 (0,701-0,861)	0,775 (0,717-0,832)	0,771 (0,688-0,853)	0,550 (0,385-0,714)
Tam sayılı Beta	0,822 (0,766-0,877)	0,752 (0,668-0,836)	0,754 (0,671-0,835)	0,751 (0,666-0,835)	0,763 (0,680-0,845)	0,753 (0,693-0,812)	0,748 (0,661-0,833)	0,509 (0,339-0,679)
OR	0,828 (0,772-0,885)	0,762 (0,679-0,846)	0,764 (0,682-0,845)	0,758 (0,674-0,842)	0,769 (0,688-0,852)	0,763 (0,704-0,822)	0,758 (0,674-0,843)	0,527 (0,359-0,694)
Tam sayılı OR	0,827 (0,772-0,885)	0,756 (0,672-0,840)	0,757 (0,675-0,840)	0,752 (0,667-0,836)	0,763 (0,681-0,846)	0,756 (0,697-0,816)	0,752 (0,667-0,837)	0,514 (0,345-0,683)
GBM	0,820 (0,762-0,878)	0,757 (0,673-0,841)	0,759 (0,676-0,841)	0,753 (0,669-0,838)	0,764 (0,682-0,847)	0,758 (0,699-0,817)	0,753 (0,668-0,838)	0,517 (0,348-0,686)
XGBoost	0,819 (0,760-0,877)	0,761 (0,678-0,845)	0,760 (0,678-0,843)	0,755 (0,671-0,839)	0,768 (0,687-0,850)	0,761 (0,702-0,820)	0,757 (0,672-0,841)	0,523 (0,355-0,690)
LightGBM	0,813 (0,754-0,872)	0,756 (0,672-0,841)	0,756 (0,673-0,838)	0,750 (0,665-0,835)	0,764 (0,681-0,846)	0,756 (0,697-0,815)	0,752 (0,667-0,837)	0,513 (0,344-0,682)
Rastgele Orman	0,690 (0,618-0,762)	0,658 (0,566-0,749)	0,658 (0,568-0,748)	0,651 (0,559-0,743)	0,666 (0,576-0,757)	0,658 (0,593-0,722)	0,652 (0,560-0,744)	0,316 (0,132-0,500)
Karar Ağacı	0,818 (0,759-0,876)	0,758 (0,674-0,842)	0,758 (0,676-0,841)	0,753 (0,669-0,838)	0,765 (0,683-0,848)	0,758 (0,699-0,818)	0,754 (0,669-0,839)	0,518 (0,349-0,686)
Bilgi Kazancı	0,817 (0,759-0,875)	0,752 (0,668-0,837)	0,754 (0,671-0,837)	0,749 (0,664-0,834)	0,761 (0,678-0,843)	0,753 (0,694-0,813)	0,748 (0,663-0,834)	0,508 (0,338-0,678)
L1 (Lasso)	0,836 (0,781-0,891)	0,775 (0,693-0,857)	0,774 (0,693-0,854)	0,768 (0,685-0,851)	0,781 (0,701-0,861)	0,774 (0,716-0,832)	0,770 (0,687-0,853)	0,549 (0,384-0,713)
DVM	0,817 (0,759-0,875)	0,753 (0,669-0,837)	0,756 (0,674-0,838)	0,751 (0,667-0,836)	0,762 (0,679-0,845)	0,754 (0,695-0,814)	0,749 (0,664-0,835)	0,511 (0,341-0,681)
Bayesci GDM	0,836 (0,781-0,891)	0,774 (0,692-0,856)	0,774 (0,694-0,855)	0,768 (0,686-0,851)	0,781 (0,701-0,861)	0,774 (0,716-0,832)	0,770 (0,687-0,853)	0,549 (0,384-0,713)



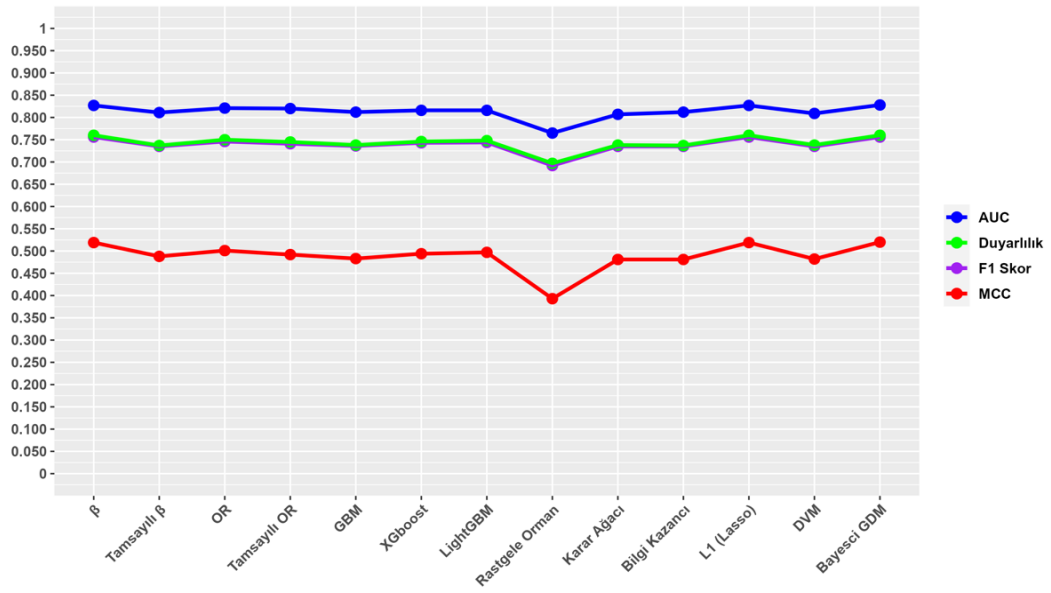
Şekil 3.31. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%50-%50 ve %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: %50-%50 olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.32’de verilmiştir.

Çizelge 3.32. Senaryo B; n=500; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,827 (0,791-0,863)	0,760 (0,706-0,813)	0,759 (0,707-0,811)	0,754 (0,700-0,807)	0,766 (0,714-0,818)	0,759 (0,722-0,797)	0,756 (0,702-0,809)	0,519 (0,413-0,626)
Tam sayılı Beta	0,811 (0,774-0,847)	0,737 (0,683-0,791)	0,745 (0,693-0,798)	0,742 (0,688-0,797)	0,751 (0,698-0,804)	0,741 (0,703-0,780)	0,735 (0,679-0,791)	0,488 (0,378-0,598)
OR	0,821 (0,785-0,857)	0,750 (0,696-0,803)	0,750 (0,697-0,803)	0,745 (0,69-0,799)	0,757 (0,704-0,810)	0,750 (0,712-0,788)	0,746 (0,692-0,800)	0,501 (0,393-0,609)
Tam sayılı OR	0,820 (0,784-0,856)	0,745 (0,691-0,799)	0,746 (0,692-0,799)	0,740 (0,685-0,794)	0,753 (0,699-0,806)	0,745 (0,707-0,784)	0,741 (0,687-0,796)	0,492 (0,383-0,600)
GBM	0,812 (0,774-0,849)	0,738 (0,684-0,793)	0,743 (0,69-0,796)	0,737 (0,682-0,792)	0,748 (0,694-0,801)	0,741 (0,702-0,779)	0,736 (0,681-0,791)	0,483 (0,373-0,592)
XGBoost	0,816 (0,779-0,853)	0,746 (0,692-0,800)	0,747 (0,694-0,800)	0,741 (0,687-0,796)	0,754 (0,701-0,807)	0,747 (0,709-0,785)	0,743 (0,688-0,797)	0,494 (0,386-0,603)
LightGBM	0,816 (0,779-0,853)	0,748 (0,694-0,802)	0,748 (0,695-0,801)	0,742 (0,688-0,797)	0,755 (0,702-0,808)	0,748 (0,710-0,786)	0,744 (0,690-0,798)	0,497 (0,388-0,605)
Rastgele Orman	0,765 (0,692-0,779)	0,697 (0,640-0,754)	0,696 (0,640-0,752)	0,690 (0,633-0,747)	0,704 (0,648-0,760)	0,696 (0,656-0,736)	0,692 (0,635-0,749)	0,393 (0,280-0,507)
Karar Ağacı	0,807 (0,770-0,845)	0,738 (0,683-0,792)	0,742 (0,688-0,795)	0,736 (0,681-0,791)	0,747 (0,694-0,801)	0,740 (0,701-0,778)	0,735 (0,68-0,79)	0,481 (0,372-0,591)
Bilgi Kazancı	0,812 (0,775-0,849)	0,737 (0,682-0,792)	0,743 (0,689-0,796)	0,736 (0,681-0,791)	0,747 (0,693-0,800)	0,740 (0,701-0,778)	0,735 (0,680-0,790)	0,481 (0,372-0,591)
L1 (Lasso)	0,827 (0,792-0,863)	0,760 (0,706-0,813)	0,759 (0,707-0,812)	0,753 (0,700-0,807)	0,766 (0,714-0,818)	0,759 (0,722-0,797)	0,756 (0,702-0,809)	0,519 (0,413-0,626)
DVM	0,809 (0,772-0,847)	0,738 (0,683-0,792)	0,742 (0,688-0,795)	0,736 (0,681-0,791)	0,747 (0,694-0,801)	0,740 (0,702-0,778)	0,735 (0,680-0,790)	0,482 (0,372-0,591)
Bayesci GDM	0,828 (0,792-0,863)	0,760 (0,706-0,813)	0,760 (0,708-0,812)	0,754 (0,700-0,808)	0,766 (0,714-0,818)	0,760 (0,722-0,797)	0,756 (0,703-0,81)	0,520 (0,414-0,626)



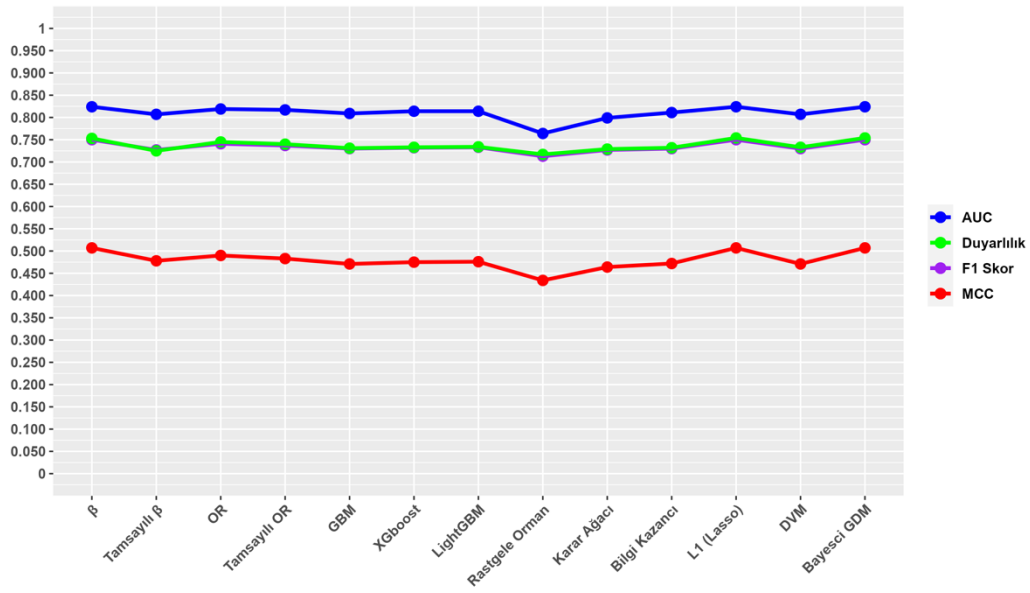
Şekil 3.32. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%50-%50 ve %50-%50)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: %50-%50 olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan 10000 benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.33’de verilmiştir.

Çizelge 3.33. Senaryo B; n=1000; bağımlı ve bağımsız değişken 1-0 oranı: %50-%50 ve %50-%50 için farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,824 (0,798-0,849)	0,753 (0,715-0,791)	0,753 (0,716-0,791)	0,747 (0,709-0,785)	0,760 (0,723-0,797)	0,753 (0,727-0,780)	0,750 (0,712-0,788)	0,507 (0,431-0,583)
Tam sayılı Beta	0,807 (0,781-0,833)	0,725 (0,687-0,764)	0,746 (0,709-0,783)	0,741 (0,702-0,780)	0,745 (0,707-0,782)	0,736 (0,708-0,763)	0,727 (0,687-0,767)	0,478 (0,399-0,557)
OR	0,819 (0,793-0,844)	0,745 (0,706-0,783)	0,745 (0,707-0,783)	0,739 (0,700-0,778)	0,752 (0,714-0,789)	0,745 (0,718-0,772)	0,741 (0,703-0,780)	0,490 (0,413-0,567)
Tam sayılı OR	0,817 (0,792-0,843)	0,740 (0,702-0,779)	0,742 (0,704-0,780)	0,736 (0,697-0,774)	0,748 (0,71-0,786)	0,741 (0,714-0,768)	0,737 (0,698-0,776)	0,483 (0,406-0,560)
GBM	0,809 (0,782-0,835)	0,731 (0,692-0,77)	0,738 (0,700-0,776)	0,731 (0,692-0,77)	0,741 (0,703-0,779)	0,735 (0,707-0,762)	0,730 (0,69-0,769)	0,471 (0,393-0,549)
XGBoost	0,814 (0,787-0,840)	0,733 (0,694-0,772)	0,741 (0,703-0,778)	0,733 (0,694-0,772)	0,743 (0,705-0,781)	0,737 (0,710-0,764)	0,732 (0,693-0,771)	0,475 (0,397-0,553)
LightGBM	0,814 (0,788-0,840)	0,734 (0,696-0,773)	0,740 (0,702-0,778)	0,733 (0,694-0,772)	0,744 (0,706-0,782)	0,737 (0,710-0,764)	0,733 (0,693-0,772)	0,476 (0,398-0,553)
Rastgele Orman	0,764 (0,735-0,794)	0,717 (0,677-0,757)	0,716 (0,677-0,755)	0,710 (0,670-0,750)	0,724 (0,685-0,763)	0,717 (0,689-0,744)	0,713 (0,673-0,753)	0,434 (0,355-0,513)
Karar Ağacı	0,799 (0,772-0,826)	0,729 (0,69-0,768)	0,734 (0,696-0,772)	0,727 (0,688-0,767)	0,738 (0,7-0,777)	0,732 (0,704-0,759)	0,727 (0,688-0,766)	0,464 (0,386-0,543)
Bilgi Kazancı	0,811 (0,785-0,837)	0,732 (0,693-0,770)	0,739 (0,701-0,777)	0,732 (0,693-0,771)	0,742 (0,704-0,780)	0,735 (0,708-0,763)	0,730 (0,691-0,770)	0,472 (0,394-0,550)
L1 (Lasso)	0,824 (0,798-0,849)	0,754 (0,716-0,792)	0,753 (0,716-0,790)	0,747 (0,709-0,785)	0,760 (0,723-0,798)	0,753 (0,727-0,780)	0,750 (0,712-0,788)	0,507 (0,431-0,583)
DVM	0,807 (0,780-0,833)	0,733 (0,694-0,772)	0,737 (0,699-0,775)	0,730 (0,691-0,770)	0,742 (0,704-0,780)	0,735 (0,707-0,762)	0,730 (0,691-0,769)	0,471 (0,393-0,549)
Bayesci GDM	0,824 (0,799-0,849)	0,754 (0,716-0,792)	0,753 (0,716-0,791)	0,747 (0,709-0,786)	0,760 (0,723-0,798)	0,754 (0,727-0,78)	0,750 (0,712-0,788)	0,507 (0,432-0,583)



Şekil 3.33. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=1000; 1-0 oranı:%50-%50 ve %50-%50)

AUC ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

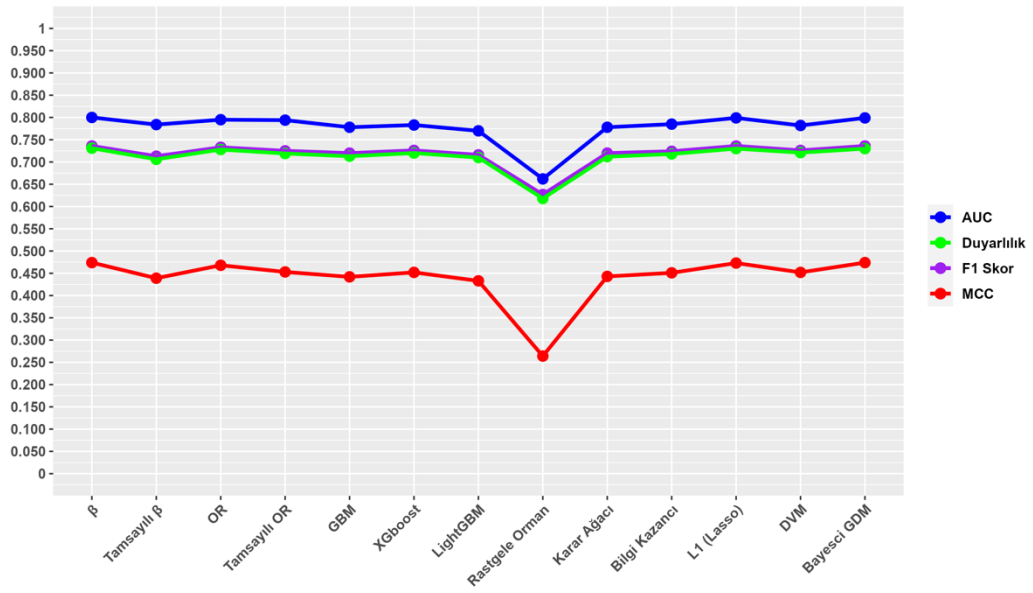
Duyarlılık ve F1 skoruna performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM ve Rastgele ormanın modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir.

3.1.12. Senaryo B, bağımlı ve bağımsız değişken oranı %50-%50 ve rastgele (n=200, n=500 ve n=1000)

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=200 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.34’de verilmiştir.

Çizelge 3.34. Senaryo B; n=200; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,800 (0,739-0,860)	0,731 (0,645-0,817)	0,744 (0,658-0,830)	0,746 (0,661-0,831)	0,728 (0,642-0,815)	0,737 (0,676-0,798)	0,736 (0,650-0,823)	0,474 (0,301-0,647)
Tam sayılı Beta	0,784 (0,723-0,844)	0,706 (0,620-0,793)	0,728 (0,642-0,814)	0,734 (0,647-0,821)	0,710 (0,621-0,798)	0,717 (0,655-0,779)	0,713 (0,624-0,803)	0,439 (0,260-0,618)
OR	0,795 (0,734-0,856)	0,728 (0,642-0,814)	0,740 (0,654-0,827)	0,742 (0,657-0,828)	0,725 (0,638-0,812)	0,734 (0,673-0,795)	0,733 (0,646-0,820)	0,468 (0,294-0,641)
Tam sayılı OR	0,794 (0,733-0,855)	0,719 (0,632-0,806)	0,733 (0,647-0,820)	0,736 (0,650-0,822)	0,717 (0,629-0,805)	0,726 (0,665-0,788)	0,725 (0,637-0,813)	0,453 (0,277-0,628)
GBM	0,778 (0,715-0,842)	0,713 (0,626-0,801)	0,729 (0,641-0,816)	0,731 (0,644-0,817)	0,710 (0,622-0,799)	0,721 (0,659-0,783)	0,720 (0,631-0,808)	0,442 (0,265-0,618)
XGBoost	0,783 (0,720-0,846)	0,720 (0,633-0,806)	0,733 (0,646-0,820)	0,736 (0,650-0,822)	0,715 (0,628-0,803)	0,726 (0,665-0,788)	0,726 (0,638-0,813)	0,452 (0,277-0,627)
LightGBM	0,770 (0,706-0,835)	0,710 (0,622-0,797)	0,723 (0,635-0,811)	0,726 (0,639-0,813)	0,706 (0,618-0,795)	0,716 (0,654-0,778)	0,716 (0,627-0,804)	0,433 (0,255-0,610)
Rastgele Orman	0,662 (0,588-0,737)	0,618 (0,525-0,711)	0,645 (0,553-0,738)	0,645 (0,552-0,738)	0,620 (0,527-0,713)	0,631 (0,565-0,697)	0,627 (0,532-0,722)	0,264 (0,074-0,454)
Karar Ağacı	0,778 (0,714-0,842)	0,712 (0,625-0,800)	0,730 (0,643-0,818)	0,733 (0,647-0,819)	0,710 (0,622-0,798)	0,721 (0,659-0,783)	0,720 (0,631-0,808)	0,443 (0,266-0,619)
Bilgi Kazancı	0,785 (0,722-0,848)	0,718 (0,631-0,805)	0,733 (0,646-0,820)	0,735 (0,649-0,821)	0,715 (0,628-0,803)	0,725 (0,664-0,787)	0,724 (0,636-0,812)	0,451 (0,275-0,626)
L1 (Lasso)	0,799 (0,730-0,860)	0,730 (0,644-0,815)	0,744 (0,658-0,830)	0,746 (0,661-0,831)	0,727 (0,640-0,813)	0,737 (0,676-0,797)	0,736 (0,649-0,822)	0,473 (0,300-0,646)
DVM	0,782 (0,719-0,846)	0,721 (0,634-0,807)	0,732 (0,645-0,819)	0,736 (0,650-0,822)	0,717 (0,629-0,804)	0,726 (0,665-0,788)	0,726 (0,638-0,813)	0,452 (0,277-0,628)
Bayesci GDM	0,799 (0,739-0,860)	0,730 (0,644-0,816)	0,744 (0,659-0,830)	0,746 (0,661-0,831)	0,727 (0,640-0,813)	0,737 (0,676-0,798)	0,736 (0,649-0,822)	0,474 (0,3012-0,647)



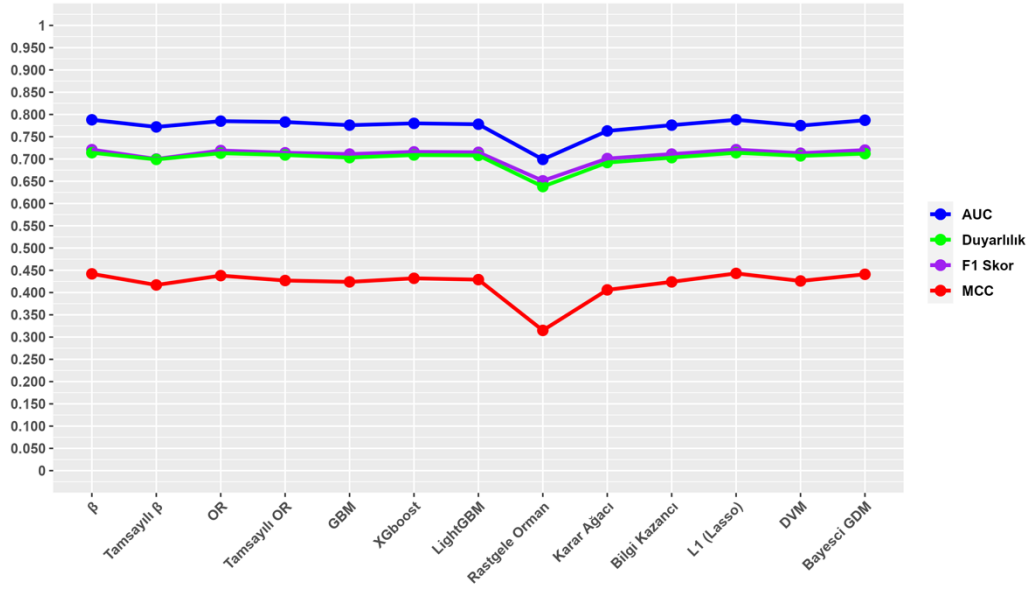
Şekil 3.34. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=200; 1-0 oranı:%50-%50 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=500 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.35’de verilmiştir.

Çizelge 3.35. Senaryo B; n=500; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,788 (0,749-0,827)	0,714 (0,658-0,769)	0,729 (0,673-0,784)	0,731 (0,676-0,786)	0,710 (0,654-0,766)	0,721 (0,682-0,760)	0,721 (0,665-0,777)	0,442 (0,330-0,553)
Tam sayılı Beta	0,772 (0,733-0,812)	0,699 (0,644-0,754)	0,714 (0,659-0,770)	0,721 (0,666-0,777)	0,700 (0,644-0,757)	0,707 (0,667-0,747)	0,700 (0,648-0,762)	0,417 (0,303-0,531)
OR	0,785 (0,746-0,825)	0,713 (0,657-0,769)	0,726 (0,670-0,782)	0,727 (0,672-0,782)	0,711 (0,655-0,766)	0,719 (0,680-0,759)	0,719 (0,663-0,775)	0,438 (0,327-0,550)
Tam sayılı OR	0,783 (0,744-0,822)	0,709 (0,653-0,764)	0,718 (0,662-0,774)	0,723 (0,667-0,778)	0,705 (0,648-0,761)	0,714 (0,674-0,753)	0,714 (0,658-0,770)	0,427 (0,315-0,539)
GBM	0,776 (0,735-0,816)	0,703 (0,665-0,777)	0,721 (0,667-0,778)	0,722 (0,667-0,778)	0,702 (0,645-0,758)	0,712 (0,673-0,752)	0,711 (0,655-0,768)	0,424 (0,311-0,537)
XGBoost	0,780 (0,740-0,820)	0,709 (0,653-0,764)	0,724 (0,668-0,780)	0,726 (0,671-0,781)	0,705 (0,649-0,761)	0,716 (0,677-0,756)	0,716 (0,660-0,772))	0,432 (0,320-0,544)
LightGBM	0,778- (0,738-0,818)	0,708 (0,652-0,764)	0,721 (0,665-0,777)	0,724 (0,669-0,779)	0,705 (0,648-0,761)	0,715 (0,675-0,754)	0,715 (0,658-0,771)	0,429 (0,317-0,541)
Rastgele Orman	0,699 (0,653-0,744)	0,638 (0,545-0,708)	0,677 (0,644-0,737)	0,672 (0,365-0,502)	0,644 (0,794-0,877)	0,657 (0,632-0,714)	0,651 (0,436-0,575)	0,315 (0,158-0,425)
Karar Ağacı	0,763 (0,722-0,805)	0,692 (0,636-0,749)	0,714 (0,658-0,770)	0,715 (0,659-0,771)	0,691 (0,635-0,748)	0,703 (0,663-0,743)	0,701 (0,644-0,759)	0,406 (0,292-0,521)
Bilgi Kazancı	0,776 (0,736-0,816)	0,703 (0,647-0,759)	0,721 (0,666-0,777)	0,723 (0,667-0,778)	0,701 (0,645-0,757)	0,712 (0,673-0,752)	0,711 (0,655-0,768)	0,424 (0,311-0,537)
L1 (Lasso)	0,788 (0,749-0,828)	0,714 (0,658-0,769)	0,729 (0,674-0,785)	0,731 (0,676-0,786)	0,711 (0,655-0,767)	0,722 (0,682-0,761)	0,721 (0,665-0,777)	0,443 (0,331-0,554)
DVM	0,775- (0,735-0,816)	0,707 (0,651-0,763)	0,719 (0,663-0,775)	0,722 (0,666-0,777)	0,703 (0,647-0,760)	0,713 (0,673-0,753)	0,713 (0,657-0,769)	0,426 (0,313-0,538)
Bayesci GLM	0,787 (0,748-0,826)	0,712 (0,657-0,768)	0,729 (0,673-0,784)	0,730 (0,675-0,785)	0,710 (0,654-0,766)	0,720 (0,681-0,760)	0,720 (0,664-0,775)	0,441 (0,329-0,553)



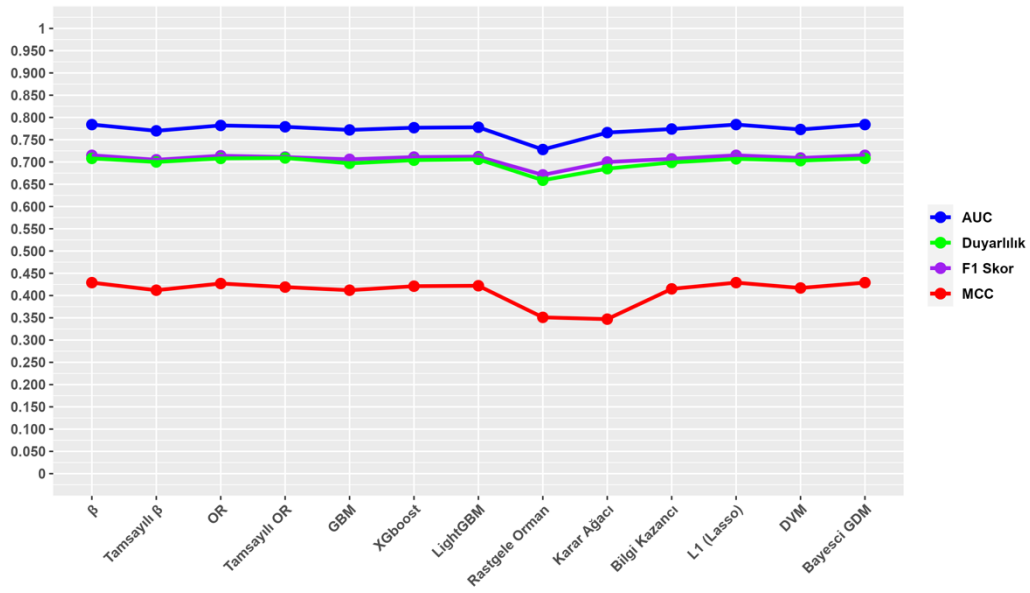
Şekil 3.35. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=500; 1-0 oranı:%50-%50 ve rastgele)

AUC, Duyarlilik, F1 ve MCC performans ölçütlerine göre beta, tamsayili beta, OR, tamsayili OR, GBM, XGboost, lightGBM, karar ağacı, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

Nominal bağımsız değişkenlerin olduğu B senaryosundaki bağımlı değişkenin 1-0 oranı %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele olduğu durumda n=1000 örneklem büyüklüğünde hesaplanan benzetim ile türetilen verilere ilişkin klinik tahmin modellerinin performansları Çizelge 3.36'de verilmiştir.

Çizelge 3.36. Senaryo B; n=1000; bağımlı değişken 1-0 oranı: %50-%50 ve bağımsız değişkenlerin 1-0 oranı: rastgele seçilen farklı algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Performans Ölçüsü (%95 GA)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
Beta	0,784 (0,756-0,812)	0,708 (0,668-0,747)	0,723 (0,683-0,762)	0,725 (0,685-0,764)	0,704 (0,664-0,744)	0,715 (0,687-0,743)	0,715 (0,675-0,755)	0,429 (0,350-0,509)
Tam sayılı Beta	0,770 (0,742-0,797)	0,700 (0,661-0,739)	0,708 (0,669-0,748)	0,718 (0,678-0,757)	0,698 (0,658-0,739)	0,705 (0,677-0,734)	0,705 (0,664-0,745)	0,412 (0,331-0,493)
OR	0,782 (0,754-0,810)	0,708 (0,668-0,747)	0,720 (0,680-0,760)	0,722 (0,683-0,762)	0,704 (0,664-0,744)	0,714 (0,686-0,742)	0,714 (0,674-0,754)	0,427 (0,348-0,506)
Tam sayılı OR	0,779 (0,751-0,807)	0,709 (0,670-0,748)	0,709 (0,669-0,749)	0,717 (0,677-0,756)	0,703 (0,662-0,743)	0,710 (0,682-0,738)	0,711 (0,672-0,751)	0,419 (0,339-0,498)
GBM	0,772 (0,743-0,800)	0,697 (0,657-0,737)	0,715 (0,676-0,755)	0,717 (0,678-0,757)	0,694 (0,654-0,734)	0,706 (0,678-0,734)	0,706 (0,666-0,746)	0,412 (0,332-0,492)
XGBoost	0,777 (0,749-0,806)	0,704 (0,664-0,744)	0,718 (0,678-0,757)	0,720 (0,680-0,759)	0,700 (0,660-0,740)	0,711 (0,683-0,739)	0,711 (0,671-0,751)	0,421 (0,341-0,501)
LightGBM	0,778 (0,749-0,806)	0,706 (0,666-0,745)	0,717 (0,677-0,757)	0,720 (0,681-0,751)	0,701 (0,661-0,741)	0,711 (0,683-0,739)	0,712 (0,672-0,752)	0,422 (0,342-0,501)
Rastgele Orman	0,728 (0,697-0,759)	0,659 (0,618-0,700)	0,692 (0,651-0,733)	0,688 (0,648-0,729)	0,662 (0,621-0,703)	0,675 (0,646-0,704)	0,671 (0,630-0,713)	0,351 (0,267-0,434)
Karar Ağacı	0,766 (0,726-0,807)	0,685 (0,616-0,754)	0,745 (0,715-0,774)	0,361 (0,309-0,414)	0,920 (0,899-0,940)	0,735 (0,707-0,762)	0,700 (0,661-0,740)	0,347 (0,245-0,449)
Bilgi Kazancı	0,774 (0,745-0,802)	0,699 (0,659-0,739)	0,716 (0,676-0,756)	0,717 (0,678-0,757)	0,697 (0,657-0,737)	0,708 (0,679-0,736)	0,707 (0,667-0,747)	0,415 (0,335-0,495)
L1 (Lasso)	0,784 (0,756-0,812)	0,707 (0,668-0,747)	0,723 (0,683-0,762)	0,724 (0,684-0,763)	0,705 (0,665-0,745)	0,715 (0,687-0,743)	0,715 (0,675-0,754)	0,429 (0,350-0,509)
DVM	0,773 (0,745-0,802)	0,703 (0,663-0,743)	0,714 (0,674-0,754)	0,717 (0,678-0,756)	0,699 (0,659-0,739)	0,709 (0,681-0,737)	0,709 (0,669-0,749)	0,417 (0,337-0,497)
Bayesci GDM	0,784 (0,756-0,812)	0,708 (0,668-0,747)	0,722 (0,682-0,762)	0,724 (0,685-0,763)	0,704 (0,665-0,744)	0,715 (0,687-0,743)	0,715 (0,675-0,754)	0,429 (0,350-0,509)



Şekil 3.36. LR, ML algoritmaları ve diğer algoritmalar ile oluşturulan modellerin performanslarını karşılaştıran çizgi grafiği (Senaryo B; n=1000; 1-0 oranı:%50-%50 ve rastgele)

AUC, Duyarlılık, F1 ve MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, karar ağacı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın sınıflama başarısının ise düşük olduğu belirlenmiştir.

MCC performans ölçütlerine göre beta, tamsayılı beta, OR, tamsayılı OR, GBM, XGboost, lightGBM, bilgi kazancı, DVM, L1 (lasso) ve bayesci GDM modellerinin sınıflama başarılarının yaklaşık olarak aynı düzeyde olduğu belirlenmiştir. Rastgele ormanın ve karar ağacı sınıflama başarısının ise düşük olduğu belirlenmiştir.

3.2. Gerçek veri setine ait sonuçlar

Araştırmada yer alan toplam 522 çocuğun 147 (%28,2)'si malnütrisyon tanısı almıştır. Tanı, Kilo, VKİ ve Sol üst orta kol çevresi değişkenlerinin z skorlarına göre konulmuştur. Bu değişkenlerin malnütrisyon tanısı alan ve almayanlarda değerleri incelendiğinde malnütrisyon tanısı alanların ortalama kilosu $17,8 \pm 4,9$ kg, VKİ'si $14,1 \pm 0,9$ ve sol üst orta kol çevresi $16,5 \pm 1,4$ cm iken aynı değerler tanı almayanlarda sırasıyla $23,1 \pm 8,7$ kg, $16,9 \pm 2,4$ ve $19,4 \pm 3,0$ cm'dir. Malnütrisyon tanısı alan çocuklarda boy ortalaması $111,2 \pm 15,1$ cm iken almayanlarda ise $114,8 \pm 15,9$ cm'dir.

Çocuğa ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması Çizelge 3.37’de sunulmuştur. Gruplar arasında cinsiyet, yaş, doğum haftası ve doğum şekli bakımından fark yokken (tüm değişkenler için $P>0,05$), doğum ağırlığı <2500 gr olanlarda malnütrisyon görülme sıklığı daha fazladır ($P<0,001$). Anne ve babanın düşüncesine göre çocuklarının yaşitlarına göre kilo ve boy durumları ile çocuklarında malnütrisyon görülmesi arasında anlamlı ilişki bulunmuştur (sırasıyla; $p<0,001$, $p<0,001$).



Çizelge 3.37. Çocuğa ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması

Değişkenler	Malnütrisyon		P değeri
	Yok (n=375)	Var (n=147)	
Cinsiyet, n (%)			
Kız	187 (71,1%)	76 (28,9%)	0,706 ^a
Erkek	188 (72,6%)	71 (27,4%)	
Yaş , ay, Ort±SS	69,9±28,2	74,8±29,4	0,079 ^c
Yaş,yıl, Ort±SS	5,4±2,3	5,8±2,4	0,084 ^c
2-5	198 (73,6%)	71 (26,4%)	0.355 ^a
6-10	177 (70%)	76 (30%)	
Yaşıtlarına göre kilosu, n (%)			
Aynı veya daha kilolu	263 (93,9%)	17 (6,1%)	<0,001 ^a
Yaşıtlarından daha zayıf	112 (46,3%)	130 (53,7%)	
Yaşıtlarına göre boyu, n (%)			
Aynı veya daha uzun	309 (81,5%)	70 (18,5%)	<0,001 ^a
Yaşıtlarından daha kısa	66 (46,2%)	77 (53,8%)	
Doğum ağırlığı, n (%)			
≥2500 g	343 (75,4%)	112 (24,6%)	<0,001 ^a
<2500 g	32 (47,8%)	35 (52,2%)	
Doğum haftası, n (%)			
37-42 hafta (Normal)	305 (72,3%)	117 (27,7%)	0,649 ^a
≤36 hafta (Prematüre)	70 (70%)	30 (30%)	
Doğum şekli, n (%)			
Normal	131 (72,8%)	49 (27,2%)	0,729 ^a
Sezeryan	244 (71,3%)	98 (28,7%)	

Ort, Ortalama, SS, Standart Sapma, VKİ: Vücut Kitle İndeksi

^aKi-kare testi

^bFisher kesin testi

^cStudent's t-testi

^dMann Whitney U testi

Anne ve babaya ilişkin sosyo-demografik özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması Çizelge 3.38'de sunulmuştur. Malnütrisyon grupları arasında çocukların annelerine ilişkin boy ve kilo değerleri istatistiksel olarak anlamlı farklı bulunmuştur (sırasıyla; $P<0,001$, $P=0,001$). VKİ değerleri ise istatistiksel olarak anlamlı farklı bulunmamıştır ($P=0,106$). Malnütrisyon grupları arasında çocukların babalarına ilişkin boy, kilo ve VKİ değerleri istatistiksel olarak anlamlı farklı bulunmuştur (sırasıyla; $P<0,001$, $P<0,001$, $P=0,001$). Malnütrisyon grupları arasında çocukların anne ve babalarının yaşları sınırda anlamlı bulunmuştur (sırasıyla; $P=0,05$, $P=0,05$)

Anne ve babanın eğitim durumları ile çocuklarında malnütrisyon görülmesi arasında anlamlı ilişki bulunamamıştır (sırasıyla; $P=0,204$, $P=0,738$). Anne ve babanın bir çalışıp çalışmama durumları ile çocuklarında malnütrisyon görülmesi arasında anlamlı ilişki bulunamamıştır (sırasıyla; $P=0,210$, $P=0,244$). Aile yapısı ile çocuklarında malnütrisyon görülmesi arasında anlamlı ilişki bulunamamıştır ($P=0,076$). Ailede yaşayan toplam kişi sayısı ve toplam çocuk sayısı ile malnütrisyon görülmesi arasında anlamlı ilişki bulunmuştur (sırasıyla; $P=0,011$, $P=0,024$) (Çizelge 3.38).



Çizelge 3.38. Anne ve babaya ilişkin sosyo-demografik özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması

Değişkenler	Malnutrition		P değeri
	Yok (n=375)	Var (n=147)	
Anne boy, cm, Ort±SS	161,1±6,1	158,6±5,6	<0,001 ^c
≥ 159,5	127 (%60,5)	83 (%39,5)	<0,001 ^a
< 159,5	248 (%79,5)	64 (%20,5)	
Anne kilo, kg, Ort±SS	68,5±12,8	64,4±12,1	0,001 ^c
≥ 67,5	185 (%65,6)	97 (34,4)	0,001 ^a
< 67,5	190 (%79,2)	50 (%20,8)	
Anne VKİ, Ort±SS	26,4±4,9	25,7±4,9	0,106 ^c
≥ 25 (kilolu veya obez)	207 (%74,2)	72 (%25,8)	0,200 ^a
≤ 24,9 (normal veya zayıf)	168 (%69,1)	75 (%30,9)	
Baba boy, cm, Ort±SS	173,7±7,3	171±7,1	<0,001 ^c
≥ 175,5	233 (%67)	115 (%33)	<0,001 ^a
< 175,5	142 (%81,6)	32 (%18,4)	
Baba kilo, kg, Ort±SS	82,5±12,9	76,1±11,7	<0,001 ^c
≥ 81,5	185 (%63,8)	105 (%36,2)	<0,001 ^a
< 81,5	190 (%81,9)	42 (%18,1)	
Baba VKİ, Ort±SS	27,3±4,0	26,0±3,7	0,001 ^c
≥ 25 (kilolu veya obez)	263 (%75,1)	87 (%24,9)	0,017 ^a
≤ 24,9 (normal veya zayıf)	112 (%65,1)	60 (%34,9)	
Annenin eğitim durumu, n(%)			
Lisans ve üzeri	58 (%77,3)	17 (%22,7)	0,204 ^a
Lise / Önlisans	154 (%74)	54 (%26)	
İlköğretim ve altı	163 (%68,2)	76 (%31,8)	
Babanın eğitim durumu, n(%)			
Lisans ve üzeri	64 (%75,3)	21 (%24,7)	0,738 ^a
Lise / Önlisans	159 (%71)	65 (%29)	
İlköğretim ve altı	152 (%71,4)	61 (%28,6)	
Anne çalışıyor mu?, n(%)			
Çalışıyor	71 (%77,2)	21 (%22,8)	0,210 ^a
Çalışmıyor	304 (%70,7)	126 (%29,3)	
Baba çalışıyor mu?, n(%)			
Çalışıyor	350 (%72,5)	133 (%27,5)	0,264 ^a
Çalışmıyor	25 (%64,1)	14 (%35,9)	
Aile yapısı, n(%)			
Çekirdek aile	310 (%73,3)	113 (%26,7)	0,076 ^a
Parçalanmış aile	19 (%79,2)	5 (%20,8)	
Geniş aile	46 (%61,3)	29 (%38,7)	
Ailede yaşayan toplam kişi sayısı, n(%)			
5 ve daha az kişi	331 (%73,9)	117 (%26,1)	0,011 ^a
6 ve daha fazla kişi	44 (%59,5)	30 (%40,5)	
Toplam çocuk sayısı, n(%)			
1 çocuk	69 (%81,2)	16 (%18,8)	0,024 ^a
2-3 çocuk	286 (%71,1)	116 (%28,9)	
4 ve üzeri çocuk	20 (%57,1)	15 (%42,9)	

SS, Standart Sapma, VKİ: Vücut Kitle İndeksi

^aKi-kare testi

^bFisher kesin testi

^cStudent's t-testi

Çocuğun beslenme alışkanlığına ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması Çizelge 3.39’de verilmiştir. Bebeklik döneminde sağlıklı beslenme (ilk 6 ay sadece anne sütü+6 aydan sonra ek gıdalara geçiş+ilk 1 yıl anne sütü alan bebekler sağlıklı olarak değerlendirilmiştir) ile çocuklarında malnütrisyon görülmesi arasında anlamlı ilişki bulunamamıştır ($P=0,357$). Et tüketimi (kırmızı et tüketen+en az bir beyaz et tüketen çocuklar sağlıklı olarak değerlendirilmiştir), sebze ve meyve tüketimi ile malnütrisyon arasında anlamlı ilişki bulunmuştur (sırasıyla; $P=0,011$, $P<0,001$, $P=0,004$). Günlük süt tüketimi (paket sütler veya normal inek sütü) ile malnütrisyon arasında anlamlı ilişki bulunamamıştır ($P=0,447$). Paketli gıda tüketim sıklığı ile malnütrisyon arasında anlamlı ilişki bulunamamıştır ($P=0,125$).

Çizelge 3.39. Çocuğun beslenme alışkanlığına ilişkin özelliklerin malnütrisyon olan ve olmayanlarda karşılaştırılması

Değişkenler	Malnütrisyon		P değeri
	Yok (n=345)	Var (n=147)	
Bebeklik döneminde sağlıklı beslenme durumu*			
Evet	172 (%69,9)	74 (%30,1)	0,357 ^a
Hayır	203 (%73,6)	73 (%26,4)	
Et tüketimi**			
Yeterli	130 (%79,3)	34 (%20,7)	0,011 ^a
Yetersiz	245 (%68,4)	113 (%31,6)	
Sebze tüketimi			
Yeterli	260 (%75,1)	86 (%24,9)	0,019 ^a
Yetersiz	115 (%65,3)	61 (%34,7)	
Meyve tüketimi			
Yeterli	295 (%75,1)	98 (%24,9)	0,004 ^a
Yetersiz	80 (%62)	49 (%38)	
Günlük süt tüketimi			
İçer	314 (%72,5)	119 (%27,5)	0,447 ^a
İçmez	61 (%68,5)	28 (%31,5)	
Paketli gıda tüketim sıklığı***			
Nadiren	26 (%83,9)	5 (%16,1)	0,125 ^a
Sık sık	349 (%71,1)	142 (%28,9)	

*İlk 6 ay sadece anne sütü alan, ek gıdalara 6 aydan sonra geçen ve en az 12 ay anne sütü alan bebekler sağlıklı beslenmiş sayılır.

** Et tüketiminde Kırmızı et ve en az bir beyaz eti (tavuk veya balık) yeterli düzeyde tüketen çocukların et tüketimi **yeterli**, beyaz eti yetersiz/kısmen yeterli düzeyde tüketen ve sadece kırmızı eti yeterli düzeyde tüketen, kırmızı eti yetersiz/kısmen yeterli düzeyde tüketen ve en az bir beyaz eti yeterli düzeyde tüketen, kırmızı ve beyaz eti yetersiz/kısmen yeterli düzeyde tüketen çocukların et tüketimi **yetersiz** sayılır.

***Paketli gıda tüketiminde sık sık “Haftada en az bir kez”i tanımlarken, Nadiren ifadesi “Ayda bir veya birkaç kez”i tanımlamaktadır.

^aKi-kare testi

^bFisher kesin testi

Çocuğun beslenme davranışları ve diğer etkili faktörlerin malnütrisyon olan ve olmayanlarda karşılaştırılması Çizelge 3.40’de sunulmuştur.

Yemek sırasında tabak kaşık görünce öğürüp öğürmeme durumu ile malnütrisyon arasında anlamlı ilişki bulunamamıştır ($P=0,777$).

Yemek yerken lokmasını ağzında bekletme durumu ve düzenli öğle yemeği yeme durumu ile malnütrisyon arasında anlamlı ilişki bulunmuştur (sırasıyla; $P=0,003$, $P=0,044$). Çocuğun yemek yeme davranışı, yemek yeme alışkanlığına karşı ebeveyn tutumu ile malnütrisyon arasında anlamlı ilişki bulunamamıştır (sırasıyla; $P=0,106$, $P=0,071$). Hareketli olup olmama durumu, günlük ekran başında kalma süresi ve sağlıklı uyku düzeni ile malnütrisyon arasında anlamlı ilişki bulunamamıştır (sırasıyla; $P=0,586$, $P=0,718$, $P=0,364$). Sık boğaz enfeksiyonu, öksürük ve ateşlenme, sık karın ağrısı, sık ishal, sık kabız ve sık sık kusma olup olmama durumları ile malnütrisyon arasında anlamlı ilişki bulunamamıştır (sırasıyla; $P=0,653$, $P=0,929$, $P=0,946$, $P=0,241$, $P=0,721$). Çocukta veya ailedeki herhangi birinde parazit, kıl kurdu ve şerit görülüp görülmemesi durumu ile çocukta malnütrisyon görülmesi arasında anlamlı ilişki bulunamamıştır ($P=0,514$). Kronik hastalık varlığı ile malnütrisyon arasında anlamlı ilişki bulunamadı ($P=0,370$).

Çizelge 3.40. Çocuğun beslenme davranışları ve diğer etkili faktörlerin malnütrisyon olan ve olmayanlarda karşılaştırılması

Değişkenler	Malnütrisyon		P değeri
	Yok (n=375)	Var (n=147)	
Yemek sırasında tabak veya kaşık görünce öğürür mü?, n(%)			
Hayır	352 (%72)	137 (%28)	0,777 ^a
Evet	23 (%6,7)	10 (%3,3)	
Yemek yerken lokmasını uzun süre ağzında bekletir mi? (15 saniyeden uzun) , n(%)			
Hayır	254 (%66,3)	79 (%23,7)	0,003 ^a
Evet	121 (%33,7)	68 (%20,3)	
Düzenli öğle yemeği yeme durumu, n(%)			
Okulda/evde/kreşte yeme	370 (%72,5)	140 (%27,5)	0,044 ^b
Atıştırmalıklarla geçiştirme	5 (%1,7)	7 (%3,3)	
Çocuğun yemek yeme davranışı, n(%)			
Aile ile sofrada kendisi yer	286 (%73,7)	102 (%26,3)	0,106 ^a
Diğer	89 (%23,3)	45 (%11,7)	
Yemek yeme alışkanlığına karşı ebeveyn tutumu*, n(%)			
Doğru	133 (%35,6)	40 (%11,7)	0,071 ^a
Yanlış	242 (%64,4)	107 (%28,3)	
Sizce hareketli bir çocuk mu?, n(%)			
Normal düzeyde veya yavaş	219 (%58,4)	82 (%21,1)	0,586 ^a
Hareketli	156 (%41,6)	65 (%16,9)	
Günlük ekran başında kalma süresi, n(%)			
≤ 2 saat	236 (%62,9)	95 (%24,5)	0,718 ^a
> 2 saat	139 (%37,1)	52 (%13,4)	
Sağlıklı uyku düzeni, n(%)**			
Evet	233 (%62,2)	85 (%21,8)	0,364 ^a
Hayır	142 (%37,8)	62 (%16,2)	
Sık boğaz enfeksiyonu öksürük ve ateşlenme olur mu?, n(%)			
Hayır	140 (%37,6)	58 (%15,0)	0,653 ^a
Evet	235 (%62,4)	89 (%22,8)	
Sık karın ağrısı olur mu?, n(%)			
Hayır	228 (%60,8)	90 (%23,2)	0,929 ^a
Evet	147 (%39,2)	57 (%14,6)	
Sık ishal olur mu?, n(%)			
Hayır	340 (%90,4)	133 (%33,7)	0,946 ^a
Evet	35 (%9,6)	14 (%3,6)	
Sık kabız olur mu?, n(%)			
Hayır	283 (%75,2)	105 (%26,5)	0,266 ^a
Evet	91 (%24,8)	43 (%11,1)	
Sık sık kusar mı?, n(%)			
Hayır	338 (%90,1)	134 (%33,7)	0,721 ^a
Evet	37 (%9,9)	13 (%3,3)	
Çocukta veya ailedeki herhangi birinde parazit kıl kurdu veya şerit görüldü mü?, n(%)			
Hayır	315 (%83,8)	120 (%30,6)	0,514 ^a
Evet	60 (%16,2)	27 (%6,9)	
Bilinen kronik bir hastalığı var mı?, n(%)			
Yok	326 (%86,9)	132 (%33,7)	0,370 ^a
Var	49 (%13,1)	15 (%3,9)	

* Yemek yeme alışkanlığına karşı ebeveyn tutumu olarak **Doğru** kabul edilen davranışlar: Müdahale etmeyerek acıktığında yemesine izin verenler **Yanlış** kabul edilen davranışlar: Kızarak zorla yedirme, başka yemek sunma, cezalandırma veya yemesi için ödül teklif etme

** Sağlıklı uyku düzeninde **Evet** ifadesi sağlıklı kabul edilen durumu: **2-5 yaş için**; gündüz uyuyan ve gece 9 saat ve üzeri uyuyan çocukları ve **6-10 yaş için**; gündüz 1 saat ve daha az uyuyan ve gece 8 saat ve üzeri uyuyan çocukları, **Hayır** ifadesi sağlıklı olmayan olarak kabul edilen durumu: **2-5 yaş için**; gündüz uyumaz, gece 9 saatten az uyuyan çocukları ve **6-10 yaş için**; gündüz 1 saatten fazla uyuyan, gece 8 saatten az uyuyan çocukları gösterir

^aKi-kare testi

^bFisher kesin testi

3.2.1. Lojistik regresyondan elde edilen skorlar ile klinik tahmin modeli

Tek deęişkenli analizde P deęeri 0,25'den düşük bulunan risk faktörleri çoklu lojistik regresyon modeline, aday deęişken olarak alınmıştır. $P < 0,25$ olduęu halde ilişki olduęu düşünölen deęişkenlerden biri seçilmiştir. Cinsiyet ve yaş deęişkenleri $P > 0,25$ olmasına karşın klinik açıdan önemli olduęu için modele eklenmiştir. Modele alınan deęişkenler aşağıda verilmiştir.

- ✓ Cinsiyet
- ✓ Yaş
- ✓ Yaşıtlarına göre boyu
- ✓ Yaşıtlarına göre kilosu
- ✓ Doğum ağırlığı
- ✓ Annenin eğitim durumu
- ✓ Anne çalışıyor mu?
- ✓ Yemek yerken lokmasını uzun süre ağızında bekletir mi?
- ✓ Düzenli öğle yemeęi yeme
- ✓ Çocuęun yemek yeme davranışı
- ✓ Yemek yeme alışkanlığına karşı ebevy tutumu
- ✓ Paketli gıda tüketimi
- ✓ Baba VKİ
- ✓ AnneVKİ
- ✓ Yeterli sebze tüketimi
- ✓ Yeterli meyve tüketimi
- ✓ Yeterli et tüketimi
- ✓ Ailede yaşayan toplam kişı sayısı

Çoklu lojistik regresyon analizi sonucunda istatistiksel olarak anlamlı bulunan deęişkenler için beta katsayıları, Wald istatistięi, SH, odds oranları (OR) ve %95 güven aralıkları Çizelge 3.41'de sunulmuştur.

Çizelge 3.41. Malnütrisyon öngörüsünde etkili olan risk faktörlerini belirlemek amacıyla gerçekleştirilen çoklu lojistik regresyon analiz sonuçları

	95% GA							
	OR							
	B	S.H.	Wald	sd	p	OR	Alt	Üst
Yaşıtlarına göre boyu	1,438	0,372	14,922	1	<0,001	4,211	2,030	8,732
Yaşıtlarına göre kilosu	3,075	0,482	40,669	1	<0,001	21,644	8,413	55,686
Paketli gıda tüketim	1,856	0,934	3,952	1	0,047	6,401	1,027	39,908
Ailede yaşayan toplam kişi sayısı	1,557	0,535	8,476	1	0,004	4,744	1,663	13,530
Constant	-5,445	1,070	25,915	1	<0,001	0,004		

LR'ye ilişkin çoklu modelden elde edilen β ve OR'lardan 4 ayrı yaklaşım kullanılarak klinik tahmin modeli için skorlar elde edilmiştir. Bunlar Beta, tamsayı Beta, OR, tamsayı OR modellerinden elde edilen skorlar Çizelge 3.42).

Çizelge 3.42. Beta, tamsayı Beta, OR, tamsayı OR

Risk Faktörü		Beta	Tamsayı Beta	OR	Tamsayı OR
Yaşıtlarına göre boyu (X ₁)	Aynı veya daha uzun	-	-	-	-
	Yaşıtlarından kısa	1,438	1	4,211	4
Yaşıtlarına göre kilosu(X ₂)	Aynı veya daha kilolu	-	-	-	-
	Yaşıtlarından zayıf	3,075	3	21,644	22
Paketli gıda tüketim (X ₃)	Nadiren	-	-	-	-
	Sık sık	1,856	2	6,401	6
Ailede yaşayan toplam kişi sayısı (X ₄)	5 ve daha az kişi	-	-	-	-
	6 ve daha fazla kişi	1,557	2	4,744	5

LR'ye dayalı oluşturulan farklı yaklaşımlardan elde edilen klinik tahmin modelleri aşağıdaki gibidir.

Skor 1 (β)

$$Y = 1,438 * X_1 + 3,075 * X_2 + 1,856 * X_3 + 1,557 * X_4$$

Skor 2 (β tamsayı)

$$Y = 1 * X_1 + 3 * X_2 + 2 * X_3 + 2 * X_4$$

Skor 3 (β / β 'nın en küçük mutlak değeri (tamsayı))

$$Y = 1 * X_1 + 2 * X_2 + 1 * X_3 + 1 * X_4$$

Skor 4 (OR)

$$Y = 4,211 * X_1 + 21,644 * X_2 + 6,401 * X_3 + 4,744 * X_4$$

Skor 5 (OR tamsayı)

$$Y = 4 * X_1 + 22 * X_2 + 6 * X_3 + 5 * X_4$$

Geliştirilen yeni klinik tahmin modellerinin malnütrisyon tanısındaki başarısını karşılaştırmak için gerçekleştirilen ROC analizi sonucu belirlenen kesim noktası değerleri ile hesaplanan duyarlılık, seçicilik, doğruluk, pozitif-negatif tahmini değerler, F1 skorları ve MCC değerleri Çizelge 3.46’da verilmiştir. ROC analizi sonucu göre malnütrisyon Beta, Tamsayılı Beta, OR, Tamsayılı OR malnütrisyon öngörüsünde anlamlı olarak bulunmuştur



Çizelge 3.45. Skorun malnütrisyon tahminindeki AUC, duyarlılık, seçicilik, PTD, NTD, doğruluk, F1 skoru ve MCC

Performans								
ölçüsü	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
(%95)								
Beta	0,835	0,858	0,719	0,547	0,928	0,759	0,668	0,524
	(0,800-	(0,802-	(0,674-	(0,483-	(0,898-	(0,722-	(0,607-	(0,414-
	0,869)	0,914)	0,765)	0,611)	0,0957)	0,795)	0,729)	0,633)
Tamsayılı Beta	0,834	0,872	0,711	0,544	0,933	0,757	0,670	0,527
	(0,800-	(0,818-	(0,665-	(0,480-	(0,904-	(0,720-	(0,610-	(0,419-
	0,869)	0,926)	0,757)	0,608)	0,962)	0,793)	0,730)	0,636)
OR	0,835	0,878	0,711	0,546	0,937	0,759	0,674	0,533
	(0,800-	(0,826-	(0,665-	(0,483-	(0,908-	(0,722-	(0,614-	(0,426-
	0,869)	0,931)	0,757)	0,609)	0,965)	0,795)	0,733)	0,641)
Tamsayılı OR	0,835	0,878	0,711	0,546	0,937	0,759	0,674	0,533
	(0,800-	(0,826-	(0,665-	(0,483-	(0,908-	(0,722-	(0,614-	(0,426-
	0,869)	0,930)	0,757)	0,609)	0,965)	0,795)	0,733)	0,641)

AUC: ROC eğrisi altındaki alan, GA: güven aralığı, PTD: pozitif tahmini değer, NTD: negatif tahmini değer

Klinik tahmin modeli tamsayılı beta kullanılarak oluşturulmuştur.

Çizelge 3.46. ML ve diğer algoritmalar ile oluşturulan model performanslarının karşılaştırılması

Makine Öğrenme algoritmaları ve diğer algoritmalar								
Performans ölçüsü (%95)	AUC	Duyarlılık	Seçicilik	PTD	NTD	Doğruluk	F1 skoru	MCC
GBM	0,906 (0,879-0,932)	0,993 (0,980-1)	0,832 (0,785-0,862)	0,690 (0,628-0,752)	0,997 (0,990-1)	0,872 (0,843-0,900)	0,814 (0,762-0,867)	0,749 (0,660-0,838)
XGBoost	0,841 (0,807-0,875)	0,878 (0,826-0,931)	0,711 (0,665-0,757)	0,546 (0,483-0,609)	0,937 (0,908-0,965)	0,759 (0,722-0,795)	0,674 (0,614-0,733)	0,534 (0,426-0,641)
LightGBM	0,841 (0,807-0,875)	0,878 (0,826-0,931)	0,711 (0,665-0,757)	0,545 (0,482-0,608)	0,936 (0,908-0,965)	0,758 (0,722-0,795)	0,673 (0,614-0,733)	0,533 (0,426-0,641)
Rastgele Orman	0,825 (0,777-0,873)	0,803 (0,738-0,867)	0,917 (0,889-0,945)	0,792 (0,726-0,857)	0,922 (0,895-0,949)	0,885 (0,858-0,912)	0,797 (0,733-0,862)	0,717 (0,605-0,829)
Karar Ağacı	0,860 (0,824-0,895)	0,802 (0,738-0,867)	0,917 (0,889-0,945)	0,792 (0,727-0,857)	0,922 (0,895-0,949)	0,885 (0,858-0,912)	0,797 (0,733-0,861)	0,717 (0,605-0,829)
DVM	0,839 (0,805-0,873)	0,878 (0,825-0,931)	0,711 (0,665-0,757)	0,546 (0,483-0,609)	0,937 (0,908-0,965)	0,760 (0,722-0,795)	0,674 (0,614-0,733)	0,533 (0,426-0,641)
Bilgi Kazancı	0,834 (0,799-0,870)	0,871 (0,816-0,924)	0,707 (0,660-0,753)	0,538 (0,474-0,601)	0,933 (0,904-0,962)	0,753 (0,716-0,790)	0,665 (0,605-0,725)	0,521 (0,413-0,630)
L1 (Lasso)	0,840 (0,807-0,875)	0,878 (0,825-0,931)	0,711 (0,665-0,757)	0,546 (0,483-0,609)	0,937 (0,908-0,964)	0,759 (0,722-0,795)	0,674 (0,614-0,733)	0,533 (0,426-0,641)
Bayesci GDM	0,907 (0,880-0,933)	0,993 (0,980-1)	0,824 (0,784-0,862)	0,690 (0,628-0,752)	0,997 (0,990-1)	0,872 (0,843-0,900)	0,814 (0,762-0,866)	0,750 (0,660-0,838)

4. TARTIŞMA

Son yıllarda klinik tahmin modellerinde makine öğrenme, derin öğrenme ve yapay zekaya dayalı algoritmalarının kullanımı çok fazla yaygınlaşmıştır. Raita vd. (2019) acil servis triyajının klinik sonuçları öngörmesi, kritik hasta ve stabil hastaları doğru bir şekilde ayırt etme ve önceliklendirme konusunda dört farklı makine öğrenimi modeli (Lasso regresyon, rastgele orman, gradyan arttırılmış karar ağacı ve derin sinir ağı) ile lojistik regresyon analizine dayalı geliştirdikleri model bulgularını AUC değerlerini kullanarak karşılaştırmıştır (Raita, 2019).

Kia vd. (2020) hastanede yatan hastalarda altı saat önce klinik kötüleşme tahmininin belirlenmesi amacıyla rastgele orman (RF), doğrusal destek vektör makinesi ve lojistik regresyon kullanılarak oluşturulan 3 farklı klinik tahmin modelinin performansını, duyarlılık, seçicilik, doğruluk, PTD, ROC AUC ve F1 skorlarını kullanarak karşılaştırmışlardır (Kia vd., 2020).

Sanchez-Pinto vd. (2022) klinik tahmin modelleri üzerinde odaklanan araştırmacıların, genellikle performansı önemseyişini ve özellikle derin öğrenme gibi yeni makine öğrenimi yöntemlerinin olağanüstü doğruluk vaadi sunduğunu belirtmişlerdir. Bu yeni yaklaşımların, klinik karar verme süreçlerinde kullanılan modeller arasında yerli yorumlanabilirlik eksikliği bulunan “kara kutu” modellerin giderek artan bir ilgiyle tercih edilmesine neden olduğunu vurgulamışlardır. Ancak, bu tür modellerin şeffaflık eksikliği, klinisyenlerin yanlış varsayımlara dayalı hatalı kararlar almasına sebep olabileceğini ve bu durumun modellerin gerçek dünya uygulamalarında etkisiz kalmasına yol açabileceğini önemli bir sorun olarak belirtmişlerdir (Sanchez-Pinto vd., 2022).

Ayrıca, yapılan bir incelemede, daha az karmaşık modellerin, örneğin lojistik regresyon ve karar ağaçları gibi yorumlanabilir modellerin, sinir ağları ve destek vektör makineleri gibi daha az yorumlanabilir modellere kıyasla benzer performans sergilediği ve bazen sadece ROC AUC'deki %2'lik bir farkla daha iyi olduğu tespit edilmiştir. Yapılandırılmış veri ve anlamlı özelliklere sahip 48 klinik tahmin çalışmasının bir incelemesinde, yorumlanabilir modeller (örneğin, lojistik regresyon, karar ağaçları, vb.) daha az yorumlanabilir modellerle (örneğin, sinir ağları, destek vektör makineleri, vb.) aynı düzeyde performans göstermiştir (Uddin vd. 2019, Sanchez-Pinto vd. 2022). Bu durum

(Sanchez-Pinto vd., 2022) tarafından aynı veri setinde makul bir doğruluk elde eden birden fazla modelin olduğunu öne süren "Rashomon etkisi" ile açıklanmıştır (Rudin C vd., 2019, Sanchez-Pinto vd. 2022). Bu sonuçlar, daha yorumlanabilir ve basit modellerin, gelecekteki doğrulama ve klinik uygulamalar için tercih edilmesi gerektiğini göstermektedir. Araştırmacılar, bu seçimi doğrulamak için kara kutu modellerini iyi yapılandırılmış ve yorumlanabilir modellerle karşılaştırmalıdır. Eğer daha az yorumlanabilir modellerin tercih edilmesi gerekliyse, açıklanabilirlik yöntemlerinin etkin bir şekilde uygulanması gerektiğini vurgulamışlardır (Sanchez-Pinto vd. 2022).

Christodoulou vd (2019) 1/2016 ile 8/2017 tarihleri arasında yayınlanmış 927 makaleden dahil edilme kriterlerine uyan 71 makale ile gerçekleştirdikleri sistematik inceleme sonucunda ML'nin LR'ye göre üstün performansına dair hiçbir kanıt bulamadıklarını belirtmiştir. Araştırma sonucunda, modelleme algoritmalarını karşılaştıran çalışmalar için metodoloji ve raporlamada iyileştirmelere ihtiyaç olduğu ayrıca vurgulanmıştır (Christodoulou vd. 2019).

Literatürde makine öğrenmesi ile gerçekleştirilen çalışmalarda modellerin klinik fayda ve açıklanabilirlik konusunda eksikleri olduğu görülmektedir. Lojistik regresyon ile oluşturulan modellerde beta, odds veya bunların tam sayılı skorlara dönüştürülerek elde edilen klinik tahmin modellerinde, bu modelleri kullanacak bireyler veya araştırmacılar hangi değişkenlere veya hangi sağlık durumlarına nasıl skor vereceği bilinmemektedir. Ancak makine öğrenme algoritmaları ile gerçekleştirilen modellerde modelin arkasındaki sınıflama doğrudan araştırmacılar tarafından görülememiği için (Sanchez-Pinto vd. 2022) "kara kutu" olarak adlandırdığı ifadedeki gibi geliştirilen modeller sadece sınıflama başarılarına odaklanmaktadır.

Çalışmamızda literatür dikkate alınarak özellikle klinik tahmin modellerinde hem açıklanabilirlik hem de yorumlanabilirlik özelliği yüksek olan lojistik regresyon modeline odaklanıldı ve lojistik regresyon modeli ile oluşturulan klinik tahmin modelindeki katsayıların ağırlıklarının belirlenmesinde makine öğrenimi ve bazı diğer algoritmaların katkısı belirlendi. Araştırma bulgularımız birlikte değerlendirildiğinde çalışmamızda öncelikle doğrudan makine öğrenme algoritmaları ile sınıflama yaparak model kurmak yerine iki düzeyli bir yanıtı sahip tanı konulacak klinik tahmin modellerinde lojistik regresyon modeli ile kurulabilecek tüm modelleri farklı senaryolar ile karşılaştırdık. Ayrıca lojistik regresyona alternatif olarak klinik skora için risk faktörlerine modelde verilecek

ağırlıkları veya skorları farklı makine öğrenme algoritmaları ile elde ettik ve tüm model sonuçlarını karşılaştırdık.

Bu çalışmada yeni bir yöntem olarak klinik tahmin modellerinde makine öğrenme algoritmaları ile modeli oluşturmak yerine makine öğrenme algoritmaları ile risk faktörleri için skorların verilebileceği önerilmektedir.

2-10 yaş arası çocuklarda malnütrisyon tanısını koymak için z skorları kullanılabilir, ancak bu yaklaşımın hem ebeveynler hem de doktorlar için bazı zorlukları bulunmaktadır. Z skorları, çocuğun boy, kilo gibi ölçümlerini yaş ve cinsiyetle karşılaştırarak istatistiksel bir değer elde etmeye dayanır. Ancak bu yöntem güvenilir olsa da bazı zorlukları vardır. Hesaplama zorluğu bunların başında gelmektedir. Z skorlarının hesaplanması, özel formüller ve standart tablolar gerektirir. Bu hesaplamaların yapılması için belirli bir uzmanlık gerekmektedir. Özellikle pratik tıp ortamlarında hızlı ve doğru hesaplama yapabilmek zaman alıcıdır. Ebeveynler için anlaşılabilirlik diğer bir kısıtlılık olarak söylenebilir. Ebeveynler için z skorları ve istatistiksel terminoloji anlaşılması zor olabilir. Çocuğun sağlığıyla ilgili verilen bilgilerin anlaşılabilir olması önemlidir. Tek boyutluluk özelliği de diğer bir kısıtlılıktır. Z skorları genellikle boy-kilo ilişkisine dayanır. Bu, çocuğun genel sağlık durumu hakkında tam bir resim sunmaktan ziyade sadece büyüme düzeyini gösterir. Ayrıca çocuğun boy ve kilosunun doğru şekilde ölçülmemesi, yanlışlıkla yapılan ölçümler veya standart olmayan ölçüm ekipmanları gibi faktörler, z skorlarının doğruluğunu etkileyebilir.

Malnütrisyon tahmini için z skorları dışında literatürde malnütrisyon ile ilişkilendirilmiş; ön kol çevresi, doğum ağırlığı, anne ve baba VKİ (vücut kitle indeksi), sosyoekonomik durum, beslenme alışkanlıkları, hastalık geçmişi, yaş ve cinsiyet, ailedeki beslenme kültürü erişilebilirlik ve besin güvenliği ve ebeveyn tahminleri gibi bir dizi risk faktörü belirlenmiştir.

Ön kol çevresi, kas kütlesi ve yağ dağılımı gibi faktörlerle ilişkilendirilmiştir. Özellikle çocuklarda ölçülen ön kol çevresi, beslenme durumu hakkında bilgi sağlayabilir. Doğum ağırlığı, çocuğun doğumda aldığı besin ve annenin gebelik sürecindeki beslenme durumuyla ilişkilendirilebilir. Düşük doğum ağırlığı, çocuğun beslenme riskini işaret edebilir. Anne ve baba VKİ'leri, genetik yatkınlık ve ailedeki beslenme alışkanlıkları gibi faktörlerle ilişkilendirilir ve çocuğun beslenme durumunu etkileyebilir. Çocuğun günlük beslenme alışkanlıkları, tüketilen besin çeşitliliği, öğün atlamaları gibi faktörler beslenme

durumunu etkileyebilir. Kronik hastalıklar, enfeksiyonlar veya sindirim sistemi problemleri gibi sađlık sorunları, beslenme alımını etkileyebilir ve malnütrisyon riskini artırabilir. Yaş ve cinsiyet, çocukların büyüme ve gelişme sürecinde önemli birer faktördür. Özellikle büyüme hızı ve büyüme dönemleri malnütrisyon riskini etkileyebilir. Ailenin geleneksel beslenme alışkanlıkları, çocuđun beslenme durumunu etkileyebilir. Örneđin, aile içindeki besin tercihleri ve yemek hazırlama yöntemleri gibi faktörler önemli olabilir. Sağlıklı ve dengeli beslenmeyi destekleyen gıdalara erişim, besin güvenliđi ve çocuđun beslenme durumunu doğrudan etkileyebilir. Ebeveynlerin çocuklarının beslenme durumu hakkındaki deđerlendirmeleri, aile içi beslenme alışkanlıkları, çocuđun yeme alışkanlıkları ve sađlık durumu gibi faktörler açısından önemli ipuçları verebilir. Ailenin gelir düzeyi, eğitim seviyesi, yaşadığı çevre gibi sosyoekonomik faktörler, çocuđun beslenme durumu üzerinde etkili olabilir.

Bu faktörler, çocuđun beslenme durumuyla ilişkilendirilebilir ancak tek başlarına malnütrisyonu doğrudan belirlemezler. Çocuđun beslenme durumunu deđerlendirmek için genellikle birden fazla faktörün bir arada deđerlendirilmesi gereklidir. Çalışmamızda malnutrsiyon tahmini için geliştirilen klinik tahmin modeli ile bu faktörler, z skorlarına alternatif olarak kullanılmıştır.

5. SONUÇ VE ÖNERİLER

Klinik tahmin modellerinin başarısı, yalnızca performans ölçütleri ile sınırlı değildir; aynı zamanda uygulanabilirlik, anlaşılabilirlik, doğruluk ve genelleme yeteneği gibi faktörler de hayati öneme sahiptir. Modelin uygulanabilirliği, tıbbi uygulamalarda pratik kullanımı sağlama yeteneğiyle ilgilidir. Örneğin, bir klinikte kullanılacak bir modelin, gerçek zamanlı verilere kolay erişim sağlaması ve hızlı bir şekilde sonuç üretmesi gerekebilir. Aynı zamanda, modelin anlaşılabilirliği, tıbbi profesyonellerin ve karar vericilerin modelin karar süreçlerini anlamalarını sağlamalıdır. Bir kanser teşhis modelinde, hangi özelliklerin kanser tahmininde ne kadar etkili olduğunu anlamak, bu modelin anlaşılabilirliği açısından kritik olabilir. Doğruluk ve güvenilirlik, modelin doğruluk oranları üzerinden değerlendirilirken, genelleme yeteneği ise modelin farklı popülasyonlar veya farklı klinik ortamlarda nasıl performans göstereceğini belirler. Örneğin, bir kalp ritmi bozukluğu tahmin modelinin, farklı yaş grupları veya coğrafi bölgelerdeki hastalar üzerindeki performansı genelleme yeteneğini belirlemek açısından önemlidir. Sonuç olarak, klinik tahmin modellerinin etkili olması, yalnızca iyi bir performans sergilemekle değil, aynı zamanda gerçek dünya koşullarında uygulanabilir, anlaşılabilir, doğru ve genellemeye açık olma özellikleriyle de ilişkilendirilir.

Bu çalışmada öncelikle iki düzeyli sonuç veren klinik tahmin modellerinin oluşturulmasında gerçek araştırmalarda yaygın olarak karşılaşılan farklı senaryolar altında lojistik regresyon analizi, makine öğrenme algoritmaları veya bazı diğer algoritmaların performansları karşılaştırılmıştır. Bulgularımıza göre; Modellerin performansı farklı senaryolarda farklı örneklem büyüklükleri ile bağımlı değişken 1-0 oranlarının %20-%80 ve %50-%50 olduğu ve bağımsız değişkenlerin 1-0 oranlarının %20-%80, %50-%50 ve rastgele olduğu durumlarda elde edilen tüm sonuçlarımızın 4 ayrı performans ölçüsü ile değerlendirildiğinde tutarlı sonuçlar verdiği gözlenmiştir.

Örneklem büyüklüğünün 200, 500 ve 1000 olarak alındığı tüm senaryolar göz önüne alındığında ilk olarak odds oranlarının dengeli olduğu (A senaryosu) ve bağımlı-bağımsız değişkenin 1-0 oranının %20-%80 olduğu durum gözönüne alındığında model performanslarında lojistik regresyonu temel alan skorlama sistemleri kullanıldığında sonuçlar birbirlerine oldukça yakın olmakla birlikte gerçek katsayılar tamsayıya çevrildiğinde skorların ağırlığında küçük düşüşler gözlenmektedir. Örneğin; AUC'de beta 0,763 değeri tamsayı beta da 0,745'e düşmüştür. Aynı şekilde odds oranları 0,759'dan

0,757'ye düřtüęü gözlemlenmiřtir. Bu düřüřler betalardan tamsayılı betalara geçiřte ya da OR'den tamsayılı OR'ye geçiřte ihmal edilebilir düzeyde olduęu gözlenmiřtir. Bu durum klinikte kolaylık saęlama aısından tamsayılı versiyonlarının kullanılabilceęini göstermektedir. İlgili senaryoda yani dengeli tasarımlara bakıldıęında makine öğrenmesi yöntemleri kendi içlerinde deęerlendirildięinde en düşük performansın rastgele orman ve karar aęacı olduęu gözlenmektedir ve bunların aynı senaryo içinde örneklem büyüklüęü deęiřtirildięi durumda performans sonuçlarında farklılık gözlenmedięi görülmektedir. Simülasyon alıřmamızda örneklem büyüklüęü $n=200$, $n=500$ ve $n=1000$ alındıęında yöntemlerin kendi içlerinde büyük farklılıklar gözlemlenmedięi ortaya ıkarılmıřtır. Aynı senaryo içerisinde Baęımsız deęiřken oranı %50-%50 olduęunda %20-%80 göre performansın arttıęı gözlenmiřtir (Örneęin; beta için F1 deęeri 0,51 den 0,59 ıkmıřtır). Baęımsız deęiřken oranları dengeliye doęru gittike model performansları artmaktadır. Lojistik regresyonda betanın ve OR direkt olarak kullanıldıęı modellerde performans daha yüksektir ve ayrıca bayesci GDM ve L1 (Lasso) modellerinde de performans bunlara benzer orandadır.

Tüm senaryolar incelendięinde tüm senaryolarda $n=200$, $n=500$ ve $n=1000$ denek sayısı ile tekrarlanmıřtır. Tüm senaryolar gözönüne alındıęında 3 farklı örneklem büyüklüęü için model performanslarının büyük bir deęiřiklik göstermedięi gözlenmektedir. İkinci durum ise dengeli dengesiz tasarımlarda yine büyük farklılıklar gözlenmemiřtir. Bu durumda arařtırmacıların dengesiz tasarımlar kullanırken ekinmemeleri önerilebilmektedir. Model performansını etkileyen durumun ise baęımlı deęiřkende 1-0 oranının ilgilenilen olay sıklıęının homojen olduęu %50-%50 olduęu durumlardır. Senaryolar içinde model performansı en fazla olduęu durum baęımlı deęiřkenin homojen olduęu durumdur. Elbette baęımsız deęiřkenlerin homojen olduęu durumda da model performansı artmaktadır. Hem baęımlı deęiřkenlerde hem baęımsız deęiřkenlerin %50-%50 olduęu durumda model performansları en yüksek seviyeye ulařılmıřtır. Baęımlı deęiřkenlerde %50-%50 arařtırmacılar tarafından saęlanabilir ancak tüm baęımsız deęiřkenlerin %50-%50 olduęu durumun saęlanması mümkün deęildir. alıřmamızda baęımsız deęiřken oranları rastgele olarak alınmıřtır (0,1-0,33; 0,34-0,66 ve 0,67-0,9 aralıęında alınmıřtır). Burada tüm senaryolarda performanslar düşmekle birlikte yine de klinik tahmin modeli kurabilecek düzeydedir. ok ok büyük bir düşüř olmamıřtır. Hem baęımlı hem baęımsız deęiřkenlerde %20-%80 olduęu durumlar model performanslarının en düşük olduęu durumlardır.

Çalışmamızda simülasyon ile karşılaştırılan modeller gerçek veri seti üzerinde de karşılaştırılmıştır. Malnütrisyon tanısı koymak için 522 kişiden veriler toplanmıştır. Bu kişilerin 147 sine malnütrisyon tanısı konmuştur. Çalışmamızda bağımlı değişken oranı %28-%72 ve bağımsız değişkenler rastgele olan modele uymaktadır. Anne babaların kendi kendilerine ölçebilecekleri değişkenler ile basit bir klinik tahmin modeli kurulması amaçlanmıştır. Klinik tahmin modelleri adımları uygulanarak anlamlı bulunan değişkenler ile model kurulmuştur. Simülasyon çalışmasıyla benzer sonuçlar bulunmuştur. Betalar ve OR oranları tamsayıya dönüştürüldüğünde performansta ihmal edilebilir düzeyde çok küçük düşüşler olmaktadır. Bu nedenle tamsayıli skorları tercih edilmesi kullanım kolaylığı için önerilmektedir.

Çalışmamızda ikinci olarak Pediatri alanında 522 çocuktan titizlikle toplanan kapsamlı bir veri seti ile benzetim çalışmasında kullanılan tüm yöntemler ile Malnütrisyon durumunu tahmin eden klinik tahmin modeli geliştirildi. En iyi başarı Lojistik regresyon analizi ile Beta katsayıları kullanılarak elde edildi. Ancak klinik tahmin modellerinin kullanılabilirlik özelliği düşünülerek ikincil olarak en başarılı model tam sayılı odds oranlarına dayalı olarak geliştirilen klinik tahmin modeli oldu.

Z skorları, çocuklarda malnütrisyonun belirlenmesinde kullanılan yaygın bir yöntemdir ve genellikle güvenilirdir. Ancak doğru ve güvenilir sonuçlar için doğru ölçümler, hesaplamalar ve profesyonel yorumlar gereklidir. Ebeveynlerin ve doktorların bu değerleri anlaması ve kullanması önemlidir, ancak bazı durumlarda doğru yorumlamak veya hesaplamak zor olabilir. Bu nedenle, tek başına z skorlarına dayalı olarak malnütrisyon tanısı koymak yerine, literatürde malnütrisyon ile anlamlı ilişkisi gösterilmiş risk faktörleri ve ebeveynlerin bilgilerine dayalı skorlar ile desteklenerek geliştirilen malnütrisyon tahmin modelinin z skorları ile birlikte değerlendirilmesi önerilmektedir.

KAYNAKLAR

- Adams, S.T and Leveson, S.H. (2012). Clinical prediction rules. *BMJ*, **344**: d8312. <https://doi.org/10.1136/bmj.d8312>
- Alabdullah, A.A., Iqbal, M, Zahid, M., Khan, K., Amin, M.N., Jalal, F.E. (2022). Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis. *Construction and Building Materials*, 345: 128296. <https://doi.org/10.1016/j.conbuildmat.2022.128296>
- Alanazi, H.O., Abdullah, A.H., Qureshi, K.N. (2017). A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *Journal of medical systems*, 41: 1-10. <https://doi.org/10.1007/s10916-017-0715-6>
- Alvarado, A. (1986). A practical score for the early diagnosis of acute appendicitis. *Annals of Emergency Medicine*, 15(5): 557-564. [https://doi.org/10.1016/S0196-0644\(86\)80993-3](https://doi.org/10.1016/S0196-0644(86)80993-3)
- Arostegui, I., Legarreta, M.J., Barrio, I., Esteban, C., Garcia-Gutierrez S., Aguirre U., Quintana J.M. (2019). A computer application to predict adverse events in the short-term evolution of patients with exacerbation of chronic obstructive pulmonary disease. *JMIR Medical Informatics*, 7(2): e10773 <https://doi.org/10.2196/10773>
- Austin, P.C, Steyerberg, E.W (2012). Interpreting the concordance statistic of a logistic regression model: relation to the variance and odds ratio of a continuous explanatory variable. *BMC Medical Research Methodology*, 12: 1-8. <https://doi.org/10.1186/1471-2288-12-82>
- Altuntaş, S.A (2022). *Makine öğrenmesi teknikleri ile yürüyüş verileri kullanılarak Parkinson hastalığının ön tespiti.* (Yüksek lisans tezi). Trakya Üniversitesi Fen Bilimleri Enstitüsü,
- Aylak, B. L., Okan, O. R. A. L., & YAZICI, K. (2021). Yapay zeka ve makine öğrenmesi tekniklerinin lojistik sektöründe kullanımı. *El-Cezeri*, 8(1), 74-93. <https://doi.org/10.31202/ecjse.776314>
- Bang, H., Edwards, A.M., Bomback, A.S., Ballantyne, C.M., Brillon, D, Callahan, M.A., Teutsch, S.M., Mushlin AI, Kern, LM (2009). A patient self-assessment diabetes screening score: development, validation, and comparison to other diabetes risk assessment scores. *Annals of Internal Medicine*, 151(11): 775. <https://doi.org/10.1059/0003-4819-151-11-200912010-00005>
- Bonnett, LJ, Snell, K.I.E., Collins, G.S., Riley, R.D. (2019) Guide to presenting clinical prediction models for use in clinical settings. *BMJ*. **17**; 365: l737. <https://doi.org/10.1136/bmj.l737>
- Boon, I.S., AU Yong T.P., & Boon C.S. (2018). Assessing the role of artificial intelligence (AI) in clinical oncology: utility of machine learning in radiotherapy target volume delineation. *Medicines*, 5(4):131. <https://doi.org/10.3390/medicines5040131>
- Borgialli D.A., Mahajan P., Hoyle J.D., JR, Powell, E.C., Nadel, F.M., Tunik, M.G., Foerster, A., Dong L, Miskin M, Dayan PS, Holmes JF, Kuppermann N (2016). Performance of the pediatric Glasgow coma scale score in the evaluation of children with blunt head trauma. *Academic Emergency Medicine*, 23(8): 878-884. <https://doi.org/10.1111/acem.13014>
- Brunekreef, B., Smit, J., De Jongste, J., Neijens, H., Gerritsen, J., Postma, D., Van Strien, R. (2002). The prevention and incidence of asthma and mite allergy (PIAMA) birth cohort study: design and first results. *Pediatric Allergy and Immunology*, 13: 55-60 <https://doi.org/10.1034/j.1399-3038.13.s.15.1.x>
- Carrington, A. M., Fieguth, P. W., Qazi, H., Holzinger, A., Chen, H. H., Mayr, F., & Manuel, D. G. (2020). A new concordant partial AUC and partial c statistic for imbalanced data in the evaluation of machine learning algorithms. *BMC medical informatics and decision making*, **20**, 1-12. <https://doi.org/10.1186/s12911-019-1014-6>

- Celi, L. A., Majumder, M. S., Ordóñez, P., Osorio, J. S., Paik, K. E., & Somai, M. (2020) (2020). Leveraging Data Science for Global Health. *Springer Nature*, 199-217
- Cerqueira, F. R., Ferreira, T. G., de Paiva Oliveira, A., Augusto, D. A., Krempser, E., Barbosa, H. J. C. & Siqueira-Batista, R. (2014) NICeSim: an open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making. *Artificial intelligence in medicine*, 62(3), 193-201. <https://doi.org/10.1016/j.artmed.2014.10.001>
- Chao, C. M., Yu, Y. W., Cheng, B. W., & Kuo, Y. L. (2014). Construction the model on the breast cancer survival analysis use support vector machine, logistic regression and decision tree. *Journal of medical systems*, 38, 1-7. <https://doi.org/10.1007/s10916-014-0106-1>
- Chen L., (2020). Overview of clinical prediction models. *Annals of translational medicine*, 8(4): 71. [doi: 10.21037/atm.2019.11.121](https://doi.org/10.21037/atm.2019.11.121)
- Chen, T., & Guestrin, C. (2016). XGboost: A scalable tree boosting system. *In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- Chowdhury, M. Z. I., & Turin, T. C. (2020). Variable selection strategies and its importance in clinical prediction modelling. *Family medicine and community health*, 8(1). doi: 10.1136/fmch-2019-000262
- Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of clinical epidemiology*, 110, 12-22. <https://doi.org/10.1016/j.jclinepi.2019.02.004>
- Collins, G. S., de Groot, J. A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., & Altman, D. G. (2014). External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC medical research methodology*, 14(1), 1-11. <https://doi.org/10.1186/1471-2288-14-40>
- Cowley, L. E., Farewell, D. M., Maguire, S., & Kemp, A. M. (2019). Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. *Diagnostic and prognostic research*, 3, 1-23. <https://doi.org/10.1186/s41512-019-0060-y>
- Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications*, 115(9). DOI:10.5120/20182-2402
- Demir, E (2019). *Randomize Olmayan Klinik Çalışmalarda En Uygun Eşleştirme Analizi İçin Makine Öğrenme Algoritmaları İle Yeni Propensity Skor Tahmin Modellerinin Geliştirilmesi*, (Doktora tezi.) Ankara Üniversitesi Sağlık Bilimleri Fakültesi,
- Demir, S., & Sahin, E. K. (2023). An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost. *Neural Computing and Applications*, 35(4), 3173-3190. <https://doi.org/10.1007/s00521-022-07856-4>
- Dogancay, M.M. (2023). *Covid-19 hastaları için Makine öğrenmesi sınıflandırma yöntemleriyle hastalık Evresinin Tahmini* (Doktora tezi) Dokuz Eylül Üniversitesi,
- Dong, Z., Wang, Q., Ke, Y., Zhang, W., Hong, Q., Liu, C. & Chen, X. (2022). Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. *Journal of Translational Medicine*, 20(1): 1-10. <https://doi.org/10.1186/s12967-022-03339->
- Eli-Chukwu, N. C. (2019). Applications of artificial intelligence in agriculture: A review. *Engineering. Technology & Applied Science Research*, 9(4):4377-4383. <https://doi.org/10.48084/etasr.2756>

- Fernández-Delgado, M., Sirsat, M. S., Cernadas, E., Alawadi, S., Barro, S., & Febrero-Bande, M. (2019). An extensive experimental survey of regression methods. *Neural Networks*, **111**:11-34. <https://doi.org/10.1016/j.neunet.2018.12.010>
- Gehrmann, S., Deroncourt, F., LI Y., Carlson E.T., Wu, J.T., Welt, J., Celi, L.A. (2018). Comparing deep learning and concept extraction based methods for patient phenotyping from clinical narratives. *Plos One*, 13(2): e0192360. <https://doi.org/10.1371/journal.pone.0192360>
- Ghassemi, M., Naumann, T., Doshi-Velez, F., Brimmer, N., Joshi, R., Rumshisky, A., & Szolovits, P. (2014, August). Unfolding physiological state: Mortality modelling in intensive care units. *In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 75-84). <https://doi.org/10.1145/2623330.2623742>
- Grobman, W.A., Stamilio, D.M. (2006). Methods of clinical prediction. *American journal of obstetrics and gynecology*, 194(3): 888-894. <https://doi.org/10.1016/j.ajog.2005.09.002>
- Grudniewicz, J., Ślepaczuk, R.. (2023). Application of machine learning in algorithmic investment strategies on global stock markets. *Research in International Business and Finance*, **66**:1020-52. <https://doi.org/10.1016/j.ribaf.2023.102052>
- Gu H.Q and Liu C. (2020). Clinical prediction models: evaluation matters. *Annals of Translational Medicine*, 8(4):72. [doi: 10.21037/atm.2019.11.143](https://doi.org/10.21037/atm.2019.11.143)
- Han, K, Song, K., Choi, B.W. (2016). How to Develop, Validate, and Compare Clinical Prediction Models Involving Radiological Parameters: Study Design and Statistical Methods. *Korean Journal of Radiology*, **17**(3): 339–350. <https://doi.org/10.3348/kjr.2016.17.3.339>
- Harrell, F (2001). Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. *New York Springer*.. <https://doi.org/10.1007/978-1-4757-3462-1>
- Henry, J, Pylypchuk, Y., Searcy, T., Patel, V. (2016). Adoption of electronic health record systems among us non-federal acute care hospitals: 2008–2015. *ONC Data Brief*, **35**: 1–9.
- Henry, K. E., Adams, R., Parent, C., Soleimani, H., Sridharan, A., Johnson, L., & Saria, S. (2022). Factors driving provider adoption of the TREWS machine learning-based early warning system and its effects on sepsis treatment timing. *Nature medicine*, 28(7), 1447-1454. [DOI: 10.1038/s41591-022-01895-z](https://doi.org/10.1038/s41591-022-01895-z)
- Hügler, M., Omoumi, P., Van Laar, J.M., Boedecker, J., Hügler, T. (2020). Applied machine learning and artificial intelligence in rheumatology. *Rheumatology Advances in Practice*, 4(1). <https://doi.org/10.1093/rap/rkaa005>
- Ichord, R.N., Bastian, R., Abraham, L., Askalan, R., Benedict, S., Bernard T.J, Jawad, A.F.(2011). Interrater reliability of the Pediatric National Institutes of Health Stroke Scale (PedNIHSS) in a multicenter study. *Stroke*, 42(3); 613-617. <https://doi.org/10.1161/STROKEAHA.110.607192>
- Islam, M.M., Haque M.R., Iqbal H., Hasan M.M., Hasan M., Kabir, M.N. (2020). Breast cancer prediction: a comparative study using machine learning techniques. *SN Computer Science*, **1**, 1-14. <https://doi.org/10.1007/s42979-020-00305-w>
- Jacobs, M.A., Umbricht, C.B, Parekh V.S, EL Khoul R.H., Cope L., Macura, K. J. & Wolff, A. C. (2020). Integrated multiparametric radiomics and informatics system for characterizing breast tumor characteristics with the OncotypeDX gene assay. *Cancers*, **12**(10): 2772. <http://dx.doi.org/10.3390/cancers12102772>
- Jadhav, S., HE, H., Jenkins, K. (2018). Information gain directed genetic algorithm wrapper feature selection for credit rating. *Applied Soft Computing*, **69**, 541-553. <https://doi.org/10.1016/j.asoc.2018.04.033>
- Janssen, K.J., Vergouwe Y., Kalkman C.J., Grobbee D.E., Moons K.G. (2009). A simple method to adjust clinical prediction models to local circumstances. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 56(3): 194-201. <https://doi.org/10.1007/s12630-009-9041-x>

- Jelovsek, J.E., Chagin K., Gyhagen M., Hagen S., Wilson D., Kattan M.W., Milsom I. (2018). Predicting risk of pelvic floor disorders 12 and 20 years after delivery. *American Journal of Obstetrics and Gynecology*, 218(2): 222-e1. <https://doi.org/10.1016/j.ajog.2017.10.014>
- Johnson, M.L., El-Serag, H.B., Tran T.T., Hartman, C., Richardson P., Abraham N.S. Adapting the Rx-Risk-V for mortality prediction in outpatient populations. *Med Care*, 2006;**44**: 793e7.
- Kapelner, A., Bleich, J. (2013). Bartmachine: Machine learning with bayesian additive regression trees. *Journal of Statistical Software*, **70**(4). <https://doi.org/10.1097/01.mlr.0000218804.41758.ef>.
- Kia, A., Timsina, P., Joshi, H.N., Klang, E., Gupta, R.R., Freeman R.M., Reich D.L., Tomlinson M.S., Dudley J.T., Kohli-Seth R., Mazumdar, M., Levin M.A. (2020). Enhancing the Prediction of Clinical Deterioration in Admitted Patients through a Machine Learning Model. *J Clin Med*. 27; **9**(2): 343. <https://doi.org/10.3390/jcm9020343>
- Kelly, CS. Andersen, CL, Pestian, JP, Wenger, AD, Finch, AB, Strope, GL, Luckstead, EF (2000). Improved outcomes for hospitalized asthmatic children using a clinical pathway. *Annals of Allergy, Asthma & Immunology*, 84(5): 509-516. doi: [10.1016/S1081-1206\(10\)62514-8](https://doi.org/10.1016/S1081-1206(10)62514-8).
- Knaus, WA, Draper, EA., Wagner DP, Zimmerman JE (1985). Apache II: a severity of disease classification system. *Critical Care Medicine*, **13**(10): 818-829. <http://dx.doi.org/10.1097/00003246-198510000-00009>
- Lankham, I., & Slaughter, M. (2020). Simple and efficient bootstrap validation of predictive models using SAS/STAT® Software. In of the SAS Global Forum (pp. 4647-2020)
- Le Gall, J. R., Lemeshow, S., & Saulnier, F. (1993). A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *Jama*, 270(24), 2957-2963. [doi:10.1001/jama.1993.03510240069035](https://doi.org/10.1001/jama.1993.03510240069035)
- Lee, B.,K. Lessler, J., Stuart, E.,A. (2010). Improving propensity score weighting using machine learning, *Stat Med*, 29(3): 337–346. <https://doi.org/10.1002/sim.3782>
- Lee, C., Lee, G.G. (2006). Information gain and divergence-based feature selection for machine learning-based text categorization. *Information Processing & Management*, 42(1): 155-165. <https://doi.org/10.1016/j.ipm.2004.08.006>
- Lee, Y.H., Bang H., Kim D.J. (2016). How to establish clinical prediction models. *Endocrinology and Metabolism*, 31(1): 38-44. <https://doi.org/10.3803/EnM.2016.31.1.38>
- Lynch, S.M., Handorf, E., Sorice, K.A., Blackman E., BealiN L., Giri V.N., Daly M. (2020). The effect of neighborhood social environment on prostate cancer development in black and white men at high risk for prostate cancer. *PloS one*, 15(8): e0237332. <https://doi.org/10.1371/journal.pone.0237332>
- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)* 9(1): 381-386. [Doi:10.21275/ART20203995](https://doi.org/10.21275/ART20203995)
- Maltarollo, V.G., Kronenberger, T., Espinoza G., Oliveira P.R., Honorio K.M. (2019). Advances with support vector machines for novel drug discovery. *Expert opinion on drug discovery*, 14(1), 23-33. [doi: 10.1080/17460441.2019.1549033](https://doi.org/10.1080/17460441.2019.1549033).
- Mehta, H.B., Mehta V, Girman C.J, Adhikari D, Johnson M.L (2016). Regression coefficient-based scoring system should be used to assign weights to the risk index. *Journal of Clinical Epidemiology*, 79: 22-28. <https://doi.org/10.1016/j.jclinepi.2016.03.031>
- Mitchell, R, Adinets, A, Rao, T, Frank, E (2018). Xgboost: Scalable GPU accelerated learning. *Computer Science*. 1806.11248. <https://doi.org/10.48550/arXiv.1806.11248>
- Moons, K.G., Kengne, A.P., Woodward, M., Royston, P., Vergouwe, Y., Altman, D.G., Grobbee, D.E. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012; **98**: 683-90. [doi: 10.1136/heartjnl-2011-301246](https://doi.org/10.1136/heartjnl-2011-301246)

- Moons, K.G., Altman, D.G., Reitsma, J.B., Ioannidis, J.P., Macaskill, P., Steyerberg, E.W., Collins, G.S. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine*, 162(1): W1-W73. <https://doi.org/10.1186/s12916-014-0241-z>
- Morita, T., Tsunoda, J., Inoue, S., Chihara, S. (1999). The Palliative Prognostic Index: a scoring system for survival prediction of terminally ill cancer patients. *Supportive care in cancer*, 7: 128-133. <https://doi.org/10.1007/s005200050242>
- Muthukrishnan, R. and Rohini, R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning. In *2016 IEEE international conference on advances in computer applications (ICACA)*, 18-20. [doi: 10.1109/ICACA.2016.7887916](https://doi.org/10.1109/ICACA.2016.7887916)
- Mythili, T., Mukherji, D., Padalia, N., & Naidu, A. (2013). A heart disease prediction model using SVM-decision trees-logistic regression (SDL). *International Journal of Computer Applications*, 68(16). [doi: 10.5120/11662-7250](https://doi.org/10.5120/11662-7250)
- Naji, M. A., El Filali, S., Aarika, K., Benlahmar, E. H., Abdelouahid, R. A., & Debauche, O. (2021). Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science*, 191: 487-492. <https://doi.org/10.1155/2017/3781951>
- Norgeot, B., Glicksberg, B.S., Trupin L., Lituiev D., Gianfrancesco M., Oskotsky B., Butte A.J. (2019). Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. *JAMA Network Open*, 2(3): e190606-e190606. [doi: 10.1001/jamanetworkopen.2019.0606](https://doi.org/10.1001/jamanetworkopen.2019.0606)
- Omuya E.O., Okeyo G.O., Kimwele MW (2021). On the feature selection for classification using principal component analysis and information gain. *Expert Systems with Applications*, 174: 114765. <https://doi.org/10.1155/2018/1407817>
- Oshiro T.M., Perez P.S., Baranauskas J.A. (2012). How many trees in a random forest? In *Machine Learning and Data Mining in Pattern Recognition: 8th International Conference Springer Berlin Heidelberg*. 8 :54-16. [doi: 10.1007/978-3-642-31537-4_13](https://doi.org/10.1007/978-3-642-31537-4_13)
- Özgür E.G., Doğanay Erdoğan, B. (2022). Comparison of Prediction Performance of Computer Adaptive Testing with Machine Learning Methods in Response Patterns Suitable for Rasch Model with Simulation Study: A Methodological Research. *Türkiye Klinikleri Journal of Biostatistics*, 14(2). ([doi: 10.5336/biostatic.2022-88297](https://doi.org/10.5336/biostatic.2022-88297))
- Pang, Y., Judd, N., O'Brien J., Ben-Avie M. (2017). Predicting students' graduation outcomes through support vector machines. In *2017 IEEE Frontiers in Education Conference (FIE)*. IEEE. 1-8. [DOI: 10.1109/FIE.2017.8190666](https://doi.org/10.1109/FIE.2017.8190666)
- Pimentel, M.A., Redfern, O.C., Malycha, J., Meredith, P., Prytherch, D., Briggs, J., Watkinson, PJ (2021). Detecting deteriorating patients in hospital: development and validation of a novel scoring system. *American Journal of Respiratory and Critical Care Medicine*, 204(1): 44–52. [doi: 10.1164/rccm.202007-2700OC](https://doi.org/10.1164/rccm.202007-2700OC)
- Probst, P., Wright, M.N., Boulesteix, A.L. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3): e1301. <https://doi.org/10.1002/widm.1301>
- Qi Y. (2012). Random forest for bioinformatics. *Ensemble machine learning: Methods and applications*, 307-323. [http://dx.doi.org/10.1007/978-1-4419-9326-7_11](https://doi.org/10.1007/978-1-4419-9326-7_11)
- Raita, Y., Goto, T., Faridi, M. K., Brown, D. F., Camargo, C. A., & Hasegawa, K. (2019). Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care*, 23: 64. [DOI: 10.1186/s13054-019-2351-7](https://doi.org/10.1186/s13054-019-2351-7)
- Ramesh, T.R., Lilhore, UK, Poongodi, M, Simaiya, S., Kaur A., Hamdim, M. (2022). Predictive analysis of heart diseases with machine learning approaches. *Malaysian Journal of Computer Science*, 132-148. <https://doi.org/10.22452/mjcs.sp2022no1.10>

- Richter, A.N., Khoshgoftaar, T.M. (2018). A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artificial Intelligence in Medicine*, **90**: 1–14. [doi: 10.1016/j.artmed.2018.06.002](https://doi.org/10.1016/j.artmed.2018.06.002)
- Rigatti, S.J. (2017). Random forest. *Journal of Insurance Medicine*, 47(1): 31-39. [doi: 10.17849/insm-47-01-31-39.1](https://doi.org/10.17849/insm-47-01-31-39.1).
- Rudin C. (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*, **1**: 206–215. [doi: 10.1038/s42256-019-0048-x](https://doi.org/10.1038/s42256-019-0048-x)
- Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. J. O. G. R. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, **7**: 804-818. <https://doi.org/10.1016/j.oregeorev.2015.01.001>
- Ronowicz, J, Thommes, M., Kleinebudde, P., Krysiński, J. (2015). A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm. *European Journal of Pharmaceutical Sciences*, **73**: 44-48. [doi: 10.1016/j.ejps.2015.03.013](https://doi.org/10.1016/j.ejps.2015.03.013).
- Royston, P, Moons, K.G., Altman D.G. (2009). Prognosis and prognostic research: Developing a prognostic model. *BMJ (Clinical research ed.)*, **338**: b604. <https://doi.org/10.1136/bmj.b604>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. **1**: 206–215. [DOI: 10.1038/s42256-019-0048-x](https://doi.org/10.1038/s42256-019-0048-x)
- Rufo, D.D., Debelee, TG, Ibenthal A, Negera WG (2021). Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM). *Diagnostics*, **11**(9): 1714. <https://doi.org/10.3390/diagnostics11091714>
- Sanchez-Pinto, L.N., Bennett, T.D. (2022). Evaluation of Machine Learning Models for Clinical Prediction Problems. *Pediatric Critical Care Medicine* **23**(5): 405-408. [doi: 10.1097/PCC.0000000000002942](https://doi.org/10.1097/PCC.0000000000002942).
- Samuel, M., (2002). Pediatric appendicitis score. *Journal Of Pediatric Surgery*, **37**(6): 877-881. [doi: 10.1053/jpsu.2002.32893](https://doi.org/10.1053/jpsu.2002.32893).
- Schmidt, M. (2005). Least squares optimization with L1-norm regularization. *CS542B Project Report*, **504**: 195-221.
- Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, **20**(1), 3-29. <https://doi.org/10.1177/1536867X20909688>
- Segar, M.W., Vaduganathan, M., Patel, K.V., Mcguire, D.K., Butler, J., Fonarow, G. C., & PANDEY A. (2019). Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: the WATCH-DM risk score. *Diabetes Care*, **42**(12): 2298-2306. [doi: 10.2337/dc19-0587](https://doi.org/10.2337/dc19-0587).
- Sekulić, A., Kilibarda, M., Heuvelink, G.B., Nikolić, M., Bajat, B. (2020). Random forest spatial interpolation. *Remote Sensing*, **12**(10): 1687. <https://doi.org/10.3390/rs12101687>
- Setoguchi, S., Schneeweiss, S., Brookhart M.A., Glynn R.J., and Cook E.F. (2008). Evaluating uses of data mining techniques in propensity score estimation: a simulation study. *Pharmacoepidemiology and Drug Safety*, **17**(6): 546–55. [doi: 10.1002/pds.1555](https://doi.org/10.1002/pds.1555).
- Shariat, S.F., Capitanio, U., Jeldres, C. and Karakiewicz, P.I. (2009). Can nomograms be superior to other prediction tools? *BJU International*, **103**(4): 492-497. [doi: 10.1111/j.1464-410X.2008.08073.x](https://doi.org/10.1111/j.1464-410X.2008.08073.x).
- Shipe, M.E., Deppen, SA., Farjah, F., Grogan, E.L. (2019). Developing prediction models for clinical use using logistic regression: an overview. *Journal Of Thoracic Disease*, **11**(4): S574. [doi: 10.21037/jtd.2019.01.25](https://doi.org/10.21037/jtd.2019.01.25).
- Shouval, R., Bondi, O., Mishan, H., Shimoni, A., Unger, R., Nagler, A. (2014). Application of machine learning algorithms for clinical predictive modeling: a data-mining approach in SCT. *Bone Marrow Transplantation*, **49**(3), 332-337. [doi: 10.1038/bmt.2013.146](https://doi.org/10.1038/bmt.2013.146).

- Shouval, R., Fein, JA, Savani, B., Mohty, M., Nagler, A. (2021). Machine learning and artificial intelligence in haematology. *British Journal of Haematology*, 192(2), 239-250. [doi: 10.1111/bjh.16915](https://doi.org/10.1111/bjh.16915).
- Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12–18.
- Steyerberg, E.W., Vergouwe, Y. (2014). Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European Heart Journal*, 35(29): 1925–1931. [doi: 10.11613/BM.2014.003](https://doi.org/10.11613/BM.2014.003)
- Steyerberg, E.W. (2010). *Clinical Prediction Models 1st ed* . New York City, NY: Springer-Verlag New York. DOI: [10.1007/978-0-387-77244-8](https://doi.org/10.1007/978-0-387-77244-8)
- Steyerberg, E.W., Steyerberg, E.W. (2009). Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating, 313-331. DOI: [10.1007/978-0-387-77244-8](https://doi.org/10.1007/978-0-387-77244-8)
- Steyerberg, E. W., Harrell Jr, F. E., Borsboom, G. J., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*, 54(8), 774-781. [doi: 10.1016/s0895-4356\(01\)00341-9](https://doi.org/10.1016/s0895-4356(01)00341-9).
- Subramanian V., Mascha E.J., Kattan M.W. (2021). Developing a Clinical Prediction Score: Comparing Prediction Accuracy of Integer Scores to Statistical Regression Models. *Anesthesia & Analgesia*, 132(6): 1603-1613. [doi: 10.1213/ANE.0000000000005362](https://doi.org/10.1213/ANE.0000000000005362).
- Sullivan, L.M., Massaro JM, D'agostin on RB SR (2004). Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in Medicine*, 23(10): 1631–1660. [doi: 10.1002/sim.1742](https://doi.org/10.1002/sim.1742).
- Şahin, E.K. (2022). Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto International*, 37(9), 2441-2465. DOI:[10.1080/10106049.2020.1831623](https://doi.org/10.1080/10106049.2020.1831623)
- Tohka, J., & Van, Gils, M. (2021). Evaluation of machine learning algorithms for health and wellness applications: A tutorial. *Computers in Biology and Medicine*, 132: 104324. [doi: 10.1016/j.combiomed.2021.104324](https://doi.org/10.1016/j.combiomed.2021.104324).
- Touzani, S., Granderson, J., Fernandes, S. (2018). Gradient boosting machine for modeling the energy consumption of commercial buildings. *Energy and Buildings*, 158: 1533-1543. <https://doi.org/10.1016/j.enbuild.2017.11.039>
- Trevethan, R. (2017). Sensitivity, specificity, and predictive values: foundations, pliabilitys, and pitfalls in research and practice. *Frontiers in public health*, 5: 307. <https://doi.org/10.3389/fpubh.2017.00307>
- Uanhoru, JO, Wang, Y, O'connell AA. (2021). Problems with using odds ratios as effect sizes in binary logistic regression and alternative approaches. *The Journal of Experimental Education*, 89(4): 670-689. <https://doi.org/10.1080/00220973.2019.1693328>
- Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making*, 19(1), 1-16. Valero III, V., Grimm, J. C., Kilic, A., Lewis, R. L., Tosoian, J. J., He, J., ... & Study, T. P. R. A. G. Novel risk scoring system reliably predicts readmission after pancreatectomy. *Journal of the American College of Surgeons*, 220(4): 701-713. [doi: 10.1186/s12911-019-1004-8](https://doi.org/10.1186/s12911-019-1004-8).
- Vidaurre, D., Bielza C., Larranaga P. (2013). A survey of L1 regression. *International Statistical Review*, 81(3): 361-387. DOI: [10.1111/insr.12023](https://doi.org/10.1111/insr.12023)
- Visweswaran S, Angus DC, Hsieh M, Weissfeld L, Yealy D. VE Cooper GF (2010). Learning patient-specific predictive models from clinical data. *Journal of Biomedical Informatics*, 43(5): 669-685. DOI: [10.1016/j.jbi.2010.04.009](https://doi.org/10.1016/j.jbi.2010.04.009)
- Yang, Q., Li, J., Zhang, Z., Wu, X., Liao, T., YU, S., ... & Sun, J. (2021). Clinical characteristics and a decision tree model to predict death outcome in severe COVID-19 patients. *BMC Infectious Diseases*, 21(1), 1-9. [doi: 10.1186/s12879-021-06478-w](https://doi.org/10.1186/s12879-021-06478-w).

- Wang YW, Ni YQ, Zhang QH, Zhang C. (2021). Bayesian approaches for evaluating wind-resistant performance of long-span bridges using structural health monitoring data. *Structural Control and Health Monitoring*, 28(4): e2699. <https://doi.org/10.1002/stc.2699>
- Win TZ., Kham NS M (2019). Information gain measured feature selection to reduce high dimensional data ([Doctoral](#) dissertation).
- Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Dphil SM. (2019). PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170 (1): 51-58. [doi: 10.7326/M18-1376](https://doi.org/10.7326/M18-1376).
- Zabor EC, ReddY CA, Tendulkar RD, Patil, S. (2022). Logistic regression in clinical studies. *International Journal of Radiation Oncology* Biology* Physics*, **112**(2):271-277. [doi: 10.1016/j.ijrobp.2021.08.007](https://doi.org/10.1016/j.ijrobp.2021.08.007).
- Zhang L, Fabbri D., Upender R, Kent D. (2019). Automated sleep stage scoring of the sleep heart health study using deep neural networks. *Sleep*, **42**(11): 159. [doi: 10.1093/sleep/zsz159](https://doi.org/10.1093/sleep/zsz159).
- Zhang Z, Zhang Khanal, MK. (2017). Development of scoring system for risk stratification in clinical medicine: a step-by-step tutorial. *Annals of Translational Medicine*, **5**(21). doi: 10.21037/atm.2017.08.22. [DOI:10.21037/atm.2017.08.22](https://doi.org/10.21037/atm.2017.08.22)
- Zhao A, Larbi M, Miller K, O'Neill S, Jayasekera J. A scoping Review of interactive and personalized web-based clinical tools to support treatment decision making in breast cancer. *The Breast*, **61**: 43-57. [DOI: 10.1016/j.breast.2021.12.003](https://doi.org/10.1016/j.breast.2021.12.003)
- Zhou H. Zhang Zhou Y, Guo X, Ma Y (2021). A feature selection algorithm of decision tree based on feature weight. *Expert Systems with Applications*, **164**: 113842. [DOI:10.1016/j.eswa.2020.113842](https://doi.org/10.1016/j.eswa.2020.113842)

EKLER

EK-1. Lojistik regresyona uygun benzetim verisi türetimi için örnek R kodları

A senaryosu, bağımlı değişken 0-1 oranı; %50-%50 ve bağımsız değişkendeki 1'lerin oranı 0,1-0,33; 0,33-0,66 ve 0,66-0,9 aralıklarının herbirinde 3 değişken olmak üzere rastgele seçilmiştir

```
F.generate <- function(n) {  
  
  b0 <- -3.4  
  b1 <- 0.9  
  b2 <- 0.6  
  b3 <- 0.6  
  b4 <- 0.7  
  b5 <- 0.8  
  b6 <- 0.7  
  b7 <- 0.7  
  b8 <- 0.9  
  b9 <- 0.8  
  
  w1 <- rbinom(n, 1, runif(1, 0.1, 0.33))  
  w2 <- rbinom(n, 1, runif(1, 0.1, 0.33))  
  w3 <- rbinom(n, 1, runif(1, 0.1, 0.33))  
  w4 <- rbinom(n, 1, runif(1, 0.34, 0.66))  
  w5 <- rbinom(n, 1, runif(1, 0.34, 0.66))  
  w6 <- rbinom(n, 1, runif(1, 0.34, 0.66))  
  w7 <- rbinom(n, 1, runif(1, 0.67, 0.9))  
  w8 <- rbinom(n, 1, runif(1, 0.67, 0.9))  
  w9 <- rbinom(n, 1, runif(1, 0.67, 0.9))  
  
  logit <- b0 + b1 * w1 + b2 * w2 + b3 * w3 +  
    b4 * w4 + b5 * w5 + b6 * w6 +  
    b7 * w7 + b8 * w8 + b9 * w9  
  
  olasiliklar <- 1 / (1 + exp(-logit))  
  
  z.a <- rbinom(n, 1, olasiliklar)  
  
  sim <- as.data.frame(cbind(w1, w2, w3, w4, w5, w6, w7, w8, w9, z.a))  
  return(sim)  
}  
  
simdata <- F.generate(500)  
  
simdata
```

LR, ML ve diğer algoritmalara ilişkin örnek R kodları

LR

```
lr <- glm(z.a ~ w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8 + w9, data = simdata, family='binomial')
```

ML

Rastgele orman

```
rf_model <- randomForest(x, as.factor(y), ntree = 5, importance=TRUE)
```

Diğer Algoritma

Bayesci GDM

```
model <- bayesglm(formula = z.a ~ w1 + w2 + w3 + w4 + w5 + w6 + w7 + w8 + w9, data = simdata,  
family = binomial())
```

LR, ML ve diğer algoritmaların toplam skor hesaplamalarına ilişkin örnek R kodları

LR

Beta katsayısı için;

```
skor <- coef(lr)
```

ML

Rastgele orman

```
var_importance <- importance(rf_model)  
var_names <- rownames(var_importance)  
importance_values <- var_importance[, 1]
```

```
skor = importance_values
```

Diğer Algoritma

Bayesci GDM için;

```
skor = coef(model)
```

Toplam skor

```
x <- simdata[,][1:9]  
y <- data.matrix(x)  
Vector1 <- c(skor[2:10])  
z <- as.matrix(Vector1)
```

```
tskor <- y %*% z
```

LR, ML ve diğer algoritmaların performans ölçütlerinin hesaplamalarına ilişkin örnek R kodları

```
veriROC <- data.frame(simdata, skor)
rocobj <- roc(veriROC$z.a, veriROC$skor)
tskor <- ifelse(skor > coords(rocobj, "best", best.method = "closest.topleft")[1,1], 1, 0)
g <- simdata$z.a
# Confusion matrix
conf_matrix <- confusionMatrix(table(tskor, g), positive = "1")

# True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN)
TP <- sum(g == 1 & tskor == 1)
TN <- sum(g == 0 & tskor == 0)
FP <- sum(g == 0 & tskor == 1)
FN <- sum(g == 1 & tskor == 0)

# Sensitivity (Recall)
sensitivity <- TP / (TP + FN)
# Specificity
specificity <- TN / (TN + FP)
# Positive Predictive Value (Precision)
PPV <- TP / (TP + FP)
# Negative Predictive Value
NPV <- TN / (TN + FN)
# Accuracy
accuracy <- conf_matrix$overall["Accuracy"]
# F1 Score
recall <- TP / (TP + FN)
precision <- TP / (TP + FP)
f1_score <- 2 * ((precision * recall) / (precision + recall))

# Standard error for each metric
se_sensitivity <- sqrt(sensitivity * (1 - sensitivity) / (TP + FN))
se_specificity <- sqrt(specificity * (1 - specificity) / (TN + FP))
se_PPV <- sqrt(PPV * (1 - PPV) / (TP + FP))
se_NPV <- sqrt(NPV * (1 - NPV) / (TN + FN))
se_accuracy <- sqrt(accuracy * (1 - accuracy) / (TP + TN + FP + FN))
se_f1_score <- sqrt(f1_score * (1 - f1_score) / (TP + FP))

# Confidence intervals
ci_sensitivity <- c(sensitivity - 1.96 * se_sensitivity, sensitivity + 1.96 * se_sensitivity)
ci_specificity <- c(specificity - 1.96 * se_specificity, specificity + 1.96 * se_specificity)
ci_PPV <- c(PPV - 1.96 * se_PPV, PPV + 1.96 * se_PPV)
ci_NPV <- c(NPV - 1.96 * se_NPV, NPV + 1.96 * se_NPV)
ci_accuracy <- c(accuracy - 1.96 * se_accuracy, accuracy + 1.96 * se_accuracy)
ci_f1_score <- c(f1_score - 1.96 * se_f1_score, f1_score + 1.96 * se_f1_score)

result <- c(ci(rocobj)[2:2], ci(rocobj)[1:1], ci(rocobj)[3:3], sensitivity, specificity, PPV, NPV,
accuracy, f1_score, ci_sensitivity[1], ci_sensitivity[2], ci_specificity[1], ci_specificity[2], ci_PPV[1],
ci_PPV[2], ci_NPV[1], ci_NPV[2], ci_accuracy[1], ci_accuracy[2], ci_f1_score[1], ci_f1_score[2])
```

EK 2. Etik Kurur Onayı

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU					
ARAŞTIRMANIN AÇIK ADI		Çocuklarda Malnutrisiyonun Tanısında Yeni Bir Klinik Tahmin Modeli			
VARSA ARAŞTIRMANIN PROTOKOL KODU					
ETİK KURUL BİLGİLERİ	ETİK KURULUN ADI	Hitit Üniversitesi Tıp Fakültesi Klinik Araştırmalar Etik Kurulu			
	AÇIK ADRESİ	Hitit Üniversitesi Tıp Fakültesi Dekanlığı Fen Edebiyat Yerleşkesi Ulukavak Mahallesi Çiftlik Çayırı Cad. No:45 Kat:1 19040 Merkez Çorum			
	TELEFON				
	FAKS				
	E-POSTA				
BAŞVURU BİLGİLERİ	KOORDİNATOR SORUMLU ARAŞTIRMACI UNVANI ADI SOYADI	Doç. Dr. Atakan COMBA			
	KOORDİNATOR SORUMLU ARAŞTIRMACININ UZMANLIK ALANI	Çocuk Sağlığı ve Hastalıkları			
	KOORDİNATOR SORUMLU ARAŞTIRMACININ BULUNDUĞU MERKEZ	Hitit Üniversitesi Tıp Fakültesi			
	VARSA İDARI SORUMLU UNVANI ADI SOYADI				
	DESTEKLEYİCİ				
	PROJE YÜRÜTÜCÜSÜ UNVANI ADI SOYADI (TUBİTAK vb. grb. kaynaklardan destek alanlar için)				
	DESTEKLEYİCİNİN YASAL TEMSİLCİSİ				
	ARAŞTIRMANIN FAZİ VE LERİ	FAZ 1	☐		
		FAZ 2	☐		
		FAZ 3	☐		
FAZ 4		☐			
Gözlemsel ilaç çalışması		☐			
Tıbbi cihaz klinik araştırması		☐			
In vitro tıbbi tanı cihazları ile yapılan performans değerlendirme çalışmaları		☐			
İlaç dışı klinik araştırma	☒				
Diğer ise belirtiniz					
ARAŞTIRMAYA KATILAN MERKEZLER	TEK MERKEZ ☒	ÇOK MERKEZLİ ☐	ULUSAL ☒	ULUSLARARASI ☐	

Etik Kurul Başkanı
Unvanı/Adı/Soyadı: Doç. Dr. İbrahim Tayfun ŞAHİNER
İmza:

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI	Çocuklarda Malnutrisyonun Tanısında Yeni Bir Klinik Tahmin Modeli
YALSA ARAŞTIRMANIN PROTOKOL KODU	

DİĞERLENDİRİLEN BELGELER	Belge Adı	Tarihi	Versiyon Numarası	Dili
	ARAŞTIRMA PROTOKOLÜ	16.08.2022	2	Türkçe ☒ İngilizce ☐ Diğer ☐
	BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU			Türkçe ☒ İngilizce ☐ Diğer ☐
	OLGU RAPOR FORMU			Türkçe ☒ İngilizce ☐ Diğer ☐
	ARAŞTIRMA BROŞÜRÜ			Türkçe ☐ İngilizce ☐ Diğer ☐
DİĞERLENDİRİLEN DİĞER BELGELER	Belge Adı	Açıklama		
	SİGORTA			
	ARAŞTIRMA BÜTÇESİ	☐		
	BIYOLOJİK MATERYEL TRANSFER FORMU	☐		
	ILAN	☐		
	YILLIK BİLDİRİM	☐		
	SONUÇ RAPORU	☐		
	GUVENLİLİK BİLDİRİMLERİ	☐		
KARAR BELGELERİ	Karar No:2022-64	Tarih: 25/08/2022		
	Yukarıda bilgileri verilen başvuru dosyası ile ilgili belgeler araştırmanın/çalışmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş ve uygun bulunmuş olup araştırmanın/çalışmanın başvuru dosyasında belirtilen merkezlerde gerçekleştirilmesinde etik ve bilimsel sakınca bulunmadığına toplanmaya katılan etik kurul üye tam sayısının salt çoğunluğu ile karar verilmiştir. İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik kapsamında yer alan araştırmalar/çalışmalar için Türkiye İlaç ve Tıbbi Cihaz Kurumundan izin alınması gerekmektedir.			

KLİNİK ARAŞTIRMALAR ETİK KURULU	
ETİK KURULUN ÇALIŞMA ESASI	İlaç ve Biyolojik Ürünlerin Klinik Araştırmaları Hakkında Yönetmelik, İyi Klinik Uygulamaları Kılavuzu
BASKAN VEKİLİNİN UNVANI/ADI /SOYADI:	Doç. Dr. İbrahim Tayfun ŞAHİNER

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet	Araştırma ile ilişkisi	Katılım *	İmza
Doç. Dr. İbrahim Tayfun ŞAHİNER	GENEL CERRAHİ	Hittit Ü.T.F.	F ☒ K ☐	F ☐ H ☒	E ☒ H ☐	
Doç. Dr. Güven GÜNEY	TIBBİ PATOLOJİ	Hittit Ü.T.F.	F ☒ K ☐	F ☐ H ☒	E ☒ H ☐	
Prof. Dr. Yusuf KARAVELİOĞLU	KARDİYOLOJİ	Hittit Ü.T.F.	F ☒ K ☐	F ☐ H ☒	E ☒ H ☐	
Doç. Dr. Havva Nur PELİT KENDİRCİ	ÇOCUK SAĞLIĞI VE HASTALIKLARI	Hittit Ü.T.F.	F ☐ K ☒	F ☐ H ☒	E ☒ H ☐	
Doç. Dr. Coşkun ÖZTEKİN	AİLE HEKİMLİĞİ	Hittit Ü.T.F.	F ☒ K ☐	F ☐ H ☒	E ☒ H ☐	
Doç. Dr. Ayta ÇAĞLIYAN TÜRK	FİZİKSEL TIP VE REHABİLİTASYON	Hittit Ü.T.F.	F ☐ K ☒	F ☐ H ☐	E ☐ H ☐	
Dr. Öğr. Üyesi Ece Yazla ASATÖV	RUM SAĞLIĞI VE HASTALIKLARI	Hittit Ü.T.F.	F ☐ K ☒	F ☐ H ☒	E ☒ H ☐	
Dr. Öğr. Üyesi Aynur KOÇ	FİZYOLOJİ	Hittit Ü.T.F.	F ☐ K ☒	F ☐ H ☐	E ☐ H ☐	izinli
Dr. Öğr. Üyesi Nihan KÜÇÜK	TIBBİ FARMAKOLOJİ	Hittit Ü.T.F.	F ☐ K ☒	F ☐ H ☐	E ☐ H ☐	izinli
Uzm. Dr. Emin RENÇBER	HALK SAĞLIĞI	Çorum Halk Sağlığı Merkezi	F ☒ K ☐	F ☐ H ☒	E ☒ H ☐	

Etik Kurul Başkanı
Unvanı Adı/Soyadı: Doç. Dr. İbrahim Tayfun ŞAHİNER
İmza:

Not: Etik Kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

KLİNİK ARAŞTIRMALAR ETİK KURULU KARAR FORMU

ARAŞTIRMANIN AÇIK ADI		Çocuklarda Malnutrisyonun Tanısında Yeni Bir Klinik Tahmin Modeli						
YARSA ARAŞTIRMANIN PROTOKOL KODU								
Hukuk, Mıv. Avukat Gülhan DURMUSBAS Biyomedikal Müh. Anıl KARATAS Sad DEMİR	HUKUK	Hitt Üniversitesi Rektörlüğü	☐	☒	☐	☒	☒	☐
	BIYOMEDİKAL	S.B. Çocuk Eğitim Araştırma Hastanesi	☒	☐	☐	☒	☒	☐
	Y. Drama Bln Uzmanı		☒	☐	☐	☐	☐	☐

* Toplantıda Bulunma

Etik Kurul Başkanı
Unvanı/Adı/Soyadı: Doç. Dr. İbrahim Tayfun ŞAHİNER
İmza:

Not: Etik kurul başkanı, imzasının yer almadığı her sayfaya imza atmalıdır.

ÖZGEÇMİŞ

