T.C.
İSTANBUL ÜNİVERSİTESİ-CERRAHPAŞA
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DOKTORA TEZİ

TÜP BEBEK TEDAVİSİNDE EMBRİYO SEÇİMİNİN MODELLENMESİ

PAKİZE YİĞİT

DANIŞMAN
PROF.DR.ABDULBARİ BENER

BİYOİSTATİSTİK ANABİLİM DALI
BİYOİSTATİSTİK VE TIP BİLİŞİMİ DOKTORA PROGRAMI

İSTANBUL-2021

ii

İÇİNDEKİLER

TABLOLAR LİSTESİ

ŞEKİLLER LİSTESİ

SEMBOLLER / KISALTMALAR LİSTESİ

ART: Yardımcı Üreme Teknikleri

AUROC: Roc eğrisinin altında kalan alan

COS: Kontrollü Overyan Hiperstimülasyon

DN: Doğru Negatif

DP: Doğru Pozitif

DVM: Destek Vektör Makineleri

Gbm: Gradyan Arttırma Makineları

IVF: İn vitro fertilizasyon (In Vitro Fertilization)

KA: Karar Ağacı

LR: Lojistik Regresyon

NB: Naive Bayes

RO: Rastgele Orman

ROS: Rastgele Yukarı Örnekleme

ROSE: Random Over-Sampling

RUS: Rasgele Az Örnekleme

SL: Super Learner

SMOTE: Sentetik Azınlık Asırı Örnekleme

VTKB:Veri tabanlarından bilgi keşfetme süreci

YN: Yanlış Negatif

YP: Yanlış Pozitif

YSA: Yapay Sinir Ağları

ÖZET

Yiğit, P. (2021). Tüp Bebek Tedavisinde Embriyo Seçiminin Moellenmesi. İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü, Biyoistatistik ABD. Doktora Tezi. İstanbul.

Tüp bebek tedavisinde başarı günden güne artış gösterse de henüz istenen başarı düzeyi elde edilememiştir. Bu nedenle, bu nedenle tedaviyi uygulayan klinisyenler ve hastalar tıbbi, finansal ve sosyal kayıplar yaşamaktadır. Bu çalışmanın amacı, IVF tedavisinde embriyonun tutunma başarısını çiflerin demografik özellikleri, embriyonun kalitesi ve IVF sürecine özgü COS değişkenleri ile makine öğrenmesi yöntemleri ile tahmin etmektir. Bunun yanında IVF başarısını etkileyen değişkenlerin bulunması ve önemine göre sıralanması da çalışmanın alt amacıdır. Araştırma, bir retrospektif kohort çalışmasıdır. Çalışmada kullanılan veri seti tek bir merkezden alınmış 939 embriyodan oluşmaktadır. Embriyonun tutunma başarısı modellendiği için model bir sınıflandırma problemi göstermektedir. Bu amaçla, Lojistik Regresyon, Naive Bayes, Rastgele Orman, Destek Vektör Makineları, Yapay Sinir ağları, Gradyan Arttırma Makineları,Bagging, Xgboosting ve Super Learner algoritmaları kullanılmıştır. Modellerin başarısı performans ölçütleri ile karşılaştırılarak en iyi tahmini yapan algoritmanın bulunması ve embriyo tutunma başarısını etkileyen değişkenlerin önem dereceleri bulunmuştur. Sonuç olarak, var olan algoritmaları bir arada kullanarak daha üstün sonuçlar veren Super Learner algoritması % 88 AUROC ile en başarılı sonucu vermiştir. IVF başarısında, embriyonu modellenmesinde en önemli değişken annenin yaşı olarak belirlenmiştir. Çalışma, IVF klinisyenlerine ve hastalara, makine öğrenmesi temelli bir klinik karar destek sistemi önermektedir.

Anahtar Kelimeler: Makine Öğrenmesi Yöntemleri, Sınıflandırma Teknikleri, Super Learner, Tüp Bebek Tedavisi,

ABSTRACT

Yiğit, P. (2021). Modeling Embriyo Selection In Vitro Fertilization Treatment . İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü, Biyoistatistik ABD. Doktora Tezi. İstanbul.

Although success in in vitro fertilization (IVF) treatment has increased recently, it has not yet achieved the desired level of success. For this reason, clinicians and patients have difficulties with medical, social and also financial. The aim of this study is to predict using machine learning methods to select the most accurate embryo for IVF success. In addition, finding the variables that affect IVF success the most is the second objective of the study. The research is a retrospective cohort study. The data set used in the study consists of 939 embryos taken from a single center. The aim is a classification problem as it is modeled whether the embryo is successful (1) or not (0). Logistic Regression, Naïve Bayes, Random Forest, Support Machine, Bagging, Gradian Boosting Trees, Xgboosting ve Super Learner methods were used to model embryo selection. As a result, the Super Learner provided the best prediction with %88 AUROC. The most important variable in the success of IVF in embryo modeling was determined as the age of the mother. The study proposes a machine learning-based clinical decision support system for IVF clinicians and patients.

Keywords: Classification, In Vitro Fertilization, Machine Learning Methods, Super Learner

1. GİRİŞ VE AMAÇ

İnfertilite, çiflerin en az 1 yıl korunmasız birleşmeden sonra gebe kalamaması olarak tanımlanmaktadır (Zegers-hochschild et al., 2006). Yardımla üreme tekniklerinden (Assisted Reproductive Technology-ART) biri olan In Vitro Fertilizasyon (IVF), (tüp bebek tedavisi), 1978 yılında ilk bebeğin bu yöntemle dünyaya gelmesinden beri en çok kullanılan infertilite tedavisidir. Dünyada yaklaşık çiftlerin % 15'inin bu sorunla karşı karşıya olduğu ve oranın hızla arttığı bilinmektedir (Sharlip et al., 2002). Yöntem, in vitro koşullarda (in vitro fertilization, IVF) sperm ve yumurtanın birleştirilip embriyonun gelişmesi ve hamileliğin gerçekleşmesi için anneye transfer edilmesi için geliştirilen tedavi yöntemine verilen genel isimdir. Her bir ART işlemi sonucunda canlı doğum olasılığı % 29.1 (McLernon, Steyerberg, Te Velde, Lee, & Bhattacharya, 2016), altıncı denemeden sonra ise bu olasılık en çok % 38 ile % 49 arasında değişebilmektedir (Dhillon et al., 2016). Sonuç olarak tedavinin % 100 başarısı söz konusu değildir. Bu nedenle tedavinin tekrar tekrar uygulanması gerekmekte ve bu da çiftlere maddi bir yükün yanı sıra ilaç tedavilerinden kaynaklanan sağlık sorunlarına, yan etki ve risklere ve sürecin belirsizliği nedeniyle de psikolojik ve ailevi sorunlara neden olmaktadır.

IVF tedavisi sırasında normalden fazla sayıda oosit ve buna bağlı olarak da çok sayıda embriyo gelişmektedir. Gelişen embriyolar içerisinde gebelik oluşturma potansiyeli en yüksek olan embriyo/ların seçimi büyük önem arzetmekte ve tedavi başarısını direkt olarak etkilemektedir (Raef, Maleki, & Ferdousi, 2020). Embriyo seçimi, gözlemsel şekilde morfolojik ve gelişimsel parametreler değerlendirilerek ve subjektif olarak yapılmakta, embriyonun tutunma potansiyeline etkili olan birçok faktörün aynı anda değerlendirilmesi mümkün olamamaktadır.

Ayrıca, IVF'in başarısına etki eden pek çok değişken olduğu (anne yaşı, dha önceki deneme sayısı, over rezervi, sperm parametreleri vs.) bilinmektedir. Bu değişkenlerden en önemlilerinin anne yaşı (Cimadomo et al., 2018), infertilite etiyolojisi (Vander Borght & Wyns, 2018) ve embriyo kalitesi (Cai, Wan, Appleby, Hu, & Zhang, 2014) olduğu bilinmektedir. Bu faktörler her bir IVF denemesinde değişik önemlerde başarıyı etkilemektedirler.

IVF'in başarısına etki eden faktörlerin ağırlıklarının bilinmesi, hastaya en uygun kontrollü overyan hiperstimülasyonun (the controlled ovarian stimulation-COS)

seçilmesini yani uygun tedavi yöntemi ve protokolünün, ilaç türü ve dozunun seçilmesini mümkün kılacağı gibi, hastanın bilgilendirilmesi, tedavi başarısının öngörülmesi ve sonraki tedavilerin planlanması açısından da önem arzetmektedir.

Makine öğrenmesi tabanlı klinik karar destek sistemleri tıbbi tanı ve tedavilerde kullanılan etkili yardımcılardır (Cimadomo et al., 2018). Bu sistemler, büyük veriden oluşan veri tabanınlarından öğrenir ve buna göre olası tahminler yapar. Bu nedenle bu karar destek sistemlerinin ART tekniklerinde kullanılması klinisyenler ve çifler için önem arz etmektedir.

Makine öğrenmesi teknikleri eğitim veri setinden öğrenen ve veriyi modelleyen, bu modeli, yeni bir veri setinde (test verisi, ya da yeni bir veri) doğrulayarak, tahminler yapan tekniklerdir. Bu yöntemlerin çok değişkenli verilerde yüksek başarısı nedeniyle klinik karar destek sistmlerinde sıklıkla kullanılmaktadırlar (Mencar et al., 2019; Tu, 1996). Makine öğrenmesi yöntemleri denetimli öğrenme ve denetimsiz öğrenme yöntemleri olmak üzere iki temel başlık altında yer alır. Bu çalışmada denetimli öğrenme tekniklerinden sınıflandırma problemi kullanılmıştır. En sık kullanılan sınıflandırma algoritmaları Karar Ağacı, Naive Bayes, Destek Vektör Makineleri ve Yapay Sinir Ağlarıdır (Raef & Ferdousi, 2019). Bu çalışmada, bu belirtilen yöntemlerin yanı sıra, IVF işleminin başarısı bir topluluk öğrenme yöntemleri olan Gradyan Arttırma Makineları, Xgboost ve Super Learner ile tahmin edilmiştir. Super Learner, bir çok algoritma ile tahmin olanağı sunup en iyi algoritmayı belirleyen ya da bu algoritmaların ağırlıklı olarak modele alınmasıyla daha iyi bir tahmin sunan algoritmadır (Laan & Rose, 2011; Van Der Laan, Polley, & Hubbard, 2007).

Literatürde, IVF tedavisinde embriyo başarısının modellenmesinde bu yöntemlerin yanı sıra başka makine öğrenmesi yöntemlerinin kullanıldığı görülmektedir, fakat bu çalışmalar yine de yeterli sayıda değildir (Barnett-Itzhaki et al., 2020; Raef et al., 2020). Super Learner algoritmasının IVF seçiminde embryo başarısının modellenmesinde kullanıldığı literarürde bir çalışma bulunmaktadır (Q et al., 2021).

Araştırmalar genellikle embriyolojik ve klinik veriyi bir arada kullanmamışlardır. Bu nedenle, çalışmaların bulgularının yorumlanması ve genellenmesini zorlaşmaktadır(Raef et al., 2020).

Türkiye'de yasal olarak 35 yaşına kadar 1, 35 yaşından sonra ise en fazla 2 embriyo transfer edilmesine izin verilmektedir. Bu düzenleme ile çoğul gebeliğin önüne

geçilmesi böylece çoğul gebeliklerden ortaya çıkan komplikasyonların (erken doğum, abortus, düşük doğum ağırlığı ve doğumsal anomaliler gibi) en aza indirilmesi hedeflenmektedir. (Spiessens et al., 2016; Uyar, Bener, & Ciray, 2015).

Transfer sayısının sınırlandırılması klinik olarak birçok avantaja sahip olmasına rağmen az sayıda embriyo transfer edileceğinden doğru embriyonun seçilmesini zorlaştırmaktadır. Bu nedenle doğru embriyonun seçimi daha da büyük öneme sahip olmaktadır.

Embriyo transferi, çoğunlukla embriyonun gelişiminin 2. ve 3. günü (bölünme evresi) ya da 5 ve 6. gününde (blastosist evresi) gerçekleştirilmektedir. Bu çalışmada transfer edilen embriyolar hem 3. Gün hem de 5. Gün transfer edilen embriyolardan oluşmaktadır. Literatürde yapılan çalışmalar genellikle 2 ve 3. Gün transfer edilen embriyolar ile yapılmıştır. Blastosit aşamasında transfer edilen embriyoların tutunma olasılığının daha yüksek olduğu iddia edildiği (Raef et al., 2020) halde bu evrede transfer sonuçlarını içeren çalışmalar sınırlıdır.

IVF tedavisinin başarısı çalışmalarda üç farklı şekilde incelenmektedir: positive beta-hCG, klinik hamilelik ve canlı doğum oranları. Pozitif beta-hCG, embryo transferinden 12-17 gün sonra yapılan beta-hCG testi 50'nin üzerinde olması başarılı bir şekilde embryonun tuttuğunun (implamantasyonun gerçekleştiğinin) göstergesidir. Klinik gebelik, gestasyonuun 7. haftasında, kesenin ve bebeğin kalp atışının ultrason ile görünülenmesi olarak tanımlanır. Canlı doğum ise gestasyonun 24. Haftasından sonra canlı doğumun gerçekleşmesi olarak açıklanır. Bu çalışmada, çalışmanın bağımlı değişkeni olan embriyonun implantasyonu, pozitif beta-hCG testine göre belirlenmiştir.

Bu çalışmanın amacı, embriyonun tutunma başarısını çiflerin demografik özellikleri, embriyonun kalitesi ve IVF sürecine özgü COS değişkenleri ile makine öğrenmesi yöntemleri ile tahmin etmektir. Bu amaçla, Super Learner algoritmasının diğer sekiz makine öğrenmesi yöntemi karşılaştırılmasının yanı sıra, IVF sürecine etki eden değişkenlerin önem dereceleri belirlenecektir.

Çalışmada kullanılan makine öğrenmesi yöntemleri Lojistik Regresyon, Naive Bayes, Rastgele Orman, Destek Vektör Makineleri, Yapay Sinir Ağları, Gradyan Artırma Makineları, XGBoost ve Super Learner'dır.

Çalışmanın sorusu “IVF tedavisinde, embriyonun tutunma başarısının modellenmesinde, çalışmada kullanılan demografik özellikler, embriyonun kalitesi ve COS değişkenlerine göre en iyi performans gösteren makine öğrenme yöntemi hangisidir/hangileridir?” Çalışmanın alt sorusu ise, “IVF tedavisinde, embriyonun tutunma başarısının modellenmesinde çalışmada kullanılan demografik özellikler, embriyonun kalitesi ve COS değişkenlerinin önem dereceleri nedir ?”

2. GENEL BİLGİLER

2.1. VERİ MADENCİLİĞİ

Veri tabanlarından bilgi keşfetme süreci (VTKB) (Knowledge discovery in databases) ve veri madenciliğinin amacı büyük veriden anlamlı bilgi ortaya çıkartmaktır. Çoğunlukla her ikisi aynı anlamda kullanılsa da, veri madenciliği veri tabanlarından bilgi keşfetme sürecinin bir parçasıdır(Mannila, 1996). Veri madenciliği VTBK sürecinde sadece modelin kurulması ve değerlendirilmesini içermektedir. VTBK süreci şu şekilde özetlenebilir (Han, Kamber, & Pei, 2012; Mannila, 1996):

Problemin tanımlanması: En önemli aşamadır, problem doğru tanımlanmadığında süreç başarısız olur,

Verilerin Birleştirilmesi: Farklı veri tabanlarından gelen verilerin birleştirilmesi,

Verinin temizlenmesi ve anlaşılması: Verinin incelenmesi, gerekli dönüşümlerin sağlanması, gürültülü verilerin temizliği, değişken seçimi gibi aşamaları içerir,

Veri madenciliği: Verinin modellenerek anlamlı bilgilerin, örüntülerin çıkarılması,

Örüntülerin Değerlendirilmesi: Elde edilen bilgilerin değerlendirildiği aşama

Raporlama: Elde edilen bilgilerin görsel ve rapor olarak sunulması aşaması

Veri madenciliği temelde, üç alanın birleşmesiyle oluşan yöntemdir: makine öğrenmesi, veri tabanları ve istatistik (Mannila, 1996).

2.2. MAKİNE ÖĞRENMESİ

Makine öğrenmesi VTBK sürecinde yer alan bir alt alandır. Makine öğrenmesi, makinenin veriden örüntüleri öğrenmesi ve bunu yeni bir veride (test verisi) tahmin etmesini ifade eden süreçtir. Makine öğrenmesinde öğrenme, çeşitli kaynaklarda farklı şekilde gruplandırılsa da temelde denetimli, denetimsiz öğrenme, yarı danışmanlı öğrenme, danışmansız öğrenme olmak üzere dörde ayrılır (Han et al., 2012).

2.2.1. Denetimli Öğrenme

Denetimli öğrenme (supervised learning), veri eğitim kümesinde etiketlenmiş veriyi öğrenerek tahminde bulunur (Mohri, Rostamizadeh, & Talwalkar, 2018). Yani, veri bir nevi eğitim kümesi ile deneyimlerden öğrenir, eğitim kümesi burada danışman görevi yapar. Bu yöntemler regresyon yöntemleri ve sınıflandırma yöntemleri olmak üzere ikiye ayrılır. Bu çalışmada da denetimli öğrenme tekniklerinden sınıflandırma teknikleri kullanılmıştır.

2.2.2. Denetimsiz öğrenme

Denetimsiz öğrenmede (unsupervised learning), veride bağımlı değişken yoktur, hiçbir veri etiketlenmemiştir (Mohri et al., 2018). Algoritma, verideki örüntüleri bulmaya odaklanır. Kümeleme yöntemleri ve faktör analizi gibi yöntemler örnek verilebilir (Mohri et al., 2018).

2.2.3. Yarı Danışmanlı Öğrenme

Yarı danışmanlı öğrenmede (Semi supervised learning), veri girişinde etiketlenmemiş veri ile birlikte az miktarda etiketlenmiş veri beraber kullanılır (Witten, 2017). Veri madenciliği ve makine öğrenmesi uygulamalarında yaygın olarak kullanılan bu yöntemler, etiketlenmemiş veri oranının fazla olduğu durumlarda tahmin oranının arttırılması için kullanılırlar (Zhu & Goldberg, 2009).

2.2.4. Aktif Öğrenme

Aktif öğrenme (Active Learning), kullanıcının (uzmanın) aktif rol oynadığı makine öğrenmesi yaklaşımıdır. Aktif öğrenme bir kullanıcıya bir dizi etiketlenmemiş veriyi etiketlenmesini ister. Amaç, model kalitesini arttırmaktır (Han et al., 2012).

2.3. TOPLULUK ÖĞRENME YÖNTEMLERİ

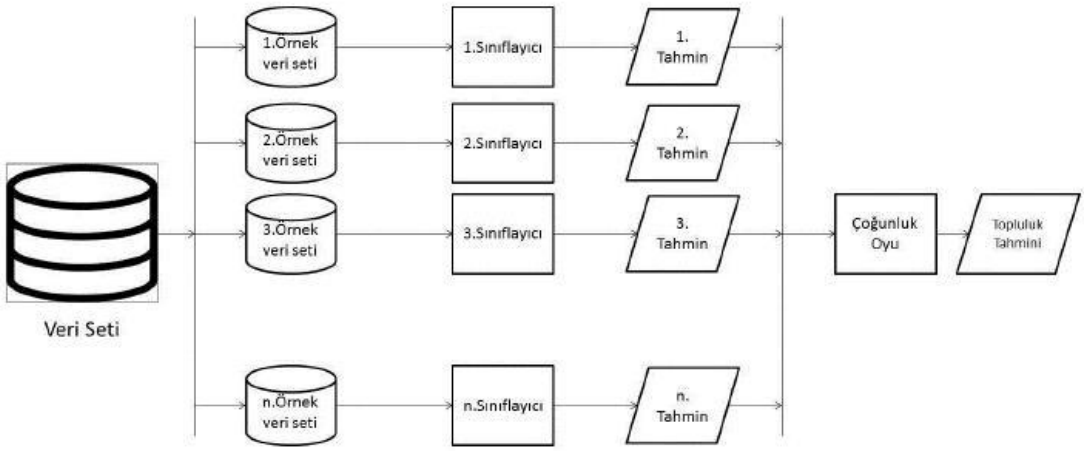
Topluluk öğrenme yöntemlerinin temel amacı bir çok algoritma ile yapılan tahminlerin birleştirilmesi ile verilen kararın, tek bir algoritma ile yapılan tahmin ile verilen karardan daha güvenilir olduğu esasına dayanır. Birden fazla uzmanın vereceği kararın hata olasılığı , tek bir uzmanın verdiği kararın hata olasılığından daha düşüktür. Topluluk öğrenme yöntemlerinin temel amacı yüksek varyansı azaltmak ve tahmin oranını arttırmak olsa da son yıllarda kayıp verilerin atanması, değişken seçimi, sınıf dengesizliği problemi, güven aralığının daraltılması gibi durumlarda da yaygın olarak kullanılmaktadır (Zhang & Ma, 2012).

ML yöntemlerinin tahmin performansını yükseltebilmesinin iki amacı vardır: yanlılığın azaltılması ve varyansın düşürülmesi (Zhang & Ma, 2012). Topluluk öğrenme yöntemlerinde bir topluluk, temel sınıflandırıcılar (base classifiers) adında çok sayıda sınıflandırma yöntemlerinden oluşur.

Topluluk öğrenme yöntemlerinde değişik birleştirme stratejileri kullanılmaktadır. Hangi yöntemin en iyi olduğuna karar vermek önemlidir. Bu topluluk öğrenme yöntemleri, boosting (artırma), bagging (torbalama), voting (oylama) ve stacking (istifleme) yöntemleridir.

2.3.1. Bagging (Torbalama)

Breiman tarafından geliştirilen bagging (bootstrap aggregating) yöntemi, eğitim verisinin bootstrap örnekleme ile farklı rastgele alt kümeler halinde eğitilmesine izin verir(L. Breiman, 1996a). Bootstrap örnekleme yönteminde eğitim veri seti içerisinden her seferinde iadeli örnekleme yöntemi ile rastgele alt kümeler oluşturulmaktadır. Oluşturulan her bir alt küme farklı bir sınıflandırıcı ile eğitilmektedir. Bu yöntemde, kullanılan sınıflandırma algotitmalarının tahminlerini birleştirmek üzere çoğunluk oyu tekniğimi kullanmaktadır. Tüm tahmincilerin sınıflandırma tahmini arasından sınıflandırıcıların verdikleri çoğunluk tahmini, bu yöntemin sonucu olarak kabul edilir.

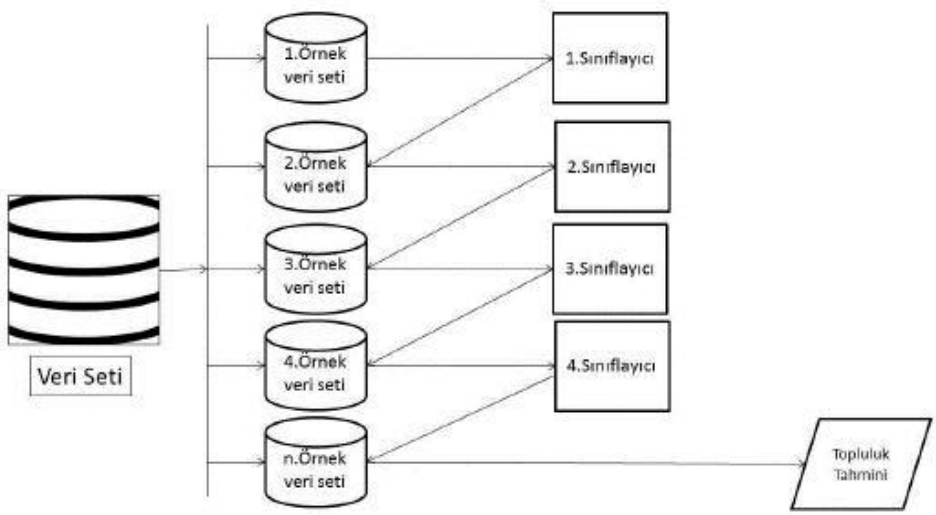


Şekil 2-1: Bagging Yönteminin İşlem Aşamaları (Doğaner, 2020)

2.3.2. Boosting (Arttırma)

Boosting ilk olarak Freund ve .Schapire tarafından 1997'de önerilmiştir (Bilgin, 2018). Boosting yöntemi, çok sayıda zayıf sınıflandırıcının birleştirilerek güçlü sınıflandırıcı elde etmek esasına dayanmaktadır. Her zayıf öğrenici bir tahmin yaptığında veri yeniden ağırlıklandırılmaktadır. Her zayıf öğrenicilerden, yanlış sınıflama yapanların

tahminleri düzeltilmeye çalışarak daha güçlü tahminciler ortaya koyar. Bagging yöntemindeki gibi sınıflandırıcılar bağımsız değildir, her bir sonuç bir sonrakini etkiler. En çok bilinen Boosting algoritması, Freund ve Shapire'ın geliştirdiği Adaboost algoritmasıdır(Freund, Schapire, & Hill, 1996). Algoritma gözlemleri ağırlıklandırarak başlar ve sınıflaması zor olanlara daha büyük, daha kolay olanlara daha küçük ağırlıklar verir. Performansı rassal daha az olan öğreniciler (karar ağaçları) zayıf öğreniciler olarak adlandırılır (Gürsakal, 2018).

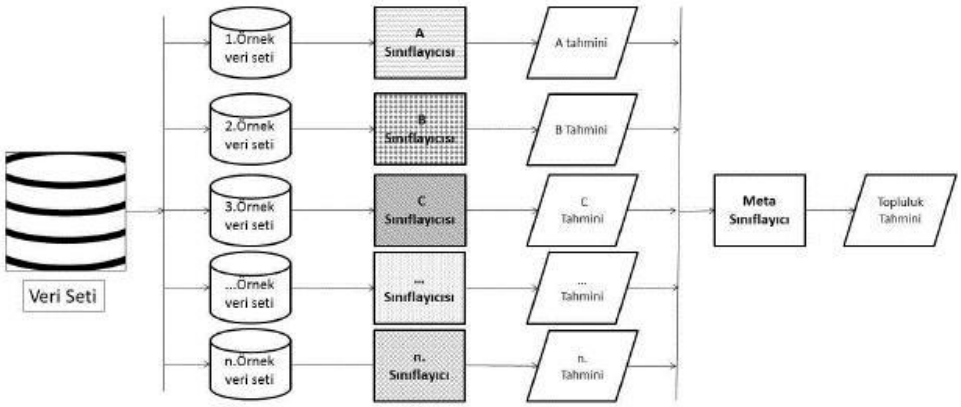


Şekil 2-2: Boosting Algoritması Çalışma Şekli (Doğaner, 2020)

2.3.3. Stacking (Yığma, İstifleme) Yöntemi

Stacking'de ML yöntemleri katmanlar halinde birbirinin üstüne yığılır, istiflenir. Her model kestirimlerini üstündeki modele aktararak sonuçta en üstteki model ile nihai

sonuçlar elde edilir (Gürsakal, 2018). Stacking bir meta öğrenme (öğrenmeden öğrenme) yöntemidir. Super Learner algoritması bir stacking yöntemidir.



Şekil 2-3: Stacking Yönteminin İşleyiş Aşamaları(Doğaner, 2020)

2.4. MAKİNE ÖĞRENMESİ ALGORİTMALARI

Bu bölümde çalışmada kullanılan sınıflandırma algoritmaları anlatılmıştır.

2.4.1. Lojistik Regresyon (LR)

LR, bağımlı değişkenin kategorik olduğu durumlarda kullanılır. Eğer bağımlı değişken, binary (ikili) ise lojistik regresyon, ikiden fazla kategorili ve ordinal ise ordinal lojistik regresyon, yine ikiden fazla nitelikli fakat nominal ise multinominal lojistik regresyon adını alır (Alpar, 2011). Lojistik regresyonda, bağımlı değişken 1, 0 değerlerini alır ve doğrusal regresyondaki gibi en küçük kareler (EKK) ile hesaplanamaz. LR, maksimum benzerlik (maximum likelihood) yöntemi kullanılır. Böylece EKK gibi sapmaların karesini en az yapmaya çalışmak yerine, olasılığı en yüksek yapmaya çalışır. Amaç, riski hesaplamaktır. P, 1 değeri verilecek şıkkın oranını tahminini belirtmek üzere ((Orhunbilge, 2010);

$$= \frac{}{+-(+\quad)} \quad \textbf{2-1}$$

Bağımlı değişkene verilen 0 ve 1 değerlerine odds dönüşümü uygulanır. Odds=(P/(1-P)).

$$\left(\frac{}{}\right) = + \quad \textbf{2-2}$$

Bu hesaplamar tek bağımsız değişken için gösterilmiştir. Çok değişkenli modeller için de benzer işlemler uygulanur.

Avantajları

LR modelini eğitmek ve kurmak kolaydır.

Değişkenlerin bağımlı değişkene etki derecesini ve ilişkinin derecesini hesaplamak mümkündür.

Değişkenlerin türü ile ilgili bir varsayımı yoktur.

Hızlıdır. Yüksek performans gösteren bilgisayarlara ihtiyaç duymaz.

Değişkenler arasındaki ilişki doğrusal olduğunda tahmin oranı yüksektir.

Dezavantajları

Değişken sayısı, gözlem sayısından fazla olduğunda aşırı uyum (overfitting) sorunu ortaya çıkar.

Doğrusal olmayan ilişkileri ölçemez.

Bağımlı değişkenin kesikli olması gerekir.

Bağımsız değişkenler arasında çoklu doğrusal bağlantı olmamalıdır.

Çalışmadaki bütün değişkenler modele alınamayabilir.

Çok değişkenli analizlerde aşırı uyum sorunu çıkabilir, bu da düzenlileştirme teknikleri ile bu sorun çözülmeye çalışılır.

2.4.2. Naïve Bayes (NB) Sınıflandırıcı

NB, Bayes teoremini kullanan istatistiksel bir sınıflandırma yöntemidir. Literatürde, karar ağacı ve bazı yapay sinir ağları sınıflandırıcıları ile benzer sonuçlar verdiği görülmüştür. Uygulamasının kısa sürmesi ve yüksek tahmin etme oranı ve yorumlanmasının basit olması nedeniyle tercih edilmektedir. Varsayımı, değişkenlerin birbirinden bağımsız olmasıdır (class-conditional independence). (Han et al., 2012). Çalışma şekli şöyledir (Han et al., 2012):

D eğitim verisi, =(1, 2,...,) n boyutlu nitelikler kümesi, =(1, 2,...,) m boyutlu sınıflar kümesi olsun. Amaç, sınıfı için (|)'i maksimum yapmaktır

(*maximum posteriori hypothesis)*. Yani NB, ’in sınıflarından hangisine ait olduğunu şöyle tahmin eder: (|)> (|) (1≤ ≤ , ≠) . Bayes teoreminden:

$$(/) = \frac{(/)()}{} \quad 2\text{-}3$$

P(X) bütün sınıflar için sabit olduğundan dolayı (|) maksimize edilir. Çok boyutlu verilerde (|) hesaplanması çok zor olduğundan, hesaplanmanın kolaylaştırılması için niteliklerin birbirlerinden koşullu bağımsız olduğu bağımsızlık varsayımı (*assumption of class conditional independence*) kullanılır. Böylece;

$$() = \prod = (\setminus) = (\setminus) (\setminus) \dots \dots \dots \dots . . (\setminus) \quad 2\text{-}4$$

X’in sınıfını tahmin etmek için sınıfları için (|) ( ) hesaplanmalıdır.

Avantajları:

Hızlıdır. Yüksek performans gösteren bilgisayarlara ihtiyaç duymaz.

Kategorik verilerde daha iyi performans sergiler.

Bağımsız değişkenlerin kendi arasında bağımsız olması varsayımı vardır. Bu varsayım sağlandığında, az sayıda veri ile yüksek tahmin oranlarına ulaşabilir.

Dezavantajları:

Değişkenler arasında bağımsızlık varsayımını sağlamak gerçek hayat uygulamarında zordur.

Eğer eğitim kümesinde olmayan bir kategori test veri setinde yer alırsa, model bu değeri tahmin edemez ve 0 olasılık olarak sunar. Bu problem sıfır frekans (zero frequency) problemi olarak adlandırılır. Bu sorunu çözebilmek için Laplace teoreminden yararlanılır.

2.4.3. Karar Ağaçları (KA)

İlk olarak, 1970’lerin sonu 1980’lerin başında, makine öğrenmesi çalışan J.Ross Quinlan ID3 (Iterative Dichotomiser) karar ağacı algoritmasını geliştirmiştir. Karar ağaçları (KA), kök düğümden başlayarak, dallar ve son olarak da yaprak düğümlerle ayrılan ağaca benzeyen bir akış diyagramıdır. KA’da düğümler, seviyelerine göre ebeveyn ve çocuk düğümler olmak üzere adlandırılır. Görselliğe dayandığı için kolay

yorumlanması, güçlü bir tahmin aracı olması diğer modellere göre daha hızlı sonuçlanması nedeniyle literatürde kullanımı yaygındır (Han et al., 2012).

Denetimli öğrenme tekniklerinden olan KA, hedef değişken kesikli olduğunda sınıflandırma ağacı (Classification Tree), sürekli olduğunda ise regresyon ağacı (Regression Tree) olarak adlandırılır (Hastie, Tibshirani, & Friedman, 2008).

Veri seti eğitim ve test verisi olarak bölündükten sonra, eğitim veri seti ile öncelikle öğrenme gerçekleşir. Ardından, test verisinde ağacın performansı değerlendirilir.

KA'da ağaç dallanması, ağacın dallarına ve yapraklarına nasıl bölüneceği, istatistiksel algoritmalar ile belirlenir. Öncelikle, bağımlı değişken için, en iyi kırılmayı sağlayan bağımsız değişkenler seçilerek dallar oluşturulur ve ilgili gözlemleri çocuk düğüme taşır. Böylelikle kök düğümden çocuk düğümler elde edilir. Ağaç dallanması, çeşitli durdurma kuralları kullanılarak ağaç daha fazla kırılma yapılamayacak duruma gelene kadar devam eder. KA'da ayrıca, modelin performansının arttırılması ve aşırı öğrenme sorunu olmasını engellemek için budama işlemi yapılır (Akpınar, 2014).

ID3, popular ve basit karar ağacı algoritmalarından biridir. C4.5 algoritması; ID3 algoritmasının hem sürekli hem de kesikli değişkenler için kullanılan genişletilmiş halidir (Akpınar, 2014) . ID3 algoritmasında eğitim kümesinin ne kadar iyi bölündüğünü ölçmek için **entropi** (*entropy*) ve **bilgi kazancı** (i*nformation gain*) algoritmaları kullanılırken, C4.5'de kazanç oranı (gain ratio) yöntemi kullanılır (Akpınar, 2014).

Denklem 2.5'de T kümesi için entropi denklemi verilmiştir (Kantardzic, 2020):

$$( ) = - \sum \left((\frac{ (,) }{ }) . \left(\frac{ (,) }{ } \right)\right) \qquad \textbf{2-5}$$

Bu T kümesi; X değişkenindeki n farklı değer için n parçaya ayrılabilir. Bu durumda beklenen bilgi ihtiyacı, entropilerin ağırlıklı toplamı şeklinde hesaplanır (Denklem 2.6):

$$( ) = \sum \left((\frac{ | | }{ }) . ( )\right) \qquad \textbf{(2-6)}$$

Gain (X)'i maksimize eden niteliği belirleyen kazanç oranı Denklem 2.7 ile hesaplanır:

$$Gain(X) = Info(T) - Info_X(T) \quad (2\text{-}7)$$

Kazanç kriteri kompakt karar ağaçların oluşturulmasında iyi bir performans sergilerken, çok sayıda çıktıya sahip olan modellerde ciddi yanlılık içerir. Bu problem normalizasyon yöntemi ile çözülmeye çalışılır (Kantardzic, 2020):

(2-8)

Yeni kazan kriteri formülü 2.9 ile hesaplanır.

(2-9)

Bu araştırmada kullanılan CART (Classification and Regressin Trees) algoritması C4.5 algoritmasının geliştirilmiş halidir. Temel farklılıkları ağacın oluşturulması, bölünme krieteri, ağacın budanması ve boş verileri nasıl ele aldığıdır. CART, C4.5'in bölünme kriteri olarak kullandığı bilgi kriteri yerine Gini index'i kullanılır. Gini Index her düğümde kullanılacak değişkeni seçer (Kantardzic, 2020).

c önceden tanımlanan sınıfları, C_i i=1,…c-1 için sınıfları, s_i her bir C_i sınıfına ait örneklem sayısını, $p_i=s_i/S$ C_i sınıfındaki bağıl frekansı göstermek üzere:

(2-10)

Bir özelliğin k alt kümesine Si bölünmesinin kalitesi, daha sonra ortaya çıkan alt kümelerin Gini indekslerinin ağırlıklı toplamı olarak hesaplanır (Denklem 2.11). n_i, bölünmeden sonra S_i altkümesindeki örneklem sayısı, n ise ilgili düğümdeki toplam örneklem sayısı olmak üzere:

(2-11)

Avantajları

Diğer modellerde yer alan normalizasyon, ölçeklendirme gibi veri önişleme adımları olmadığı için bu aşama diğer modellere göre daha kısa sürer.

Bulguları yorumlamak kolaydır.

Görsel olarak da bulguları görmek mümkündür.

Hem sayısal hem de kategorik verilere uygulanabilir.

Doğrusal olmayan ilişkiler de incelenebilir.

Belli derecede boş veriyi analiz tolere edebilir.

Dezavantajları

Aşırı öğrenme sorunu vardır.

Aşırı öğrenme sorunundan kaynaklı yüksek varyans sorunu ortaya çıkar. Büyük veride tahmin yapmak uygun değildir.

Öğrenme süreci uzun sürer.

2.4.4. Rastgele Orman (Random Forest-RO) Sınıflandırıcı

RO, ilk defa Ho (Ho, 1998)'nun önerdiği rastgele uzay yöntemini Breiman(L. E. O. Breiman, 2001)'ın torbalama yöntemi ile desteklediği bir karar ağacı algoritmasıdır. RF, bir topluluk öğrenme yöntemidir (ensemble learning) ve ağaçlar CART (Sınıflandırma ve Regresyon Ağaçları), bootstrap ve bagging yötemleri ile oluşturulur(L. E. O. Breiman, 2001). Bir rastgele orman algoritması şöyle çalışır (L. E. O. Breiman, 2001): eğitim veri seti Bootstrap yöntemi ile 2/3'ü karar ağaç oluşturacak, 1/3'ü ise ağaç oluşturmayacak (test veri seti) şekilde ikiye bölünür. Eğitim veri setinde CART karar ağacı yöntemi ile ağaç oluşturulur. Ağaç oluşturulurken her bir düğümde m sayıda bağımsız değişken seçilir. M bağımsız değişken sayısı olmak üzere m<M olmalıdır. "m" değerine karar verirken bilgi kazancı en yüksek olan tercih edilir, bu amaçla gini katsayısından yararlanılır. "k" toplam sınıf sayısı, P_k, düğümdeki k sınıfına ait örneklem sayısı olmak üzere,

$$()=\Sigma= (-) \qquad 2\text{-}12$$

Karar ağacı kurulduğunda her bir yaprak düğüm bir sınıfa atanmış olur. Model kurulduktan sonra, test verisi için sınıflar atanır. Bundan önceki adımlar B defa tekrarlanır. Her bir gözlemin denemeler sonucundan kaç kez hangi sınıfa atandığı çoğunluğa göre karar verilerek her bir gözlemin nihai sınıf ataması yapılır.

RF algoritması aşırı öğrenme sorununu çözdüğü ve topluluk öğrenmesi yöntemi olduğu için çok iyi bir tahmincidir. RF algoritmasının bir diğer önemli özelliği değişkenlerin önem derecelerini belirlemesidir. Değişkenlerin önem derecelerini yine bilgi kazancı, gini yöntemi ile yapar(Archer & Kimes, 2008).

Avantajları

DT'deki aşırı öğrenme ve yüksek varyans sorunları, topluluk öğrenme yöntemleri kullanıldığı için RF'da yoktur.

RF hem sayısal hem de kategorik veriler için kullanılabilir.

Veri dönüşümü gerektirmez.

Boş verilerle çalışabilir.

Doğrusal olmayan ilişkileri ortaya çıkartabilir.

Uç değerlerle çalışabilir.

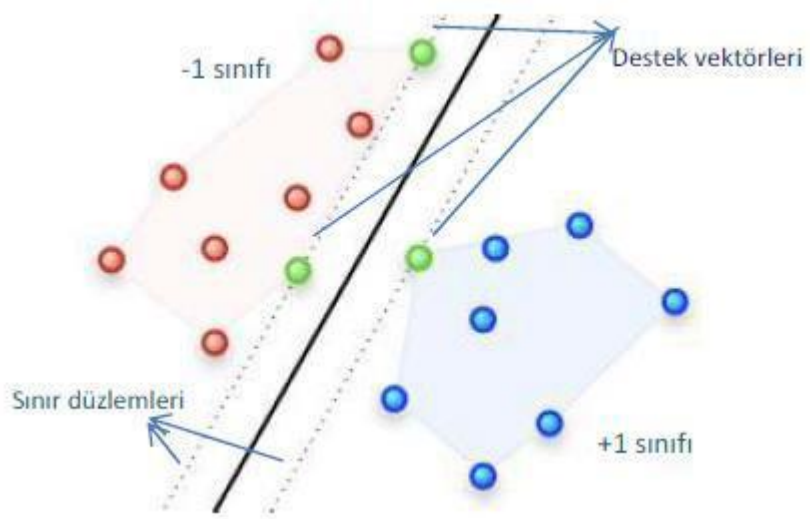
Dezavantajları

Fazla sayıda hiper parametresi olduğu için en iyi performansı gösteren modeli bulmak zaman alır.

Modelin öğrenme süresi fazladır. Yüksek performanslı bilgisayarlara ihtiyaç duyar.

2.4.5. Destek Vektör Makineleri (DVM)

DVM, 1960'lı yılların sonunda Vladimir Vapnik ve Alexey Chervonenkis tarafından geliştirilmiş, temeli istatistiksel öğrenme teorisine dayanan bir makine öğrenmesi yöntemidir (Yilmaz Akşehirli, Ankaralı, Aydın, & Saraçlı, 2013). DVM'de sınıfları birbirinden ayıracak optimal hiper ayırma düzlemi elde etmek amaçlanır. Yani, farklı sınıflara ait destek vektörleri arasındaki uzaklığı maksimize eder (Ayhan & Erdoğmuş Şenol, 2014).



Şekil 2-4: İki Sınıflı Doğrusal DVM Örneği (Ayhan & Erdoğmuş Şenol, 2014)

DVM ilk olarak, eğitim kümesini iki sınıfa ayıran hiper ayırma düzlemi doğrusal olarak geliştirilmiştir. Eğitim kümesi $(x_1, y_1),\ldots(x_l, y_l)$, $x \varepsilon R^n$, $y_i \varepsilon \{-1, +1\}$ olmak üzere, $\omega.x+b=0$ hiper düzlemi ile ayrılabilir. Eğer gruplar hatasız olarak ayrıldıysa ve en yakın vektör ile hiper ayırma düzlemi arasındaki mesafe maksimum ise bu vektörler seti optimal hiper düzlem ile ayrılmıştır denir. Bu ayıran hiper düzlem formülü(X. Wang & Zhong, 2003):

(. +) ≥ ,= , …. . **(2-13)**

Buna göre, optimal hiperdüzlem, Denklem 2.13'ü sağlayan ve $^1_2 \| \|^2$'ü minimize eden fonksiyondur. Vapnik (Vapnik, 1995; X. Wang & Zhong, 2003) bu minimizasyonun denklem 2.14'ü maksimize ederek sağladığını göstermiştir (α_i değişkenlere göre):

(2-14)

$0 \leq \alpha_i$, i = 1,…l and $\sum_{=1} = 0$. Bu x_is'ler ($0<\alpha_i$) Destek Vektörlerdir (Support Vectors, SVs. Bunlar genellikle X_{SVM} olarak ifade edilen eğitim veri setindeki küçük alt gruplardır. Bilinmeyen x_i vektör sınıflandırması için Denklem 2.15'deki gibi bulunur:

( ) = Σ ϵ (.) +) **(2-15)**

ve $0 < \alpha_i$ olan ve toplam, sıfır olmayan DV'ler için

$= \Sigma \epsilon$ **(2-16)**

Eğer optimal hiper düzlem doğrusal olarak ayrılamıyorsa DVM iki farklı yöntem kullanır: Birincisi, hataları eğitir. İkinisi ise, φ (x) fonksiyonu ile veriyi daha yüksek boyutlu doğrusal olmayan bir özellik uzayına dönüştürür(X. Wang & Zhong, 2003). Bu yüksek uzayda, veriyi doğrusal olarak ayırmak daha mümkün hale gelir. Böylelikle problem denklem 2.17'deki gibi ifade edilir:

$\_ \| \| + \Sigma$ **(2-17)**

Bu bir kuadratik optimizasyon problemidir ve esnek (slack) değişkenler eklenerek kısıtlar dikkate alınır.

Avantajları

Çok boyutlu verilerde tahmin yöntemi olarak kullanılabilir.
Değişken sayısı örneklem sayısından fazla olan durumlarda bile iyi performans sergiler.
Kernel fonksiyonu ile karmaşık problemleri de çözer.
Aşırı uyum sorunu azdır.

Dezavantajları

Hiper parametresi çok olduğu için doğru modeli bulmak zordur.
Öğrenme süresi uzun sürebilir. Özellikle büyük verilerde çalışması zordur.
Çıktısını yorumlamak zordur.

2.4.6. Yapay Sinir Ağları (YSA)

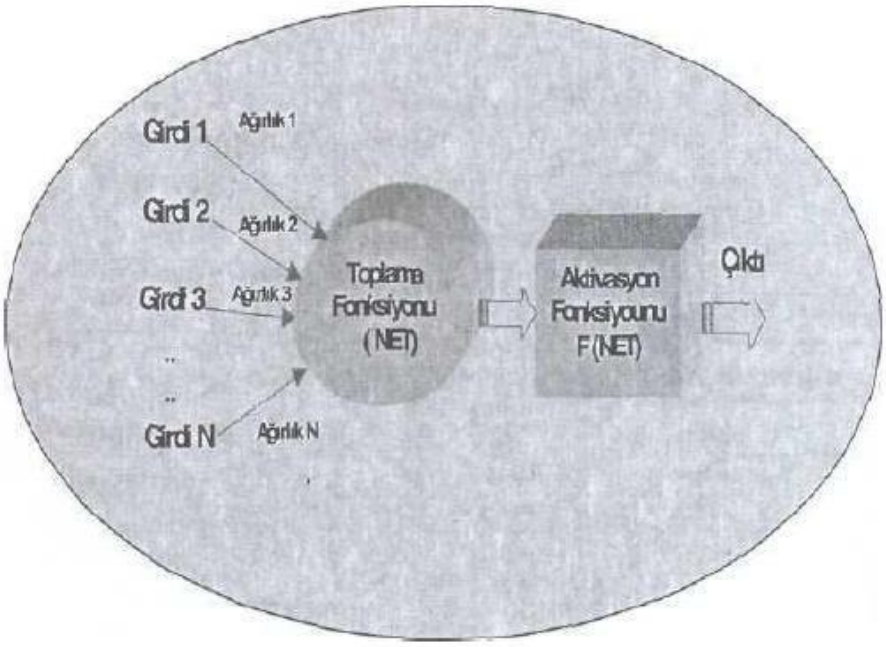
YSA, öğrenme ile insan beyninin yetenekleri olan yeni bilgiler üretebilme, yeni bilgiler oluşturma ve keşfedebilme gibi özellikleri otomatik olarak gerçekleştirebilecek adaptif bilgi işleme ile ilgilenen bilgisayar bilim dalı olduğu söylenebilir (Öztemel, 2006). YSA, beyindeki sinir hücresi olan nöronların çalışma şeklini bilgisayarda modelleyerek beynin öğrenme, ilişkilendirme, genelleme gibi özelliklerini simüle etmeye çalışır.

Şekil 2.2'de görüldüğü gibi temel yapay sinir ağı hücresinin, girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve hücrenin çıktısı olmak üzere beş temel elemanı vardır (Öztemel, 2006).

Girdiler x_i , ağırlıklarla çarpılır (w_{ij}) ve eşik değer (b_j) ile toplanır. Ardından doğrusal olmayan hesaplamaların kullanılmasına izin veren aktivasyon fonksiyonu f () yardımı ile çıkış (y) değeri elde edilir. Bu işlemler denklem 2.18-19'da belirtilmiştir.

=Σ **+b** **(2-18)**

o-œ) **(2-19**



Şekil 2-5: Yapay Sinir Hücresinin Yapısı(Öztemel, 2006)

Yapay sinir hücreleri bir araya gelerek yapay sinir ağını oluştururlar. Genellikle üç katmandan oluşan bir yapı gösterirler: girdi katmanı (çalışmadaki değişkenler), gizli katmanlar (birden fazla olabilir) ve çıkış katmanıdır. Gizli katman sayısı araştırmacı tarafından belirlenir.

Avantajları

Büyük veride çalışma performansı yüksektir.

Doğrusal olmayan ilişkileri tahmin etmekte başarılıdır.

Boş veriler olduğunda iyi performans sergiler.

Uç değerler olduğunda iyi performans sergiler.

Hiper parametrelerden dolayı öğrenme uzun sürer.

Dezavantajları

Modeli nasıl kurduğu bilinmediği için kara kutu özelliği vardır. Bu en önemli dezavantajıdır. Bu durum, yöntemin güvenilirliğini azaltır.

Aynı anda pek çok işi yapma özelliğinden dolayı (paralel processing) yüksek performanslı makinelere ihtiyaç duyar.

Diğer makine öğrenmesi yöntemlerine göre aşırı uyum sorununun olmaması için daha fazla sayıda örnekleme ihtiyaç duyar.

Hiperparametre sayısı fazladır, doğru modeli bulmak zordur.

Modelin öğrenmesi uzun sürer.

2.4.7. Gradyan Artırma Makineları (Gradient Boosting Machine-GBM)

GBM algoritması, 2001 yılında Friedman tarafından geliştirilmiş, temel amacı fonksiyon uzayında türevsel azalış (gradient descent) yoluyla riski (loss function) en aza indirgemek olan bir karar ağacı tabanlı boosting algoritmasıdır (Friedman, 2001).

GBM zayıf sınıflandırıcıyı ağırlıklandırarak eksikliklerini tamamlar ve böylece güçlü bir sınıflandırıcı oluşturur (C. Wang, Kou, & Song, 2019). Kayıp fonksiyon L(y,F(x))=log(1+e$^{-yF(x)}$, F(x) tahmin edilen değer, $( ) = \sum_{=1} h(; ,)$'dir. Kayıp fonksiyonu minimize etmek için F(x)'e her bir iterasyonda hata terimleri de eklenir. Modelin işleyici aşağıdaki gibi özetlenebilir (Ding, Cao, & Næss, 2018; Friedman, 2001; C. Wang et al., 2019):

1. Model sabit sayı c ile başlar:

$$( ) = \Sigma \quad (,) \qquad \textbf{(2-20)}$$

2. m=1…M;

   2.1. i=1…N pseudo-residuals hesaplanır:

$$= - [\quad \frac{(, ( ))}{(\quad)}] \qquad \textbf{(2-21)}$$

2.2. h_{mi} karar ağacını eğitim setine $\{(\ ,\ )\}_{=1}$ uydur:

$$= \frac{}{+\ (\ -()\ )} \quad \textbf{(2-22)}$$

2.3. L kayıp fonksiyonunu minimize ederek βm ağırlıkları hesaplanır:

$$=(,\ -()+) \quad \textbf{(2-23)}$$

2.4. Fm'i, güncele:

$$= -1() + \qquad h$$

3. Çıktı $F_m(x)$.

GBM algoritması ağaç sayısı, her ağaçtaki gözlem sayısı, terminal düğümde yer alacak minimum gözlem sayısı, öğrenme oranı ve her ağaçtaki düğüm sayısı olmak üzere beş hiperparametreye sahiptir.

Avantajları

Veri dönüşümüne gerek duymaz.

Hem sayısal hem kategorik verilerde iyi çalışabilir.

Boş verilerle çalışabilir.

Tahmin gücü yüksektir.

Dezavantajları

GBM'de hatalar sürekli minimize edildiği için çapraz geçerleme mutlaka kullanılmalıdır.

Hiper-parametresi fazladır.

Modelin öğrenmesi uzun sürer.

Çıktılarını yorumlamak güçtür.

2.4.8. XGBoost Sınıflandırıcı

Karar ağacı tabanlı topluluk öğrenmesi yöntemlerinden XgBoost algoritmasının çalışma şekli GBM'ye benzemektedir. Her iki algoirtma da temelde zayıf sınıflandırıcıları iterative ederek güçlü sınıflandırıcılar elde etmeye çalışır. Fakat iki temel

farklılıkları vardır: Birincisi, Xgboost'da kayıp fonksiyonunun yanında, ağacın karmaşıklığını azaltan başka bir fonksiyon mevcuttur. İkinci olarak da GBM'de ağaç sadece kayıp fonksiyonunun birinci dereceden gradyanına uydurulur. XGBoost'da ise her iterasyonda yeni bir karar ağacı oluştururken birinci dereceden gradyanı hem birinci dereceden hem de ikinci dereceden gradyanlara uydurur(C. Wang et al., 2019).

Algoritmanın hiperparametereleri öğrenme oranı, gamma (min kayı fonksiyonu), yaprakta ihtiyaç duyulan minimum gözlem sayısı, ağaç sayısı, ağacı oluşturmak için gerekli satır oranı, ağacın derinliği, ağaçları oluşturmak için gerekli alt örnek oranı olarak sayılabilir.

Avantajları:

Xgboost kendi içinde düzenlileştirme (Lasso and Ridge Regressions) tekniklerini uyguladığı için GBM'nin düzenleştirilmesi olarak adlandırılır. Bu nedenle aşırı uyum sorununun önüne geçilir.

Paralel öğrenme özelliği ile GBM'den daha hızlıdır.

Boş veri, normalizasyon gibi veri ön işleme aşamaları gerektirmez.

Kendi içinde çapraz geçerleme özelliği kullanır.

Dezavantajları:

Çıktıları yorumlamak güçtür.

Hiper-parametre sayısı çok olduğu için veriyi eğitmek güçtür.

Aşırı öğrenme sorunu yine de ortaya çıkabilir.

2.4.9. Super Learner (SL)

Süper öğrenme (Super Learning) teorik olarak Van der Laan vd. (Van Der Laan et al., 2007) tarafından geliştirilen genel kayıp temelli öğrenme tahmin yöntemidir. Super Learner (SL) tahmin yöntemlerinin içinden çapraz geçerlilik riskini minimum yapan en uygun kombinasyonunu bulmak için tasarlanmış bir tahmin yöntemidir(Polley, Hubbard, & Laan, 2010). SL, Van der Laan vd. (Van Der Laan et al., 2007) tarafından geliştirilmiş olsa da, temelleri yapay sinir ağları stacking algoritmasına (Wolpert, 1992) ve regresyon adaptasyonu olan stacking regresyondur(L. Breiman, 1996b).

SL, Van der Laan vd.(Van Der Laan et al., 2007)'daki haliyle anlatılacaktır.

Gözlenen öğrenme veri seti $X_i= (Y_i,W_i)$, i=1,……n Y çıktı değişkeni, W ise p boyutlu kovariate seti. Amaç, $\psi_0(W)=E(Y\backslash W)$ fonksiyonunu tahmin etmektir. Beklenen kaybı minimize eden fonksiyon:

$$\psi(W)=[\ (\ ,\ (\ ))] \qquad \textbf{2-24}$$

Kayıp fonksiyonu genellikle kayıp hatalarının karesi olarak alınır: L_2: $(Y-\psi(W))^2$.

Bir problem için, bir çeşitli tahmin yöntemlerinden oluşan algoritmalar kütüphanesi önerilir. Algoritma kütüphanesi $\mathcal{L}$ ile gösterilsin ve K(n), $\mathcal{L}$'nin kardinalitesi olsun.

1. $\hat{\ }$ $\mathcal{L}$ içindeki her bir $(\ ),\ =1,\ldots,\ \ (\ )$ 'i tahmin etmek için algoritma bütün veri setine $X=\{X_i: i=1,….,n\}$ uygulanır.
2. Veri seti X, V kat çapraz geçerleme (V-fold cross validation) kuralına göre eğitim ve validasyon veri setleri olmak üzere ikiye ayrılır. (n) gözlem içeren veri, V eşit parçaya bölünür v=1,…,V, v. veri seti validasyon veri seti geri kalan veri ise eğitim veri seti olarak alınır,. T(v) v. eğitim veri seti ve V(v) eş validasyon seti olsun. T(v)=X/V(v), v=1,…V.
3. V. kat için, $\mathcal{L}$ içindeki her bir algoritmayı T(v) için uygula ve tahminleri $\hat{\ }$ eş validasyon veri seti için kaydet, $\ (\ ;\ )(\ )$ v =1,..V.
4. K matrisi ile bir n oluşturmak için her algoritmadan gelen tahminler bir arada istiflenir (stack)

$\hat{\ }\ (\ )$ notasyonu $V=\{\ ,\ (\ )(\ \ (\ )),\ =1,\ldots,\&\ =1,\ldots,\ \}\ (\ )=(\ :)$ (v) validasyon kümesin kovarans vektörü için kullanılır.

5. Ağırlık vektörü α ile indekslenmiş aday tahmin edicilerin ağırlıklı kombinasyonlarından oluşan bir takım önerilir.

$\hat{\ }$

$$(\ \backslash\ )=\sum=\ \ .(\ )\ (\ (\ )),\ \geq\ ,\forall\ ,\sum=\ \ = \qquad \textbf{2-25}$$

6. Bütün olası tahminlerlerden, aday tahmin edicilerin $\sum=1$ $\hat{\ }$ çapraz geçerleme riskini minimize eden belirlenir.

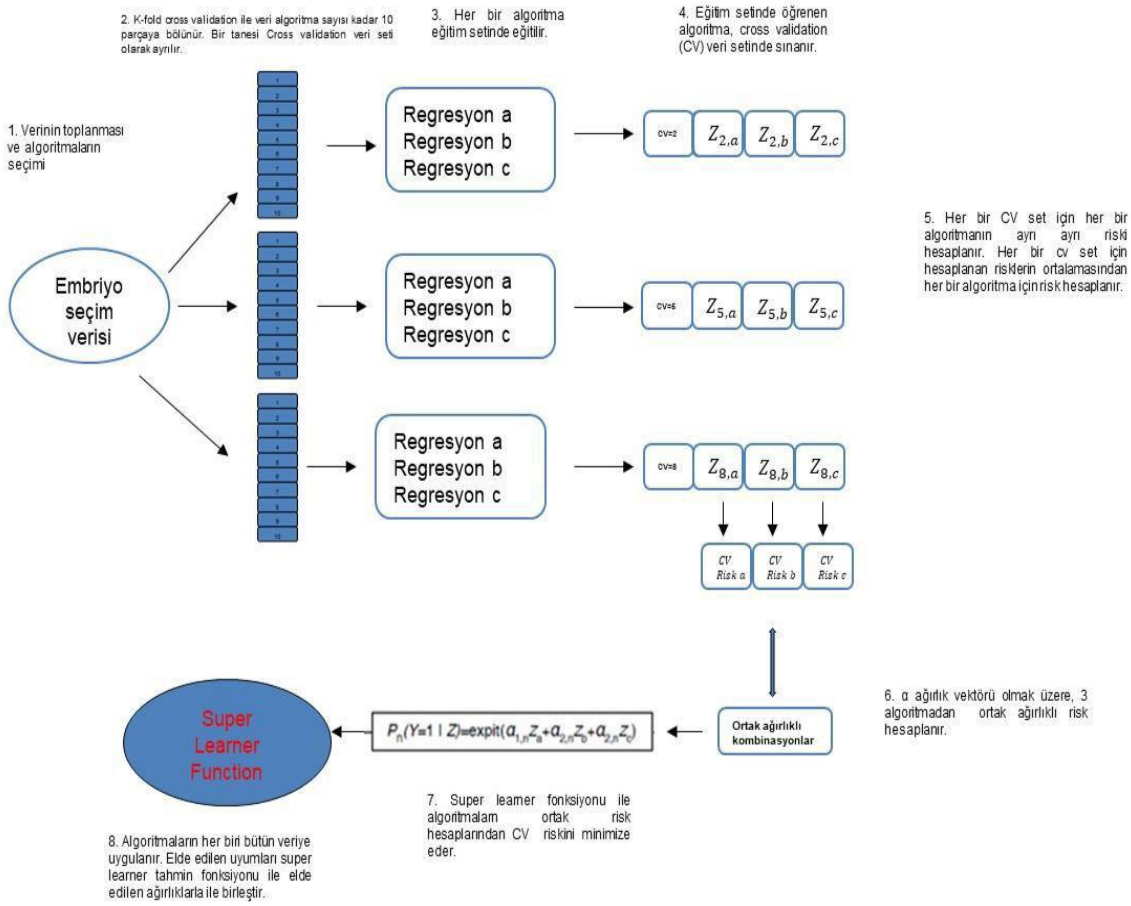
ˆ =Σ (− (\)) **2-26**

7. Nihai super learner'ın uyumu için, m(z|) ağırlıklı kombinasyonları

içinden ''lan ( ), k=1,…K ile birleştirilir:

ˆ ˆ

( ) = Σ ˆ ( ) **2-27**



Şekil 2-6: Embriyo seçimi için super learner algoritması Akış Diyagramı (Laan & Rose, 2011)

Bu algoritmanın akış diyagramı Şekil-2.2.'de gösterilmiştir. Çalışmada öncelikle K-Çapraz Geçerleme yöntemi kullanılarak ve veri seti 10 eşit parçaya bölünür. Bu veri kümelerinden bir tanesi cross-validation (çapraz doğrulama) veri seti olarak ayrılacaktır. Geri kalan dokuz tanesi eğitim veri seti olarak kullanılır. Eğitim veri setinden elde edilen en uygun model, geri kalan bir çapraz doğrulama veri setine uygulanacaktır. Çapraz doğrulama setinde (verinin % 10'undan) öngörülen olasılıkları kullanarak tahmin riski

hesaplanır. Bu süreç, modelde kullanılan algoritma sayısı kadar tekrar edilir. Her bir algoritma için her bir eğitim seti için ayrı ayrı model kurulur. Modeller, her bir validasyon seti için uygulanıp tahmin riski hesaplanır. Ardından, Cross-Validation setlerinden elde edilen ağırlıklar her bir model için eğitim setinde tekrar kullanılır (diğer 9 parça) ve ağırlıklar hesaplanır. Böylelikle bütün veri seti tahmin riskinin hesaplanmasında yer alır. Böylelikle her bir algoritma için 10 farklı risk tahmin edilir ve bu riskler , her bir algoritma için CV veri setlerinden elde edilen ortalama risktir. Şekil-1'de kullanılan model , 3 algoritma kullanıldığındaki durumdur.

Ardından, modeldeki algoritmalar için ortak bir ağırlık vektörü belirleyebilmek için ağırlık vektörü α'dan index hesaplanır. Amaç, bu ortak kombinasyonlardan elde edilen ağırlıklardan hangisinin minimum çapraz doğrulama riskine sahip olduğunu bulmaktır $\alpha = \{\alpha_a, \alpha_b, \alpha_c\}$). Her bir algoritma için hesaplanan ağırlık değerinden sonra , embriyo seçiminin modellenmesinde bütün veriye bu model uygulanır. Böylelikle, en düşük tahmin riskine sahip olan model elde edilmiş olur.

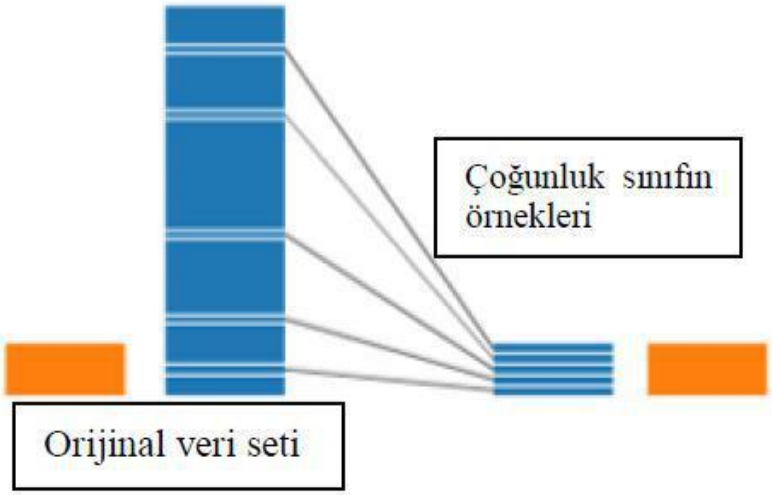
SL'de kullanılan algoritmalar içerisinden en uygun modelin seçilmesine Kesikli SL (Discrete Super Learner), istiflemede kullanılan algoritmalara bir ağırlık vererek yeni bir hesaplama yapılarak tahmin yapılıyorsa Super Learner (SL) adını alır(Baćak & Kennedy, 2018).

2.5. Sınıf Dengesizliği Problemi Yöntemleri

Bir sınıflandırma modelinde gruplardan bir tanesinin örneklem sayısı diğerinden büyük ölçüde az olduğunda, bu durum sınıflandırma algoritmaları açısından sınıf dengesizliği problemi ortaya çıkartmaktadır. Bu sınıflar azınlık ve çoğunluk sınıfları adıyla adlandırılır. Literatürde bu problemi ortadan kaldırmak için kullanılan pek çok yöntem vardır. Bu çalışmada analizde kullanılan rastgele az örnekleme (RUS), rastgele yukarı örnekleme (ROS), ROSE ve SMOTE yöntemleri açıklanacaktır.

2.5.1. Rastgele Az Örnekleme (RUS)

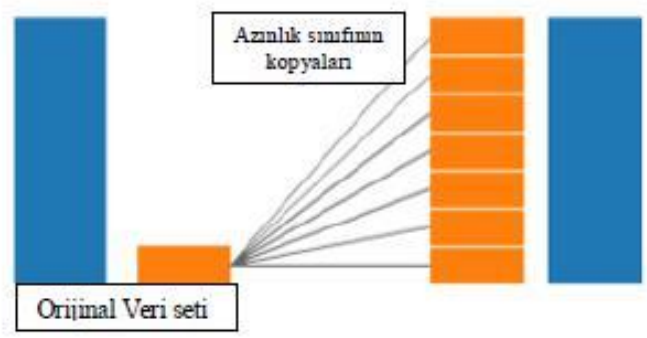
Azınlık sınıfı aynı kalmak koşuluyla, çoğunluk sınıfı azınlık sınıfı ile eşit örnekleme sahip olacak şekilde rastgele örneklemlerin seçilmesine rastgele az örnekleme adı verilir(Han et al., 2012).



Şekil 2-7: RUS örnekleme (Koç, 2019)

2.5.2. Rastgele Yukarı örnekleme (ROS)

ROS, çoğunluk sınıfında bulunan gözlemleri sabit bırakarak, azınlık sınıfını rastgele gözlemler atayarak gözlem sayısının artmasını sağlayan yöntemdir(Han et al., 2012). Burada azınlık sınıfında tekrarlı örnekler yer almasından dolayı aşırı örnekleme sorunu ortaya çıkabilmektedir.



Şekil 2-8: ROS örnekleme

2.5.3. SMOTE Örnekleme (Sentetik Azınlık Asırı Örnekleme)

Chawla (Chawla, Bowyer, Hall, & Kegelmeyer, 2002) ve arkadaşları tarafından önerilen SMOTE yöntemi, ROS'un zayıf yönlerine karşın önerdikleri, sentetik verilerin üretildiği bir örnekleme yöntemidir. Sentetik örnekler oluşturulurken azınlık sınıfındaki örneklerin k en yakın komşulukları (kNN) rastgele belirlenir. Adımlar şu şekildedir (Chawla et al., 2002): (1) azınlık sınıfındaki gözlemler için k yakın komşu değerleri bulunur, (2) azınlık sınıflarının gözlemleri ile k yakın komşulukları arasındaki fark değeri alınır, (3) α (0-1) arasında rastgele seçilir ve 2. adımdaki fark değerleri ile çarpılır,(4) $= + (-) *$ formülüyle sentetik gözlemler istenen sayıda elde edilir.

2.5.4. ROSE (Random Over-Sampling) Örnekleme

ROSE örnekleme tekniğinde, azınlık sınıfı bootstrap yaklaşımı ile RUS ve ROS yöntemleri birleştirilerek örnekler oluşturulur. Öncelikle, çoğunluk sınıfının yarısı bootstrap ile azaltılır (RUS), ardından azınlık sınıfına bootstrap ile örneklem sayısı arttırılır (ROS)(Tantithamthavorn, Hassan, & Matsumoto, 2018).

2.6. Sınıflandırma Problemlerinde Kullanılan Performans Ölçütleri

Bu çalışmada sınıflandırma problemlerinin performanslarının ölçülmesinde ve karşılaştırılmasında genel duyarlılık (sensitivity / recall), seçicilik (specificity), kesinlik (precicion-positive predictive value), doğruluk (Accuracy),F ölçütü (F-measure), AUROC(ROC eğrisinin altında kalan alan) ve kappa istatistiklerinden yararlanılmıştır. Bu istatistikler hesaplanırken karmaşıklık matrisinden yararlanılmıştır (confusion matrix). karmaşıklık matrisi şu şekildedir (Han et al., 2012):

Tablo 2-1: Karmaşıklık Matrisi(Han et al., 2012)

Karmaşıklık Matrisi		**Tahmin**	
		Negatif	**Pozitif**
Gerçek	**Negatif**	Doğru Negatif (DN)	Yanlış Pozitif (YP)
	Pozitif	Yanlış Negatif (YN)	Doğru Pozitif (DP)

2.6.1. Doğruluk

Doğru sınıflandırılmış gözlemlerin toplam örneklem sayısına oranıdır. 0 ile 1 arasında bir değer alır, 1'e yakın olması doğruluk oranının yüksek olduğunu gösterir. Sınıf dengesizliği problemi olduğu durumlarda güvenilir olmadığı literatürde yer almaktadır.

ğ = + 2-28

2.6.2. Seçicilik

Seçicilik, gerçekte negatif olan gözlemlerin, tahmin sonucunda da negatif olma olasılığıdır(Han et al., 2012). 0 ile 1 arasında bir değer alır, 1'e yakın olması seçicilik oranının yüksek olduğunu gösterir.

ç = **2-29**

2.6.3. Duyarlılık

Duyarlılık (recall), gerçekte pozitif olan bir gözlemin, tahmin sonucunda da pozitif olma olasılığıdır(Han et al., 2012). Yine 1'e yakın olması yüksek duyarlılık anlamına gelir.

= **2-30**

2.6.4. Kesinlik

Gerçekte pozitif olan değerlerin, tahmin sonucunda da pozitif olma olasılığıdır (Pozitif Predictive Value) .

= **2-31**

2.6.5. F Değeri

Var olan yöntemlerin tek başına yeterli olmayacağı görüşünden yola çıkarak ortaya çıkmış, kesinlik ve duyarlılık değerlerinin harmonik ortalaması ile hesaplanan değerdir. Yüksek olması yüksek performansı temsil etmekte, literatürde özellikle sınıf dengesizliği problemlerinde dikkati alınması tavsiye edilen ölçüttür.

= * * **2-32**

2.6.6. Kappa İstatistiği

Gözlemciler arası uyumu test etmek için kullanılan Cohen (Cohen, 1960) tarafından geliştirilen kappa istatistiği yapılan sınıflandırmanın başarısı hakkında da fikir verdiği için sıklıkla kullanılır.

2.6.7. ROC Eğrisinin Altında Kalan Alan (AUROC)

(1-seçicilik), y ekseninde ise duyarlılık ile oluşturulan eğridir. AUROC, farklı duyarlılık ve seçicilik değerleri ile elde edilen ROC eğrisi altında kalan alandır(Han et al., 2012). AUROC değeri, 0 ile 1 arasında değişmektedir. Bu değerin yüksek olması performansın yüksek olduğunu göstermektedir.

2.6.8. Modelin Performansının Değerlendirilmesi

Modellerin tahmininde temel yaklaşım, modelin öğrenebilmesi için eğitim veri seti ve tahmin gücünü ölçmek için ise test veri seti olarak iki gruba ayırmaktır. Veri madenciliğinde veriyi yeniden örneklemek için pek çok strateji geliştirilmiştir. Bu yöntemler kısaca aşağıdaki gibidir, araştırmacı problemine göre en uygun yöntemi kullanmaya çalışır (Kantardzic, 2020):

Yeniden ikame yöntemi, bu yöntemde hem eğitim hem de test veri setinde aynı veri seti kullanılır. Nadir olarak kullanılsa da küçük örneklemler olduğunda kullanılır, hata çok fazladır.

Holdout, eğitim ve test veri seti farklı kümelerden oluşur. Örneklem sayısı az olduğunda ve eğitim ve test veri seti çalışmanın başında ayrıldığından hata oranı artabilir. Farklı ayrılmalar farklı tahmin sonuçları üretebilir.

Leave-one-out, bir model eğitim setinde (n-1) örneklemle eğititilir ve kalan örneklemde test edilir. Bu yöntem kullanıldığında, (n-1) örneklem sayısına sahip eğitim veri seti n kez tekrarlanır. Bu yaklaşımda n farklı model bulunduğundan, hesaplaması, tasarlanması ve karşılaştırılması zordur.

K- fold Cross Validation (Çapraz geçerleme), bu yöntem holdout ve leave-one out yönemlerinin birleştirilmesidir. Eğitim veri seti k parçaya bölünür ve her defa bir parça dışarda bırakılarak model eğitilir. Sonuçta k tane hata oranı bulunur ve bunların ortalaması alınarak nihai tahmin hatası bulunur. Küçük örneklemler için de iyi bir tahmim aracıdır.

Bootstrap, bu yöntemde veri tekrarlı olarak yeniden örneklenir. Böylelikle yüzlerce yeni veri seti oluşur. Deneysel araştırmalarda çapraz geçerlemeden daha iyi sonuç verdiği bulunmuştur. Özellikle küçük örneklem sayılarında kullanılması yararlıdır.

2.7. İNFERTILITE VE TUP BEBEK TEDAVİSİ

Günümüzün önemli sağlık sorunlarından biri olan infertilite, Dünya Sağlık örgütü tanımına göre ‘korunmasız olarak 1 yıl cinsel ilişkide bulunulmasına rağmen çocuk sahibi olunamaması’ olarak tanımlanmaktadır(Zegers-Hochschild et al., 2009). Evlenme ve kariyer planları nedeniyle çocuk sahibi olma yaşının artması, ekonomik ve sosyal sorunlar, sedanter yaşam ve çevresel nedenler gibi birçok farklı nedenden dolayı infertilite oranları, hem ülkemizde hem de dünyada her geçen yıl hızla artmaktadır. Bu durumdaki çiftlerin tedavisi için farklı yöntemleri içeren ‘yardımla üreme teknikleri’ ne başvurulabilmektedir. Bu yöntemlerin arasında başarı şansı olarak en yüksek ve en etkili yöntem, genel olarak ‘tüp bebek yöntemi’ olarak bilinen ‘İn vitro fertilizasyon’ yöntemidir. Bu yöntem hem kadın hem de erkek infertilitesinin tedavisinde etkili olarak kullanılabilmekle birlikte, her tedavi siklusu için başarı şansı ~%40 olup, başarısızlık ile sonuçlanan sikluslarda izlenebilecek tek yol tekrar denemekten öteye gidememektedir. Yapılan her yeni denemede hem maliyet, tedavi riskleri ve ilaçların yan etkilerinden kaynaklanan riskler artmakta, hem zaman kaybı yaşanmakta, hem de psikolojik, ailevi ve sosyal sorunlar meydana gelmektedir. Ayrıca devletin sosyal güvenlik politikaları kapsamında maliyetini karşıladığı bir tedavi yöntemi olması nedeniyle, devlete maliyetin

artması ile birlikte, tedavi maliyeti kapsanan siklus sayıları sınırlandırılmak zorunda kalınmaktadır. Tüm bu boyutlar birlikte değerlendirildiğinde IVF siklusundaki başarı şansının artırılması, büyük önem arz etmektedir. Başarı şansını etkileyen parametreler arasında hastaya ve eşine ait değişkenler, infertilite sürecine ait değişkenler, tedavi siklusuna ait değişkenler, hormonal durum, embriyoya ait değişkenler, endometrial durum sıralanabilir. Hangi faktörün ne kadar etkili olduğu ve sonuçları predikte etmedeki etkisi konusu tam olarak bilinmemekte ve tartışmalar devam etmektedir.

Hangi faktörlerin ne kadar etkili olduğunun bilinmesi, hastanın doğru bilgilendirilmesi, klinik ve laboratuvar kararlarının doğru verilmesini sağlayacağı gibi, bu faktörlerin desteklenmesi için stratejilerin geliştirilmesi ve ileri araştırmaların planlanması yoluyla başarının artmasına önemli katkı sağlayacaktır.

2.7.1. Değerlendirilen parametreler:

2.7.2. Embriyonun Morfolojik Özellikleri

Mikroskop altında embriyoların morfolojik özelliklerine bakarak implantasyon potansiyelinin belirlenmesi, rutin IVF laboratuvar uygulamalarında en çok kullanılan seçim kriteridir. Morfolojik olarak embriyonun birçok parametresinin embriyo canlılığı üzerine etkili olduğu düşünülmektedir. Bu parametreler, Pronükleer değerlendirme (pronukleus varlığı ve zamanlaması), Bölünme evresi değerlendirmesi (Bölünme hızı, Blastomer boyutu, Fragmentasyon derecesi, Blastomer multinükleasyonu, Sitoplazmik görüntü, Perivitellin alan ve zona pellusida özellikleri) ve Blastokist evresi değerlendirmesi (Blastosöl büyüklüğü, İç hücre kitlesi, Trofektoderm) olarak sıralanabilir.

Pronükleer değerlendirme embriyonun zigot dönemi değerlendirmesi olup bu evrede değerlendirilen parametrelerin, ileri embriyo gelişimi ve embriyonun implantasyon potansiyeli ile ilgili önemli bilgi sağladığı bilinmektedir (Scott & Smith, 1998; Tesarik & Greco, 1999; Van Blerkom, Bell, & Weipz, 1990). Yumurta sitoplazmasında (Ooplazmada) 2 pronükleusun izlenmesi, oositin fertilize olduğu anlamına gelir. (Şekil-2-3)

Şekil 2-9: Ooplazmasında 2 pronükleus izlenen fertilize olmuş bir oosit.

Bölünme evresi değerlendirmesinde, erken bölünme varlığı, bölünme hızı, blastomer boyutu, fragmentasyon oranı, blastomerlerin nükleer durumu, sitoplazmik görüntü, perivitellin alan ve zona pellusida özellikleri değerlendirilmektedir.

Literatürde birçok farklı bölünme evresi embriyo skorlama sistemi bildirilmiştir. Bölünme evresinde embriyolar 2. Günde 2-4 hücre, 3. Günde 6-8 hücre, 4. Günde kompaksiyon ve 5. Günde blastosist oluşumunun gözlenmesi beklenir.

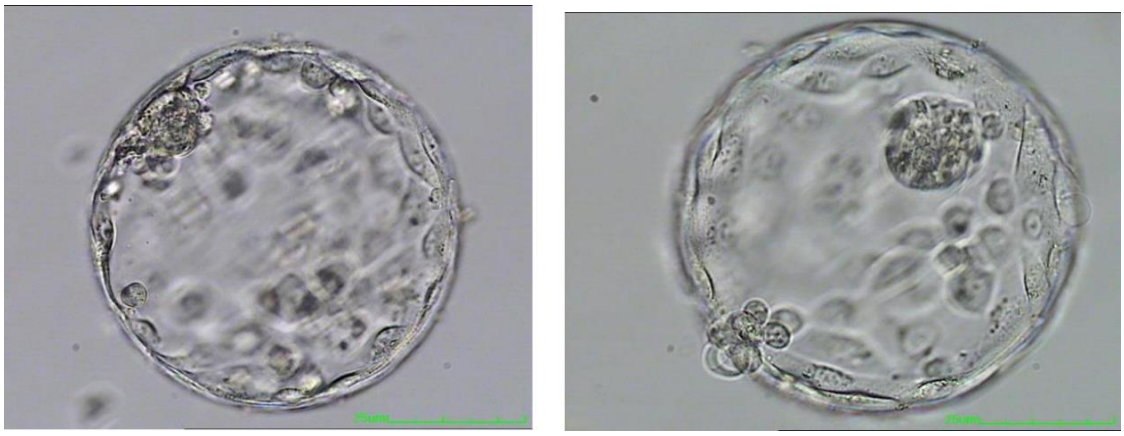
Şekil 2-10: Bölünme evresinde 2, 4 ve 8 hücreli bir embriyo

Yapılan çalışmalar, çok yavaş ya da çok hızlı embriyo gelişiminin embriyo implantasyon oranlarını olumsuz yönde etkilediğini göstermişlerdir(Cummins et al., 1986; Desai, 2002; Lewin et al., 1994; Ziebe et al., 1997). Bu çalışmaların tümünde 2. günde 4-hücreli olan embriyoların transfer edilmesiyle, diğer blastomer sayılı

embriyolara göre anlamlı derecede daha yüksek implantasyon oranları elde edildiği gösterilmiştir. Aynı şekilde 3. gün optimal bölünme hızı izlenen embriyolar transfer edildiğinde çok yüksek implantasyon oranları elde edildiği bildirilmiştir(Gerris & Evers, 1999; E. Van Royen et al., 2001; Eric Van Royen et al., 1999).

Bölünme hızı ve blastomer boyutuyla birlikte embriyo fragmentasyonu da embriyo seçim kriterleri arasına yerini almaktadır(Roseboom, Vermeiden, Schoute, Lens, & Schats, 1995). Nükleusu olmayan fragmentlerin sayısı ve miktarının embriyo hacmine oranı, embriyo kalitesi belirlenirken değerlendirilmektedir. Hücre fragmentasyonunun nedeni ve embriyo gelişimine etkisi tam olarak bilinmemektedir fakat hem embriyo hem de hastaya spesifik bir durum olarak ortaya çıkabilmektedir (Rienzi et al., 2005). Fragmentasyonun, apoptoz ya da programlı hücre ölümünden kaynaklandığı en yaygın görüştür.

Bölünme evresindeki embriyolarda olduğu gibi blastokist değerlendirme ve seçiminde de zaman ve morfoloji çok büyük öneme sahiptir(Balaban et al., 2000; Dokras, Sargent, & Barlow, 1993; Gardner, Lane, Stevens, Schlenker, & Schoolcraft, 2000; Richter, Harris, Daneshmand, & Shapiro, 2001). Blastokist skorlama sistemi, genişleme derecesi, iç hücre kitlesi ve trofektoderm hücrelerinin yoğunluğu esas alınmaktadır. İyi kaliteli bir blastokistin, inseminasyondan yaklaşık 112-114 saat sonra belirli ve geniş bir blastosöle (embriyo hacminin en az yarısı kadar), kavite içerisine doğru uzanan belirli bir iç hücre kitlesine ve birbirine bağlı bir epitel tabakası oluşturacak şekilde düzgün yerleşmiş trofektoderm hücrelerine sahip olması gerekmektedir(Gardner & Schoolcraft, 1999). Gardner ve ark., blastokist skorunun implantasyon ve gebelik oranlarını olumlu yönde etkilediğini bildirmişlerdir (Gardner et al., 2000).



(a) (b)

Şekil 2-11 (a) Genişlemiş blastokist, (b) Hatching başlamış bir blastokist

2.7.3. Embriyo Transfer Günü

IVF ya da ICSI sonrası embriyoların transferi, pn evresinde, bölünme evresinde ve blastokist evresinde yani 1-6. günlerde yapılabilmektedir. Embriyoların hangi gün ve hangi evrede transferinin optimum sonucu vereceği konusunda tartışmalar hala devam etmektedir.

Pn evresinde zigot transferi, bu konuda yasa ile kısıtlama uygulanan ülkelerde uygulanmakta ve zigotlar pronükleer morfoloji kriterlerine göre seçilerek transfer edilmektedirler. Ancak bu evrede embriyoların ileri gelişimleri ve implantasyon potansiyelleri ile ilgili sınırlı bilgi sahibi olunabildiğinden, bu konuda kısıtlama getirilmemiş ülkelerde pn evresinde transfer tercih edilmemektedir. Rutin laboratuvar uygulamalarında en sık 2., 3. ve 5. günlerde embriyo transferi tercih edilmektedir.

2.7.4. Estradiol seviyesi

Hcg enjeksiyonu günü, kandaki estradiol seviyesi de araştırılan bir başka parametredir. E2 seviyesinin yüksek ya da düşük olmasının siklus sonuçlarını etkilediğine dair bulgular bulunmakla birlikte klinik önemi tam olarak bilinmemektedir. E2 dolaylı olarak ovaryan cevap ile bağlantılı olduğundan siklus başarısı için önemli bir belirteç olarak nitelendirilebileceği, bununla birlikte yüksek E2 seviyesinde endometrial problemlere neden olarak implantasyon başarısızlığına yol açabileceği bilinmektedir.

2.7.5. Önceki deneme sayısı

Çiftlerin daha önce IVF denemesi olup olmadığı, var ise kaç denemesinin olduğu da klinik sonuçlar açısından anlamlı olabilmektedir. Önceki denemelerdeki tedavi sonucu da diğer denemelerin başarı şansı için belirleyici olabilmektedir.

2.7.6. Toplam gonadotropin dozu

IVF tedavisi sırasında oosit gelişimini tetiklemek amacıyla kullanılan hormon ilaçları olan gonadotropinlerin hangi süreyle ve dozda kullanıldığı da araştırılan bir başka parametredir. Bu parametre de yine ovaryan cevap ile ilgili dolaylı bilgi vereceğinden dolayı, sonuçlar üzerine etki etmesi beklenen önemli göstergelerden biri olabilmektedir

2.7.7. Toplam folikül sayısı

Toplam folikül sayısı olarak belirtilen gelişen folikül sayısı da oosit sayısı ile direk bağlantılı olan ve sonuçları etkilediği bilinen bir parametredir. Folikül sayısının optimum seviyede olması sonuçları olumlu yönde etkilemekle birlikte, bu seviyenin altında ya da üzerinde olması da sonuçları olumsuz etkilemektedir.

2.7.8. Stimülasyon süresi

Kontrollü ovaryan stimülasyon olarak bilinen ve oositlerin gelişimini tetiklemek için uygulanan tedavinin süresi de bir diğer önemli parametredir. Bu parametrenin, folikül gelişimi ve implantasyon potansiyelini etkilediği bilinmekle birlikte ne derecede başarıyı etkilediği bilinmemektedir.

3. GEREÇ VE YÖNTEM

Bu bölümde, araştırmanın modeli, örneklem, verileri, veri toplama araçları ve analizi ile ilgili bilgiler yer almaktadır.

3.1. Araştırmanın Amacı

Araştırmanın temel amacı, infertil bireylerden elde edilen embriyolardan elde edilen demografik, COS, sperm, embriyoya ait değişkenlerle, embriyonun tutma olasılığını makine öğrenmesi (machine learning, ML) yöntemleri ile tahmin etmektir. Çalışma, başarı (1) ve başarısızlığı (0) tahmin etmeyi amaçladığı için bir sınıflandırma problemidir. Araştırmada kullanılan modeller içerisinde en başarılı olan model bulunarak, embriyonun implamante etmesinde değişkenlerin önem sırası belirlenecektir.

Literatürde pek çok ML algoritması vardır ve hangisinin kullanılacağına karar vermek uzmanlık ve çok sayıda deneme gerektirmektedir. Bu çalışmada da, sınıflandırma problemini bir topluluk öğrenmesi yöntemi olan Super Learner (SL) algoritması ile çözülmesi çalışmanın birincil amacıdır. Bu algoritma, pek çok ML algoritmasını aynı anda kullanarak var olan yöntemlerden daha yüksek performanslı tahminler yapmaktadır. Çalışmada, sınıflandırma probleminin çözümünde, Lojistik Regresyon, Naive Bayes, Rastgele Orman, Destek Vektör Makineleri, Bagging, Gradient Boosting ve Xgboosting algoritmaları ile problem çözülecek, ardından SuperLearner yöntemi kullanılarak tahmin yapılmıştır.

3.2. Araştırmanın Türü

Çalışma, Mart 2014-Mart 2018 tarihleri arasında bir tüp bebek merkezine başvuran hastalar üzerinde yapılmış retrospektif kohort bir araştırmadır.

3.3. Araştırmanın Yeri ve Zamanı

Araştırma, Mart 2014-Mart 2018 tarihleri arasında İstanbul Çamlıca Medicana Hastanesi tüp bebek merkezine başvuran kişilerle yapılmıştır.

3.4. Araştırmanın Örneklemi

Çalışmada kullanılan veri seti bu dönemde başvuran 1146 hastadan transfer edilen 1952 embryodan oluşmaktaydı. Fakat, bazı hastalarda çok fazla sayıda boş veri olduğu

için transfer edilen 939 (%48,10) embryo ile çalışma yapıldı, 1013 (% 51,90) embriyo verisi kullanılamadı.

Tedavinin başarı oranının arttırılması için hastalara birden fazla embriyo transferi yapılabilmektedir. Daha önce yapılan çalışmalarda embriyoların implamantasyon olasılıklarının birbirinden bağımsız olduğu bulunmuştur (Speirs, Baker, & Abdullah, 1996; Trimarchi, 2001; Uyar et al., 2015). Bu nedenle aynı anda transfer edilen embriyoların tutunma olasılıkları farklı olduğu ve birbirini etkilemedikleri sonucuna varılmaktadır.

3.5. Araştırmanın Varsayımları

Araştırmada kullanılan verilerin doğru olduğu varsayılmıştır.

3.6. Araştırmanın Kısıtlılıkları

Çalışmanın en büyük sınırlılığı tek merkezli bir çalışma olması ve örneklem sayısının bir genelleme yapılabilmesi için yetersizliğidir Tek merkezli olması öte yandan bir avantajdır çünkü hastalara yapılan müdahelenin standardizasyonu sağlanmıştır.. Fakat, model farklı merkezlerden gelen yeni bilgilerle sürekli kendini güncelleyerek yenilenebilir ve genelebilir duruma getirebilir ve performansı arttırılabilir.

Çalışmada, merkezin bütün değişkenleri içeren bir elektronik sağlık verisi bulunmadığı için veri hasta dosyalarından el ile girilmiştir. Böylece hem insan hatası oranı artmış hem de çalışmanın süresi uzamıştır. Bu nedenle de bazı önemli değişkenlere ya ulaşılamamış ya da bazı değişkenlerde boş değerler oluşmuştur.

Yapılan gelecek çalışmalarda tahmin oranının arttırılması için değişken sayısı arttırılabilir.

Çalışmada sadece dokuz ML yöntemi kullanılmıştır. Bunun yanında, kullanılan ML yöntemlerinin çok sayıda hiper parametre değeri vardır, çalışmada yapılan denemeler belirtilenlerle sınırlıdır.

3.7. Etik İzinler

Çalışma için etik kurul 14.06.2019 tarihinde Medipol Üniversitesi Girişimsel olmayan etik kurulundan 494 referans numarası ile alınmıştır. Ayrıca, tedavi için gelen kişilerden bilgilendirmiş onam formu alınmıştır.

3.8. Veri Toplama Araçları

Kuruluşun hasta bilgi formu Ek-1'de verilmiştir. Çalışmada, uzman görüşü alınarak ve hasta bilgi formunda eksik gözlem sayılarının %50'ın üzerinde olan değişkenler dışarda bırakılmıştır. Toplam araştırmada, anneden, babadan ve embriyodan gelen 17 değişken kullanılmıştır.

3.9. Verinin Toplanması

Veriler, kurumdan alınan hasta bilgi formlarından bilgisayara excel formatında kaydedilmiştir. Çalışmada kullanılan R ve R studio programları veriyi .csv formatında istediği için veri .csv formatına çevrilmiştir. Aşağıda embriyonun implamantasyonunda kullanılan bazı yöntemler ve aşamaları verilmiştir.

3.9.1. Ovülasyon indüksiyonu ve embriyo manipülasyonu

Tedaviye alınan tüm kadınlara uzun (21. gün) veya antagonist protokolleri kullanılarak ovulasyon indüksiyonu yapıldı. En az üç folikül ≥18 mm çapa ulaştığında ß-hCG (Ovitrelle, Merck Serono) uygulanmış ve 35.5 saat sonra oosit toplama işlemi gerçekleştirilmiştir. Tüm oositlere intrasitoplazmik sperm enjeksiyonu (ICSI) uygulanmıştır. Oositlerin ve embriyoların kültürü ve manipülasyonu için LifeGlobal (IVFOnline, ABD) kullanılmış ve Labotect C 200 inkübatöründe (Labotect GmbH, Almanya) kültüre edilmiştir.

Mikroenjeksiyon prosedüründen 24 saat sonra fertilizasyon kontrolü yapıldı ve transfer edilene kdar embriyolar takip edilerek skorlanmıştır. Aynı embriyolog ve klinisyen tüm IVF adımlarını gerçekleştirmiştir.

3.9.2. Embriyo transferi

Transfer edilecek embriyo sayısı yasal düzenlemelere göre belirlendi. Türkiye'deki mevzuata göre, embriyo kalitesi ne olursa olsun otuz beş yaşın altındaki hastaların birinci ve ikinci ART tedavi girişimlerinde bir embriyo transferine izin verilmektedir; Bu yaş grubunun üçüncü veya daha sonraki denemelerinde en fazla iki embriyo, 35 yaşın üzerindeki hastalara en fazla iki embriyo transfer edilebilir.

Embriyo transferleri 5. günde bir Wallace 1816N (Smiths Medical, ABD) kateteri ile abdominal ultrason rehberliğinde yapıldı. Serum ß-hCG konsantrasyonları embriyo transferinden 12 gün sonra ölçüldü ve> 50 mIU / ml pozitif gebelik olarak tanımlandı. Gebelik kesesinin gözlenmesi klinik gebelik ve implantasyon olarak tanımlandı. Bu değere göre araştırmanın bağımlı değişkeni (Başarı/Başarısız) belirlendi.

3.9.3. Estradiol ölçümü

HCG uygulamasının yapıldığı gün kan örnekleri alındı. Serum E2 konsantrasyonları (nmol / L) enzim immün tekniği (AXSYM; Abbott, Almanya) ile belirlendi.

3.9.4. Endometrium ölçümü

Endometriyum, ultrason ile hCG uygulamasının yapıldığı gün ölçüldü.

3.10. Veri Önişleme

İstatistiksel analizlerde olduğu gibi veri madenciliği süreçlerinde de mutlaka verilerin analize hazırlanması gerekmektedir. Veri madenciliğinde bu süreçler veri temizleme, veri entegrasyonu (birleştirme), veri dönüşümü ve veri indirgeme süreçlerinden oluşmaktadır(Han et al., 2012). Veri indirgeme ve veri entegrasyonları bu çalışmada kullanılmamıştır.

3.10.1. Değişken Seçimi

Araştırmada kullanılan değişkenler için uzman görüşlerine başvurulmuş ve literatür çalışması yapılmıştır. Sonuç olarak, eksik gözlemler sorunu olmayan, herhangi bir değişken dışarıda bırakılmamıştır. ART tedavi sonucunu modelleyen çalışmalarda genellikle az sayıda değişkene odaklanılmıştır (Raef et al., 2020) . Bu çalışmada tedavi sonucunu etkileyen demografik, kontrollü overyan hiperstimülasyon verisi, embriyo ve sperm verisi olmak üzere 17 değişken kullanılmıştır (Tablo 3.1).

Raef vd. (Raef & Ferdousi, 2019) tarafından, 2019 yılında ART başarısında makine öğrenmesi yöntemleri üzerinde yapılan bir derleme makalesinde, 1997-2018 yılları arasında yayınlanmış 20 çalışma incelenmiştir. Bu çalışmalardan, dokuz tanesinde annenin yaşı, kadının yaşı ve yaş (erkek ve kadının yaşı) değişkenler en önemli değişkenler olarak belirlenmiştir. İki çalışmada FSH dozu, 1 çalışmada embriyoda oluşan hücre sayısı, 1 çalışmada infertilitenin süresi, 1 çalışmada kadında açıklanamayan

infertilite, 1 çalışmada laparoskopik ameliyat, 1 çalışmada ise infertilitenin ilk mi ikinci gebelikte olduğudur.

Araştırmamızın değişkenler arasında anne yaşı, infertilite nedeni, infertilitenin kaçıncı gebelikte oluştuğu yer almaktadır

3.10.2. Veri Temizleme

Bu aşamada veride bulunun eksik gözlemler ve uç değerler gözden geçirilir ve uygun yöntemler kullanılarak düzeltilir(Han et al., 2012).

3.10.2.1. Eksik Gözlemler

Araştırmanın değişkenlerinin türü ve boş değer oranları Tablo-1'de verilmiştir. Literatürde değişkendeki boş değer oranı % 50'nin üzerinde olduğunda, boş değer doldurma yöntemlerini kullanmanın doğru olmadığı belirtilmektedir. Çalışmada kullanılan değişkenlerdeki boş değer oranlarına bakıldığında oranın %0,32 ile % 47,84 arasında değişmektedir. Eğer bir değişkenin eksik veri oranı %12'nin altındaysa, normal dağılıma uygun ise değişken ortalama değilse medyam ile eğer kategorikse moduyla doldurulmuştur (Hachesu, Ahmadi, Alizadeh, & Sadoughi, 2013). Daha yüksek oranda uç değere sahip olan değişkenler ise IBM SPSS 23.0 'de Data Imputation kullanılarak doldurulmuştur. Bu yöntemde, boş değer içeren değişken bağımlı değişken olup, diğer araştırma değişkenleri kullanılarak lojistik ve multi-nominal regresyon analizi kullanılarak doldurulmuştur.

Tablo 3-1: Değişken türleri ve Değikenlerde bulunan Boş Değerler

Değişkenler	**Değişken türü**	**Boş Değerler**
Yaş	Sayısal	-
Beden Kitle Indeksi (BKI)	Sayısal	-
Infertilite Tipi (Birinci-İkinci)	Nominal	267 (% 39,73)
İnferfertilite Nedeni	Nominal	288 (%47,84)
Toplam gonatropin dozu	Sayısal	36 (%3,99)
Simülasyon Süresi (gün)	Sayısal	32 (% 3,53)

E2	Sayısal	15 (% 1,62)
Toplam Yumurta Sayısı	Sayısal	-
Olgun Yumurta sayısı (M2)	Sayısal	-
HCG günündeki Toplam Folikül Sayısı	Sayısal	3 (% 0,32)
HCG günündeki Endemetrium Kalınlığı	Sayısal	8 (%0,86)
Deneme Sayısı	Ordinal	34 (% 3,76)
Transfer günündeki embriyo kalitesi	Ordinal	-
Embriyo Transfer günü	Nominal	-
Toplam Sperm Sayısı	Ordinal	31 (% 3,41)
Sperm Hareketliliği	Ordinal	41 (%4,57)
Sperm Progresi	Ordinal	65 (% 7,44)

3.10.2.2. Aykırı değerler ve Veri Dönüşümü

Araştırmada kullanılan, beden kitle indeksi, E2 değeri, endemetrium kalınlığı, toplam yumurta sayısı, olgun yumuta sayısı (m2) değişkenlerinde uç değerler gözlenmiştir. Bu nedenle bu değişkenlere ln dönüşümü uygulanmış ve değişkenlerdeki uç değer sorunu çözülmüştür. Ayrıca, spern sayısı, sperm hareketliliği ve sperm progresi Dünya Sağlık Örgütü (WHO)'nun belirlediği değerler çerçevesinde yeterli/yetersiz olmak üzere iki gruba ayrılmıştır.

Ayrıca, nominal ve ordinal ölçekli değişkenler dummy değişken haline dönüştürülmüştür. Modele bu şekilde dahil edilmiştir. Çalışmada hem sürekli hem kategorik değişkenler bir arada bulunmaktadır. Sonuçların güvenilirliği açısından, modeller hem sürekli ve kategorik değişkenlerin bir arada kullanıldığı Model-1, bütün değişkenlerin kategorik dummy değişkenlere dönüştürüldüğü Model-2 olmak üzere ayrı ayrı modeller kurulup karşılaştırılmıştır. Model-1'de daha hassas tahminlerin yapıldığı

tahmin edilmekle birlikte, bazı algoritmaların dummy değişkenlerle tahmin gücünün daha yüksek olduğu bilinmektedir. Bu nedenle her iki model karşılaştırılmıştır.

3.11. Modelleme Aşaması

Veri madenciliği ve makine öğrenmesi yöntemlerinde temel amaç veride yer alan gizli örüntülerin en doğru şekilde belirlenmesidir. Bu nedenle çeşitli yöntemler geliştirilmiştir. Bunlardan ilki hold-out yöntemidir. Bu yönteme göre veri seti eğitim ve test veri seti olarak ikiye ayrılmakta, model eğitim veri seti için geliştirilmekte, test setinde ise bu modelin doğruluğu test edilmektedir. Bir diğer yöntem ise çapraz geçerleme (cross-validaton) yöntemleridir. Bu yöntemlerde ise eğitim kümesi k sayıda parçaya bölünür. Bir diğer yöntem ise veride sınıf dengesizliği sorunu olduğunda, eğitim veri setinde sınıfların dengelenmesi durumudur.

Bu çalışmada, öncelikle hold-out yöntemi ile veri seti %70-%30 oranlarında eğitim ve test veri seti olarak ayrılacaktır. Çapraz geçerleme yöntemi olarak ise 10 kat çapraz geçerleme 10 tekrar için kullanılacaktır.

Araştırmamızda sınıf dengesizliği problemi vardır. Embriyonun implante olmama (% 79,6) ve olma olasılığı (%20,4) oranlarında gözlenmiştir. Bu problemi giderebilmek için yine eğitim kümesinde up-sampling, down-sampling, ROSE ve SMOTE yöntemleri yine lojistik regresyon analizinde denenecek ve en uygun olan örnekleme yöntemi diğer modeller için de kullanılacaktır.

Embriyonun başarısının modellenmesi bir ikili sınıflandırma problemidir. Bu amaçla Lojistik Regresyon, Naive Bayes, Rastgele Orman, Destek Vektör Makinesi, Yapay Sinir Ağları, Bagging, Gradient Boosting , Xgboosting ve Super Learner yöntemleri ile tahnin edilecektir.

Modellerin başarı oranlarının değerlendirilnesinde ise, özgüllük, seçicilik, kesinlik, doğruluk, F değeri, Kappa istatistiği ve ROC eğrisi altında kalan alanlar değerlendirilecektir. Bunun yanında Super Learner algoritmasında Brier risk skorları da modellerinde değerlendirilmesinde kullanılacaktır.

3.12. Çalışmada Kullanılan Analiz Programı

Çalışmada kullanılan makine öğrenmesi yöntemlerinin uygulanabilmesi için R programı kullanılmıştır. Bu program, istatistiksel analizler için geliştirilen açık kaynak kodlu bir programdır. Kodların yazımında kolaylık olması açısından RStudio ortamı kullanılmıştır.

Bunun yanında, kayıp verilerin doldurulmasında IBM SPSS 23.0 programı kullanılmıştır.

4. BULGULAR

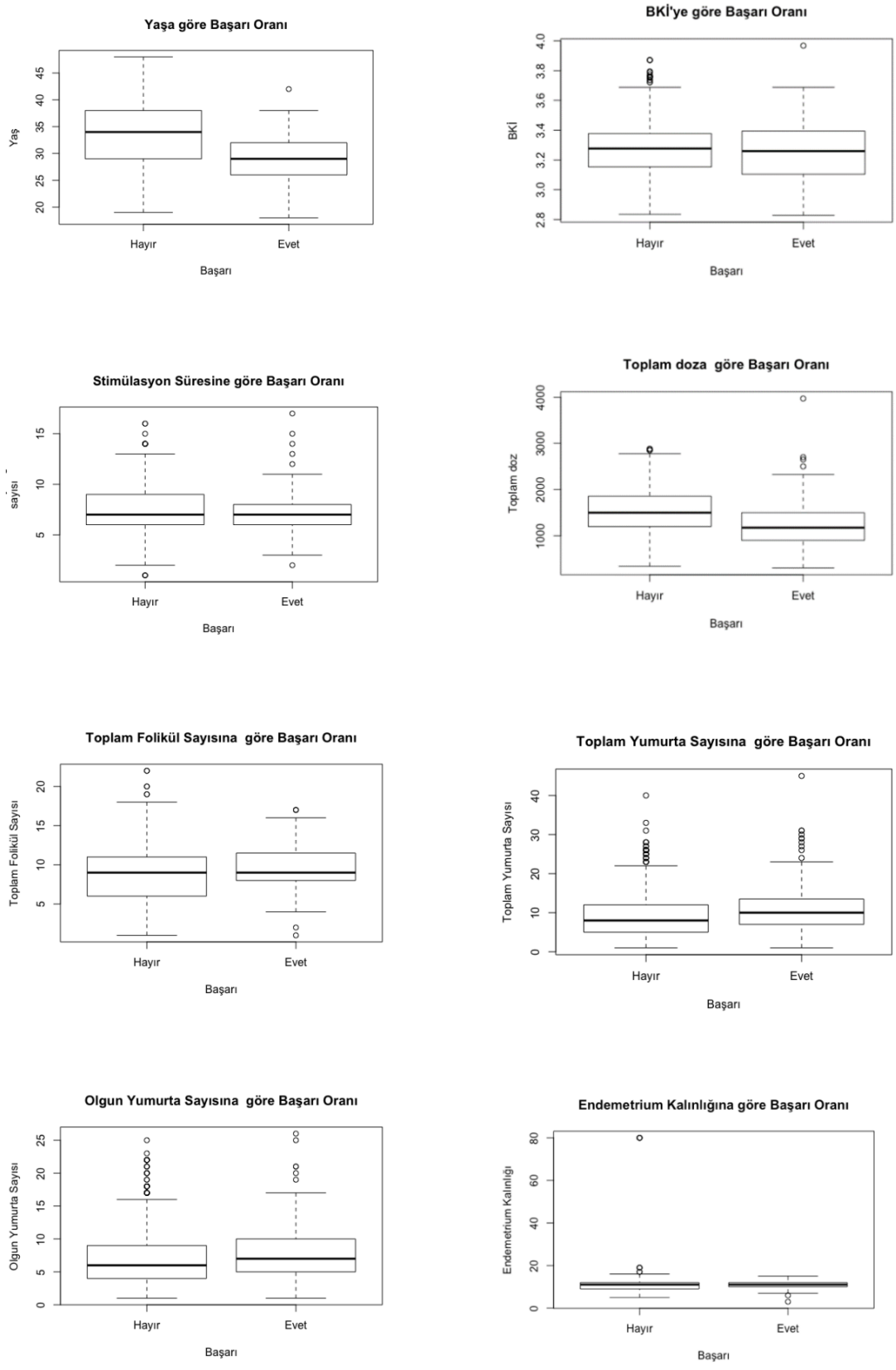
Değişkenler arasındaki korelasyonlar Ek-1'de verilmiştir. Korelasyonlar hesaplanırlen, sayısal değişkenler arasındaki korelasyon pearson, iki kategorik değişken arasımda phi korelasyon katsayısı, bir nümerik, bir kategorik değişken arasında point biserial korelasyon uygulanmıştır. Yapılan önemlilik testlerinde etki büyüklüklerine göre annenin yaşı, geçirilen hamilelik sayısı, embriyo transfer günü ve toplam tonadotropin dozu embriyo başarısının modellenmesinde sırasıyla en önemli değişkenler olarak bulunmuştur. Tablo-1'de araştırmanın değişkenlerine ilişkin tanımlayıcı istatistikleri bulabilirsiniz. Çalışmanın bulgular kısmında, Embriyoda implamantasyonun olması başarı olup olmaması olarak isimlendirilmiştir. Şekil-4.1 ve Şekil-4.2'de değişkenlerin grafikleri bulunmaktadır.

Tablo 4-1: Değişkenlerin tanımlayıcı İstatistikleri

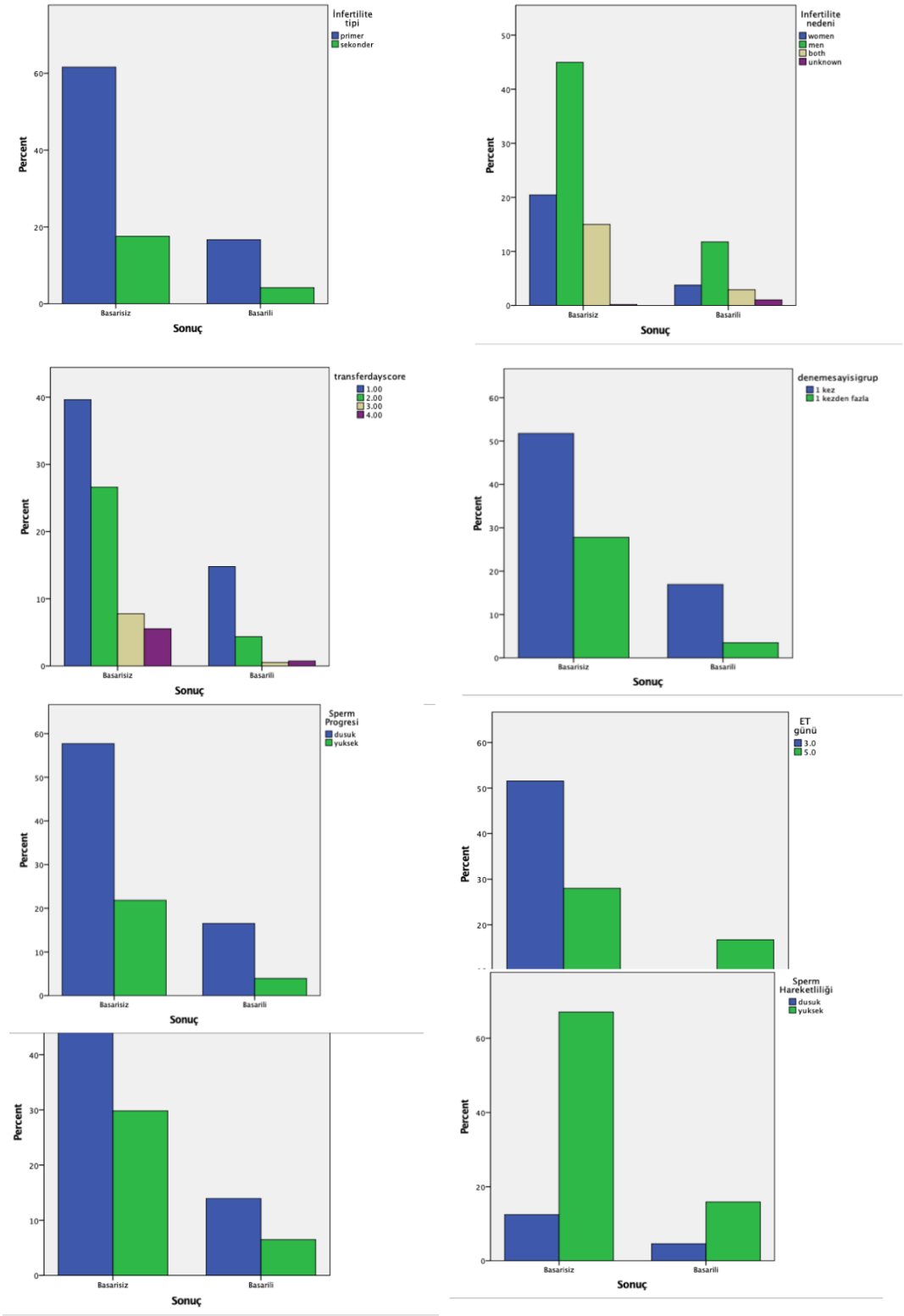
Değişken Türü	Değişken	Kategori	Implamantasyon var (n=747; 79,6%)	Implamantasyon yok (n=192;20,4 %)	Test İstatistiği	p	Etki Büyüklüğü
Demografik Veri	Yaş (age)		28,79 (4.05)	33,65 (5,96)	13,329	p<0,001	0,541
	Beden Kitle İndeksi (bmi)		26,38 (5,37)	26,96 (4,97)	1,736	0,083	0,057
	Hamilelik	İlk (primary)	152 (79,2%)	571 (76,4%)	0,642	0,444	0,423
		İkinci (secondary)	40 (20,8%)	176 (23,6%)			
	İnfertilite nedeni	Kadın (woman)	36 (18,8%)	204 (27,3%)	23,517	p<0,001	0,158
		Erkek (men)	114 (59,4%)	394 (52,7 %)			
		Hem kadın hem erkek (both)	32 (16,7%)	143 (19,1%)			
		Bilinmeyen (unknown)	10 (5,2%)	6 (0,8 %)			
Kontrollü Overyan Hiperstimülasyon Verisi	Toplam Gonadotropin dozu (tdose)		1225,66 (505,32)	1530,40 (498,41)	7,474	p<0,001	0,237
	Stimülasyon gün sayısı (days)		7,34 (2,06)	7,48 (2,04)	0,848	0,397	0,028
	E2 Level (lnetwo)		976,40 (1212,75)	1260,76 (1343,36)	2,93	0,005**	0,163
	Toplam yumurta sayısı (total)		11,31 (6,35)	9,58 (5,87)	3,57	p<0,001	0,116

	Olgun yumurta sayısı (mtwo)		8,23 (4,24)	7,37 (6,57)	4,164	p<0,001	0,221
	HCG günündeki Toplam Folikül Sayısı (tfolcount)		9,55 (2,97)	8,80 (3,46)	2,992	0,003**	0,16
	HCG günündeki Endemetrium Kalınlığı (end)	end	10,86 (1,79)	10,93 (5,43)	2,032	0,042*	0,066
	Deneme Sayısı	İlk (first)	159 (82,8%)	486 (65,1%)	22,381	p<0,001	0,154
		Birden Fazla (mtonce)	33 (17,2 %)	261 (34,9%)			
Embryo Verisi	Transfer günündeki embriyo kalitesi	1 (tdsone)	139 (72,4%)	372 (49,8%)	33,683	p<0,001	0,189
		2 (tdstwo)	41 (21,4%)	250(33,5 %)			
		3 (tdsthree)	5 (2,6 %)	73 (9,8 %)			
		4 (thdsfour)	7 (3,6 %)	52 (7 %)			
	Embryo Transfer Günü	Third day (thirdday)	35 (18,2 %)	263 (35,2%)	133,955	p<0,001	0,378
		Fifth day (fifthday)	157 (81,8%)	484 (64,8%)			
Sperm Verisi	Toplam Sperm Sayısı	Düşük (sc_low)	131 (68,2%)	467 (62,5%)	2,155	0,142	0,042
		Normal (sc_normal)	61 (31,8 %)	280 (37,5 %)			
	Sperm Hareketliliği	Hareketsiz (immotile)	43 (22,4%)	117 (15,7 %)	4,899	0,027*	0,072
		Hareketli (motile)	149 (77,6%)	630 (84,3%)			
	Sperm Progresi	Var (prog_motile)	155 (80,7 %)	542 (72,6%)	5,332	0,021*	0,075
		Yok (pro_immotile)	37 (19,3%)	205 (27,4 %)			

*p<0,05 ** p<0,01 Not: Sayısal değişkenler ortalama (standart sapma) şeklinde gösterildi, karşılaştırılmasında t testi, kullanıldı. Kategorik değişkenler n(%)şeklinde gösterildi karşılaştırmalar için Chi-Square kullanıldı.



Şekil 4-1: Sayısal Değişkenlerin Kutu Diyagramları



Şekil 4-2: Kategorik Değişkenlerin Sütun Grafikleri

4.1. Lojistik Regresyon (LR) Analizi Bulguları

Aşağıda çalışmada kullanılan sınıf dengesizliği problemini gidermek için kullanılan Smote, Balanced Over, Balanced Under, Balanced Both ve ROSE yöntemlerinin eğitim veri setindeki dağılımları verilmiştir. Lojistik regresyon'da model-1 için bütün yöntemler için kullanılacak ve en doğru tahmini yapan örnekleme tekniği ile diğer modellerde de kullanılacaktır. Çalışmada R'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb]
& 1, 2020) paketi kullanılmıştır. Dengesiz veriler için kullanılan paketler Rose(Lunardon, Menardi, & Torelli, 2014), DMwr(Torgo, 2015)'dır. Tablo 4.2'de kullanılan sınıf dengesizliği algoritmalarının örneklem sayıları verilmiştir.

Tablo 4-2: Kullanılan Sınıf Dengesizliği Algoritmlarındaki Örneklem sayıları

Başarı	**Hayır**	**Evet**
Orijinal	523 (%79,6)	134 (%20,4)
Smote	268 (%50)	268 (%50)
Balanced Over	523 (% 52,3)	477 (% 47,7)
Balanced Under	166 (% 55,33)	134 (%54,67)
Balanced Both	789 (%52,6)	711 (% 47,4)
ROSE	342 (% 52,05)	315 (%47,95)

Tablo-4.3'de algoritmaların lojistik regresyon analizi sonucundaki test veri setinin sınıflandırma algoritması karşılaştırma metrikleri verilmiştir. Yapılan değerlendirme sonucunca, En yüksek sensivity, specificity, precision, accuracy değerleri Smote yönteminde, F ve AUROC ise balanced both veri setinde kullanılmıştır. Sınıf dengesizliği problemi olan durumlarda, F değerinin daha güvenilir bir ölçüt olduğu literatürde vurgulanmıştır. Smote ve Balanced Both yöntemlerinde F değeri birbirine çok yakındır. Bundan sonraki modellerde sınıf dengesizliği problemlerinin çözümünde Smote yöntemi kullanılacaktır. Tablo 4.3'de ise algoritmaların performansları karşılaştırılmıştır. En iyi performansa sahip olan algoritmanın SMOTE yöntemi olduğu görülmektedir. Diğer ML yöntemleri için de eğitim kümesinde SMOTE yöntemi kullanılacaktır.

Tablo 4-3: LR Model-1 Test Veri Seti Metrikleri

Örnekleme Yöntemeleri	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Orijinal	0,517	0,955	0,750	0,865	0,612	0,736	0,534
Smote	**0,793**	**0,83**	**0,547**	**0,822**	**0,647**	0,812	**0,635**
Balanced Over	0,793	0,79	0,494	0,791	0,608	0,792	0,477
Balanced Under	0,741	0,804	0,494	0,791	0,593	0,772	0,46
Balanced Both	0,81	0,821	0,54	0,819	**0,648**	**0,816**	0,533
ROSE	0,793	0,768	0,469	0,773	0,589	0,78	0,447

Buna göre araştırmanın değişkenlerinin önem derecesi Tablo-4.5'deki gibidir. Buna göre en önemli üç değişken transfer günü (5.gün) , infertilite nedeninin erkek olması ve embryo skorunun yüksek (1) olmasıdır. En önemsiz değişken ise toplam gonadotropin dozudur.

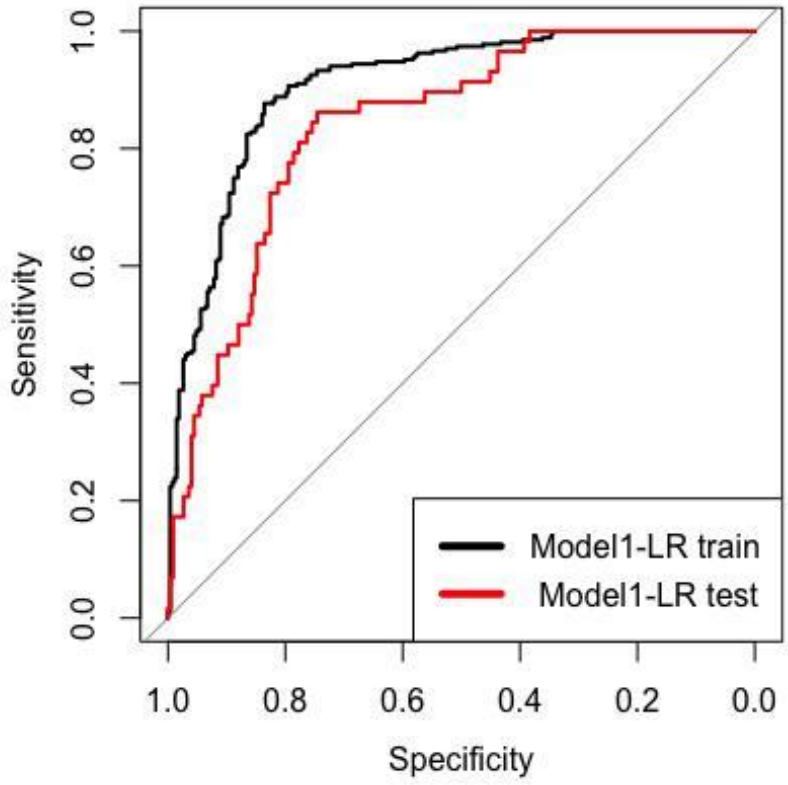
Tablo 4-4: LR Model-2 Performans Metrikleri

Model	LR	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-2	Train-10 fold cV	0,832	0,787	0,796	0,81	0,814	0,886	0,619
	Test	0,897	0,737	0,469	0,77	0,616	0,876	0,473

Model-2'de ise çapraz doğrulama ile eğitim veri setinde doğruluk % 76,5 iken, test veri setinde % 63,8 'e düşmüştür. F değeri ise % 77,3 iken % 47'ye düşmüştür. Sensitivity eğitim veri setinde % 79,9 iken, test veri setinde % 77,6'ya, duyarlılık ise % 79,9'den % 77,6' ya düşmüştür. Çalışmadan elde edilen metriklerin sonuçlarına göre, Model-1'in performansı Model-2'nin performansından iyidir.

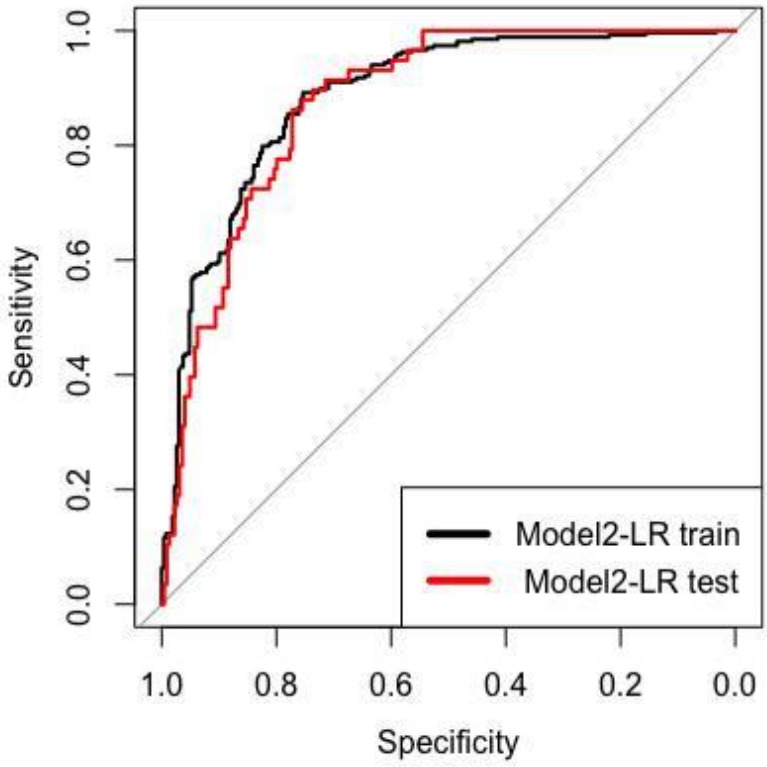
Tablo 4-5: LR Model-1 için Değişkenlerin Önem Derecesi

1	fifthday	6	women	11	first	16	bmi
2	men	7	tdsthree	12	mtwo	17	prog_motile
3	tdsone	8	motile	13	primary	18	sc_normal
4	both	9	tdstwo	14	age	19	stime
5	end	10	lnetwo	15	tfolcount	20	tdose



Şekil 4-3: Model-1 Lojistik Regresyon Eğitim ve Test Kümeleri Roc Eğrisi

Tablo 4.4'de Model-2 Lojstik Regresyon bulguları verilmiştir. Çapraz doğrulama ile eğitim veri setinde doğruluk % 81 iken, test veri setinde % 77 'ye düşmüştür. F değeri ise % 81,4 iken % 61,6'ya düşmüştür. Duyarlılık eğitim veri setinde % 83,2 iken, test veri setinde % 89,7'ye yükselmiş, seçicilik ise % 78,7'den % 73,7' ye düşmüştür. Çalışmadan elde edilen metriklerin sonuçlarına göre, Model-1'in performansı Model-2'nin performansından yüksektir.



Şekil 4-4: Model-2 Lojistik Regresyon ROC Eğrisi

4.2. Naïve Bayes Sınıflandırıcı Bulguları

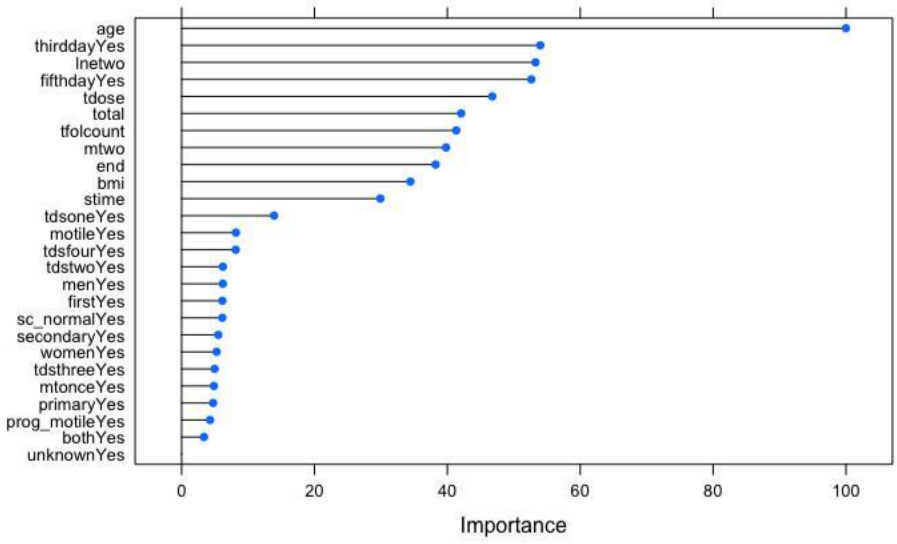
Naïve Bayes sınıflandırıcı uyguluması için R studio'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb] & 1, 2020) ve e1071(Hornik, Weingessel, Leisch, & Davidmeyerr-projectorg, 2020) paketleri kullanılmıştır. 10 kat çapraz geçerleme 10 tekrar için eğitim veri seti ve test veri setinin sonuçları Tablo-4.6'da verilmiştir.

Tablo 4-6: Naive Bayes Eğitim ve Test Veri seti Metrikleri

Modeller	Naive Bayes	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	0,855	0,791	0,804	0,823	0,823	0,823	0,646
	Test	0,741	0,741	0,426	0,741	0,541	0,741	0,379
Model-2	Train-10 fold cV	0,799	0,731	0,748	0,765	0,773	0,765	0,53
	Test	0,776	0,603	0,334	0,638	0,47	0,69	0,255

Sonuçlar incelendiğinde, Model-1'de çapraz doğrulama ile eğitim veri setinde doğruluk % 82,3 iken, test veri setinde % 74,1'e düşmüştür. F değeri ise % 82,3 iken % 54,1'e düşmüştür. Sensitivity eğitim veri setinde % 85,5 iken, test veri setinde % 74,1'e, duyarlılık ise % 79,1'den % 74,1' düşmüştür. Model-2'de ise çapraz doğrulama ile eğitim veri setinde doğruluk % 76,5 iken, test veri setinde % 63,8 'e düşmüştür. F değeri ise % 77,3 iken % 47'ye düşmüştür. Sensitivity eğitim veri setinde % 79,9 iken, test veri setinde % 77,6'ya, duyarlılık ise % 79,9'den % 77,6' ya düşmüştür. Çalışmadan elde edilen metriklerin sonuçlarına göre, Model-1'in performansı Model-2'nin performansından yüksektir.

Model-1'de değişkenlerin önem derecesi incelendiğinde (Şekil 4-5), en önemli değişkenin annenin yaşı olduğu görülmmektedir. Sonrasında ise sırasıyla, embriyo transfer günü ve E2 seviyesi gelmektedir.



Şekil 4-5: Model-1 NB Değişkenlerin Önem Derecesi

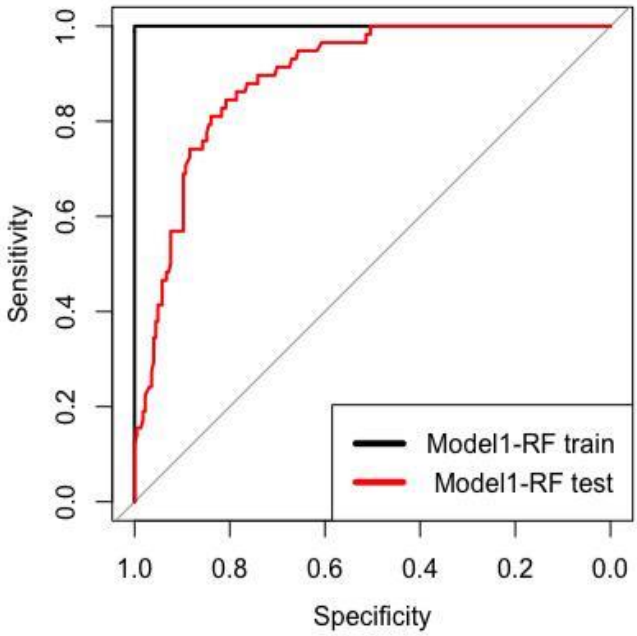
4.3. Rastgele Orman (Random Forest) Bulguları

Rasgele Orman (RO) sınıflandırıcı uyguluması için R studio'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb] & 1, 2020) ve randomForest(Zaremba & Smoleński, 2000) paketleri kullanılmıştır. Caret'de randomForest algoritasında iki temel parametre değerlerleri (tuning parametresi) vardır: mtry (Her bölünmede rastgele örneklenen değişken sayısı) ve ağaç sayısı. Çalışmada, her iki model için de 1'den 14'e kadar mtry, 100,500,1000,1500,2000, 2500 ağaç sayısı denenmiştir. En uygun performans Model-1 için 6 mtry, 500 ağaç, Model-2 için ise 4 mtry ve 500 ağaç en iyi performansı vermiştir. Bu modellerin 10 kat çapraz geçerleme 10 tekrar için eğitim veri seti ve test veri setinin sonuçları Tablo 4-7'de verilmiştir.

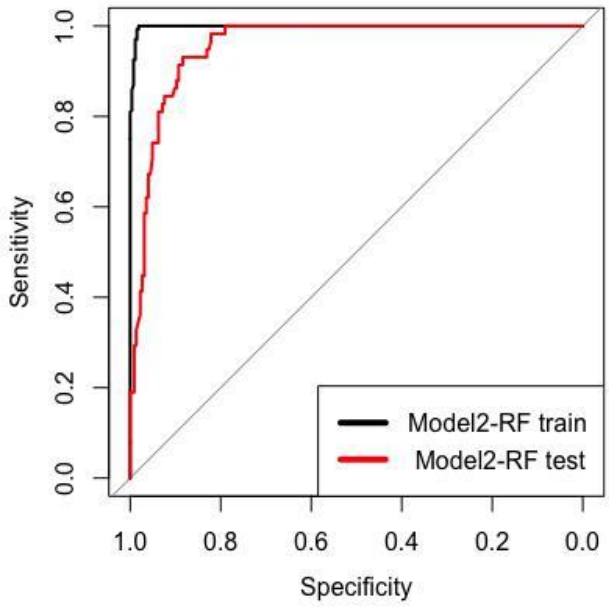
Tablo 4-7: RO Eğitim ve Test Veri seti Metrikleri

Modeller	Rastgele Orman	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold CV	1	1	1	1	1	1	1
	Test	0,81	0,839	0,566	0,833	0,666	0,825	0,56
Model-2	Train-10 fold CV	1	0,981	0,982	0,991	0,991	0,987	0,981
	Test	0,914	0,71	0,449	0,752	0,602	0,81	0,451

Ratgele Orman sonuçları incelendiğinde, Model-1'de eğitim veri setinde model tamamen doğru tahmin etmişken, test verisinde % 83,3 doğruluk, % 81 duyarlılık, % 66,6 F değerine sahiptir. Model-2'de ise eğitim veri setinde % 99,1 doğruluk, % 100 duyarlılık ve % 99,1 F değerine sahipken, test veri setinde % 75,2 doğruluk, % 91,4 duyarlılık ve F değeri % 60,2' ye düşmüştür. Model-1'in performansı yine Model-2'nin performansından daha iyidir.



Şekil 4-6: Model-1 Random Forest ROC Eğrisi



Şekil 4-7: Model-2 Random Forest ROC Eğrisi

4.4. Destek Vektör Makinesi (DVM) Bulguları

Dstek Vektör sınıflandırıcı uyguluması için R studio'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb] & 1, 2020) paketi kullanılmıştır. DVM algoritmasında grid yöntemi kullanılarak en uygun parametre değerleri bulunmaya çalışılmıştır. Radyal temelli çekirdek fonksiyonu kullanılmıştır. DVM'de parametre değerleri maliyet (cost) ve gamma (γ) değerleridir.

Tablo 4-8: DVM Eğitim ve Test Veri seti Metrikleri

Modeller	DVM	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	0,963	0,963	0,963	0,963	0,963	0,963	0,925
	Test	0,655	0,817	0,481	0,784	0,555	0,736	0,416
Model-2	Train-10 fold cV	0,914	0,634	0,714	0,774	0,802	0,774	0,549
	Test	0,983	0,477	0,328	0,582	0,492	0,73	0,265

Yapılan uygulamalar içerisinde, en yüksek performansı veren değerler Model-1 için cost=10 ve γ=1, Model-2 için ise cost=0,1 ve γ=1 için bulunmuştur. Bu modellerin 10 kat çapraz geçerleme 10 tekrar için eğitim veri seti ve test veri setinin sonuçları Tablo-4.8'de verilmiştir.

Model-1'de doğruluk % 96,3 iken eğitim veri setinde, test veri setinde % 78,4'de düşmüştür. F değeri % 96,3'den % 73,6'ya düşmüştür. Duyarlılık ise % 96,3'den % 65,5'e düşmüştür.

Model-2'de ise doğruluk % 77,4 iken eğitim veri setinde, test veri setinde % 58,2 'ye düşmüştür. F değeri % 80,2'den % 49,2'ye düşmüştür. Duyarlılık ise % 91,4'den % 98,3'e çıkmıştır. Model-1'in performansının, Model-2'den daha iyi olduğu görülmektedir.

4.5. Yapay Sinir Ağları (YSA) Bulguları

YSA uyguluması için R studio'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb] & 1, 2020) paketi kullanılmıştır. YSA algoritmasında grid yöntemi kullanılarak en uygun parametre

değerleri bulunmaya çalışılmıştır. YSA'nın iki temel parameter değerleri gizli katman sayısı (size) ve aşırı öğrenmeyi azaltmak için kullanılan eksilme (decay) değerleridir. Çalışmada, gizli katman sayısı için 1 ve 10 arasında, eksilme değerleri için ise 0,1-0,5 arasında 0,1 azaltma yaparak grid yöntemi kullanılarak denenmiştir. Maksimum yineleme sayısı 10.000 olarak belirlenmiştir. Sonuç olarak Model-1' de en iyi performansı 9 gizli katman ve 0,4 eksilme, Model-2'de ise gizli katman sayısı 10, eksilme değeri ise 0,5 bulunmuştur.

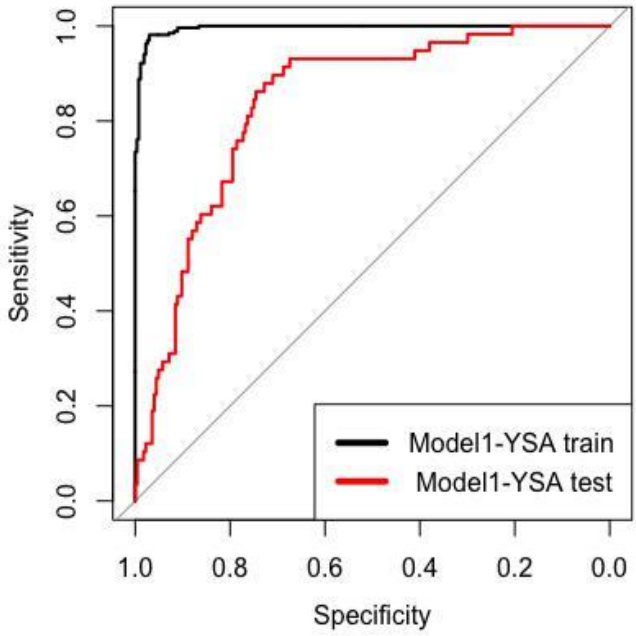
Tablo 4-9: YSA Eğitim ve Test Veri seti Metrikleri

Modeller	YSA	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	0,985	0,974	0,974	0,98	0,98	0,979	0,959
	Test	0,757	0,821	0,524	0,809	0,619	0,79	0,497
Model-2	Train-10 fold cV	0,978	0,944	0,946	0,961	0,962	0,961	0,922
	Test	0,862	0,705	0,431	0,738	0,575	0,784	0,414

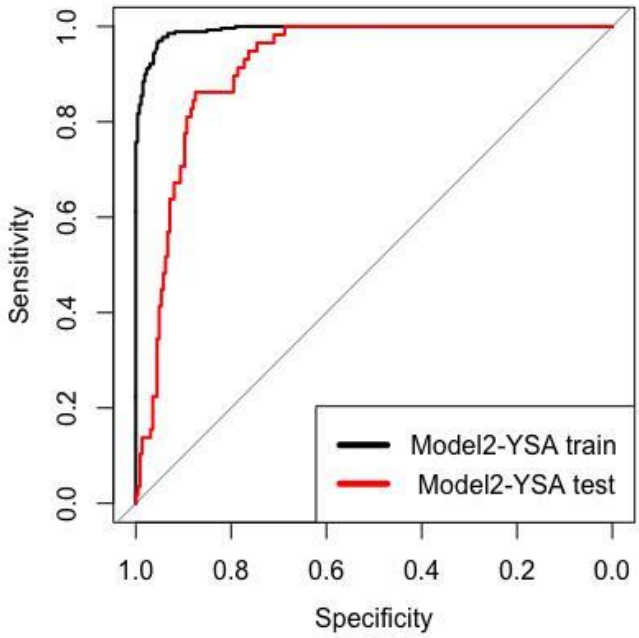
Model-1'de doğruluk % 98,5 iken eğitim veri setinde, test veri setinde % 75,7'ye düşmüştür. F değeri % 98'den % 61,9'a düşmüştür. Duyarlılık ise % 98,5'den % 75,7'ye düşmüştür (Tablo 4-9).

Model-2'de ise doğruluk % 96,1 iken eğitim veri setinde test veri setinde % 57,5 'e düşmüştür. F değeri % 96,2'den % 57,5'e düşmüştür. Duyarlılık ise % 97,8'den % 86,2'ye düşmüştür. Model-1'in performansının, Model-2'den daha iyi olduğu görülmektedir (Tablo 4-9).

.



Şekil 4-8: Model-1 Yapay Sinir Ağları ROC Eğrisi



Şekil 4-9: Model-2 Yapay Sinir Ağları ROC Eğrisi

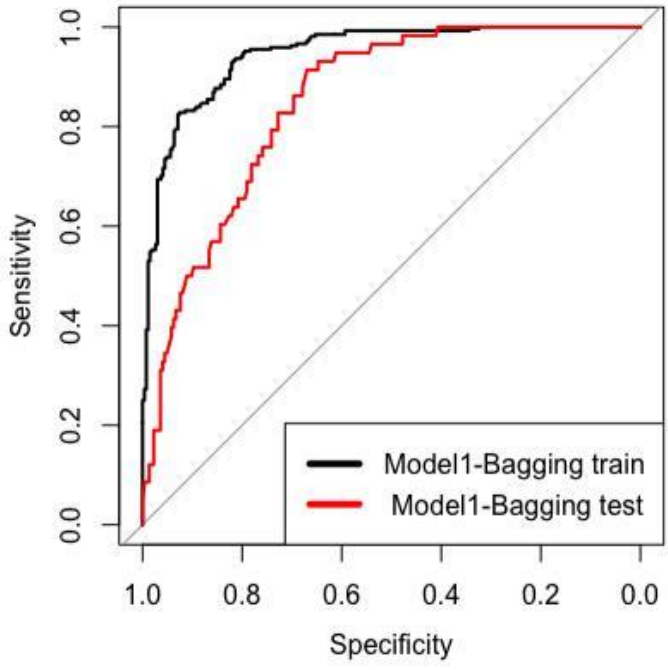
4.6. Bagging Bulguları

Bagging sınıflandırıcı uyguluması için R studio'da randomForest(Zaremba & Smoleński, 2000) paketi kullanılmıştır. Bagging RO'nun mtry= değişken sayısı olduğu özel bir halidir. Bu nedenle modelde, Model-1 için mtry=20, Model-2 için ise mtry=23 alınmıştır. Model metrikleri Tablo-4-10'daki gibidir.

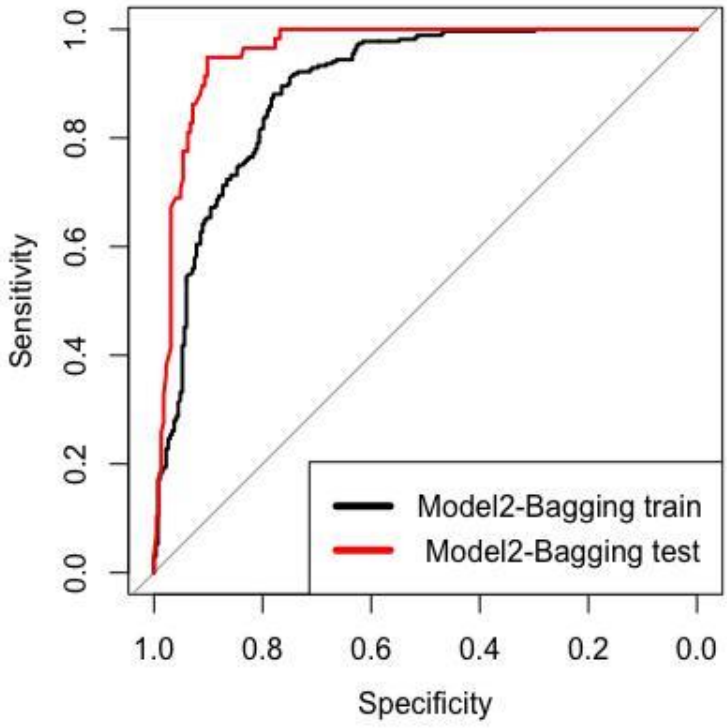
Tablo 4-10: Bagging Eğitim ve Test Veri seti Metrikleri

Modeller	Bagging	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	1	1	1	1	1	1	1
	Test	0,707	0,79	0,466	0,773	0,562	0,749	0,417
Model-2	Train-10 fold cV	1	0,996	0,996	0,998	0,998	0,998	0,996
	Test	0,862	0,688	0,417	0,723	0,562	0,775	0,394

Bagging sonuçları incelendiğinde, Model-1'de eğitim veri setinde model tamamen doğru tahmin etmişken, test verisinde % 70,7 doğruluk, % 79 duyarlılık, % 56,2 F değerine sahiptir. Model-2'de ise eğitim veri setinde % 99,8 doğruluk, % 100 duyarlılık ve % 99,8 F değerine sahipken, test veri setinde % 72,3 doğruluk, % 86,2 duyarlılık ve F değeri % 56,2' ye düşmüştür. Model-1'in performansı yine Model-2'nin performansından daha iyidir.



Şekil 4-10: Model-1 Bagging ROC Eğrisi



Şekil 4-11: Model-2 Bagging ROC Eğrisi

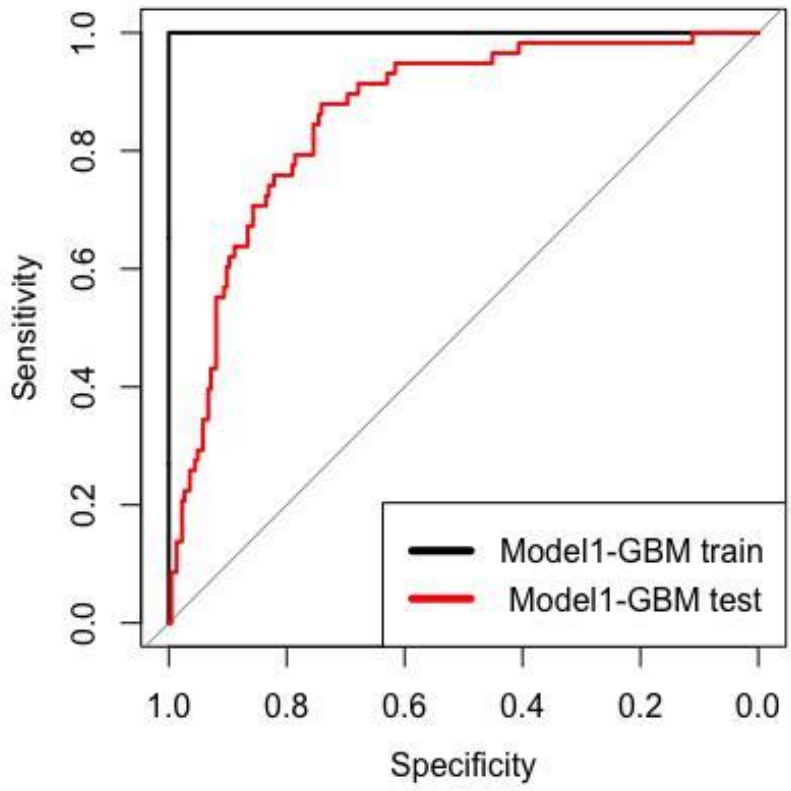
4.7. Gradyan Artırma Makineleri (Gradient Boosting Machines - GBM) Bulguları

GBM uyguluması için R studio'da caret(Max Kuhn [aut, cre], Jed Wing [ctb], Steve Weston [ctb], Andre Williams [ctb], Chris Keefer [ctb], Allan Engelhardt [ctb], Tony Cooper [ctb], Zachary Mayer [ctb], Brenton Kenkel [ctb], R Core Team [ctb] & 1, 2020) paketi kullanılmıştır. Model parametreleri ise grid yöntemi ile belirlenmiş, en uygun parametreler ise Model-1 için ağaç sayıları parametresi (.) 1100, yaprak sayıları (. h) 7, öğrenme oranı (shrinkage) 0,1, ağaçlardaki yaprak düğümlerinin sayısı (.) 6; Model-2 için se sırasıyla 2000;7;0,01;1 olarak bulunmuştur. Bulgular Tablo-4.11'da verilmiştir.

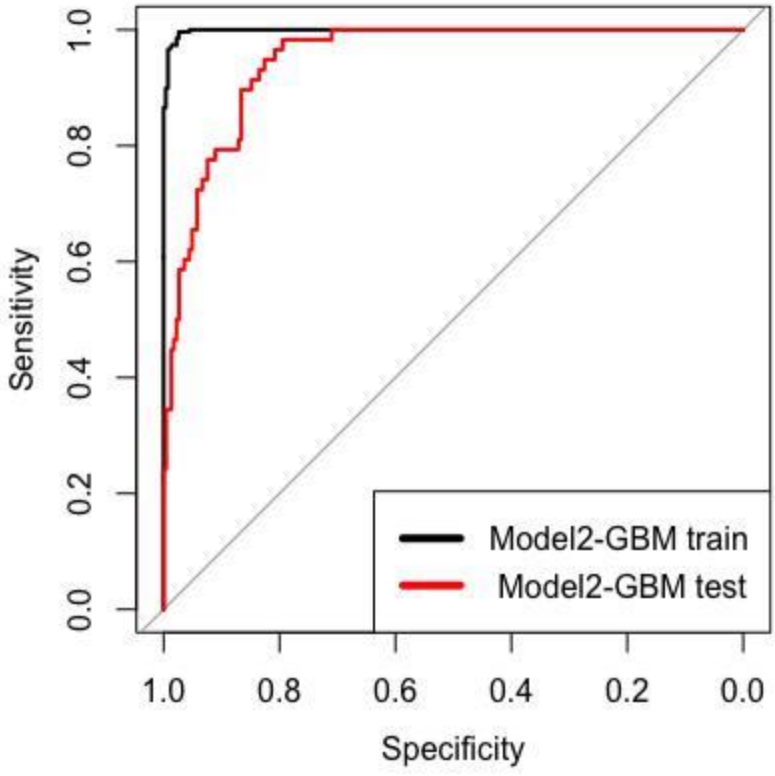
Tablo 4-11: GBM Eğitim ve Test Veri seti Metrikleri

Modeller	GBM	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	1	1	1	1	1	1	1
	Test	0,707	0,835	0,526	0,809	0,603	0,771	0,48
Model-2	Train-10 fold cV	0,985	0,978	0,978	0,978	0,981	0,981	0,963
	Test	0,828	0,768	0,48	0,78	0,608	0,798	0,47

GBM sonuçları incelendiğinde, Model-1'de eğitim veri setinde model tamamen doğru tahmin etmişken, test verisinde % 80,9 doğruluk, % 70,7 duyarlılık, % 60,3 F değerine sahiptir. Model-2'de ise eğitim veri setinde % 97,8 doğruluk, % 98,5 duyarlılık ve % 98,1 F değerine sahipken, test veri setinde % 78 doğruluk, % 82,8 duyarlılık ve F değeri % 60,8' e düşmüştür. Metrikler incelendiğinde Model 2'nin F değeri, AUROC değeri Model-1'den yüksekken, Kappa ve Kesinlik değerleri Model-1'de daha yüksektir.



Şekil 4-12: Model-1 GBM ROC Eğrisi



Şekil 4-13: Model 2 GBM ROC Eğrisi

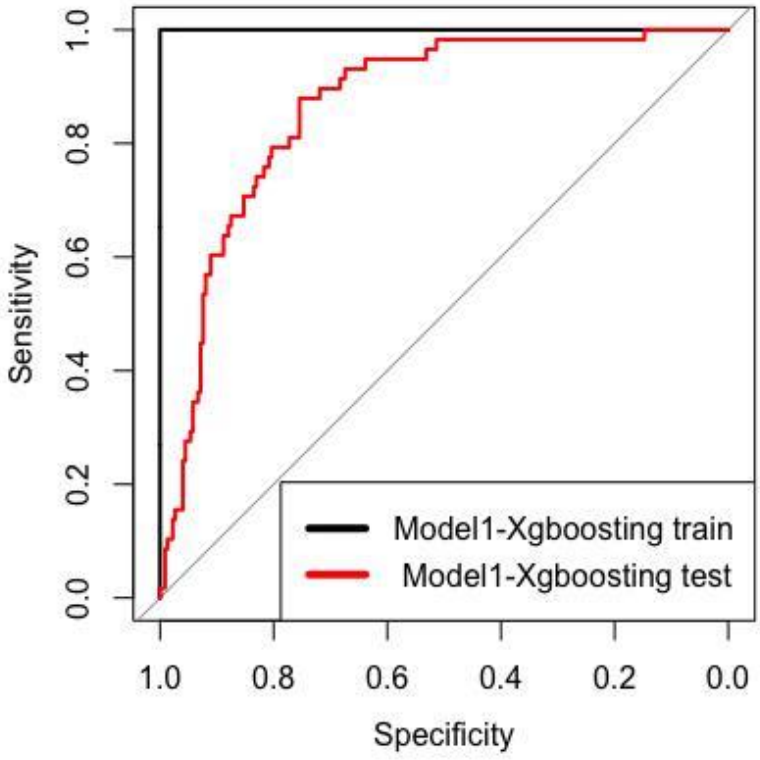
4.8. Xgboosting Bulguları

Xgboosting uyguluması için R studio'da xgboost(Chen & He, 2014) paketi kullanılmıştır. Model parametreleri ise grid yöntemi ile belirlenmiştir. En uygun parametreler Model-1 için ağaç sayısı, maximum iterasyon sayısı (nrounds) 100, ağacın derinliği (max_depth) 3, eta 0,3, gamma 0, ağaçtaki değişken sayısının kontrolü (colsample_bytree) 0,6, ağacı durdurma ağırlığı (min_child_weight) 1, alt örneklem (subsample) 1 olarak bulunmuş, Model-2'de ise parametreler sırasıyla 100;3;0,3;0;0,8;1;1 olarak bulunmuştur. Bulgular Tablo-4.12'de verilmiştir.

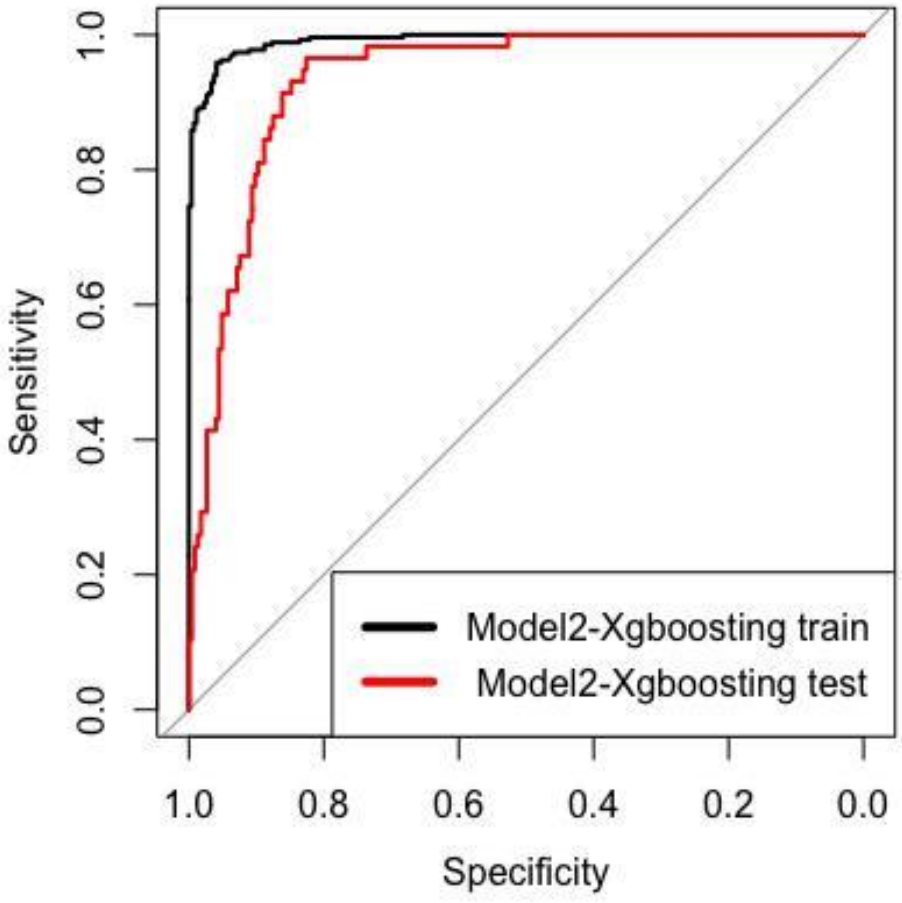
Tablo 4-10: Xgboosting Eğitim ve Test Veri seti Metrikleri

Modeller	Xgboosting	Duyarlılık	Seçicilik	Kesinlik	Doğruluk	F	AUROC	Kappa
Model-1	Train-10 fold cV	1	1	1	1	1	1	1
	Test	0,707	0,839	0,533	0,812	0,608	0,773	0,487
Model-2	Train-10 fold cV	0,97	0,933	0,935	0,952	0,952	0,951	0,903
	Test	0,862	0,723	0,446	0,752	0,588	0,793	0,435

Buna göre, Model-1'de eğitim veri setinde model tamamen doğru tahmin etmişken, test verisinde % 81,2 doğruluk, % 70,7 duyarlılık, % 60,8 F değerine sahiptir. Model-2'de ise eğitim veri setinde % 95,2 doğruluk, % 97 duyarlılık ve % 95,2 F değerine sahipken, test veri setinde % 75,2 doğruluk, % 86,2 duyarlılık ve F değeri % 58,8' e düşmüştür. Metrikler incelendiğinde, Model-1'in performansının Model-2'den daha iyi olduğu bulunmuştur.



Şekil 4-14: Model-1 Xgboosting ROC Eğrisi



Şekil 4-15: Model-2 Xgboosting ROC Eğrisi

4.9. Super Learner (SL) Bulguları

SL için algoritmaların birleştirilmesinde R'da SuperLearner(Eric Polley, Erin LeDell, Chris Kennedy, Sam Lendle, & Mark van der Laan, 2019) paketi kullanılmıştır. Bu modelde, SL algoritmaları içerisinden bireysel olarak en iyi sonucu veren LR, RO, GBM ve Xgboosting seçilerek analiz yapıldı. Tablo 4.11'de elde edilen risk ve katsayı analiz sonuçları verildi. SL tahminde sadece RO (ranger) ve Xgboosting algoritmasını kullandığı görüldü. Diğer LR, GBM algoritmalarının modelde ağırlıklarının olmadığı, yani tahminde kullanılmadıkları görüldü (Tablo 4.13). Xgboost ve RO için değişik parametreler denenerek tahminler yapıldı. En iyi sonucu veren SL algoritması raporlandı ve SL, Model-1 için uygulandı.

Tablo 4-11: Super Learner Algoritmaların Ağırlıkları

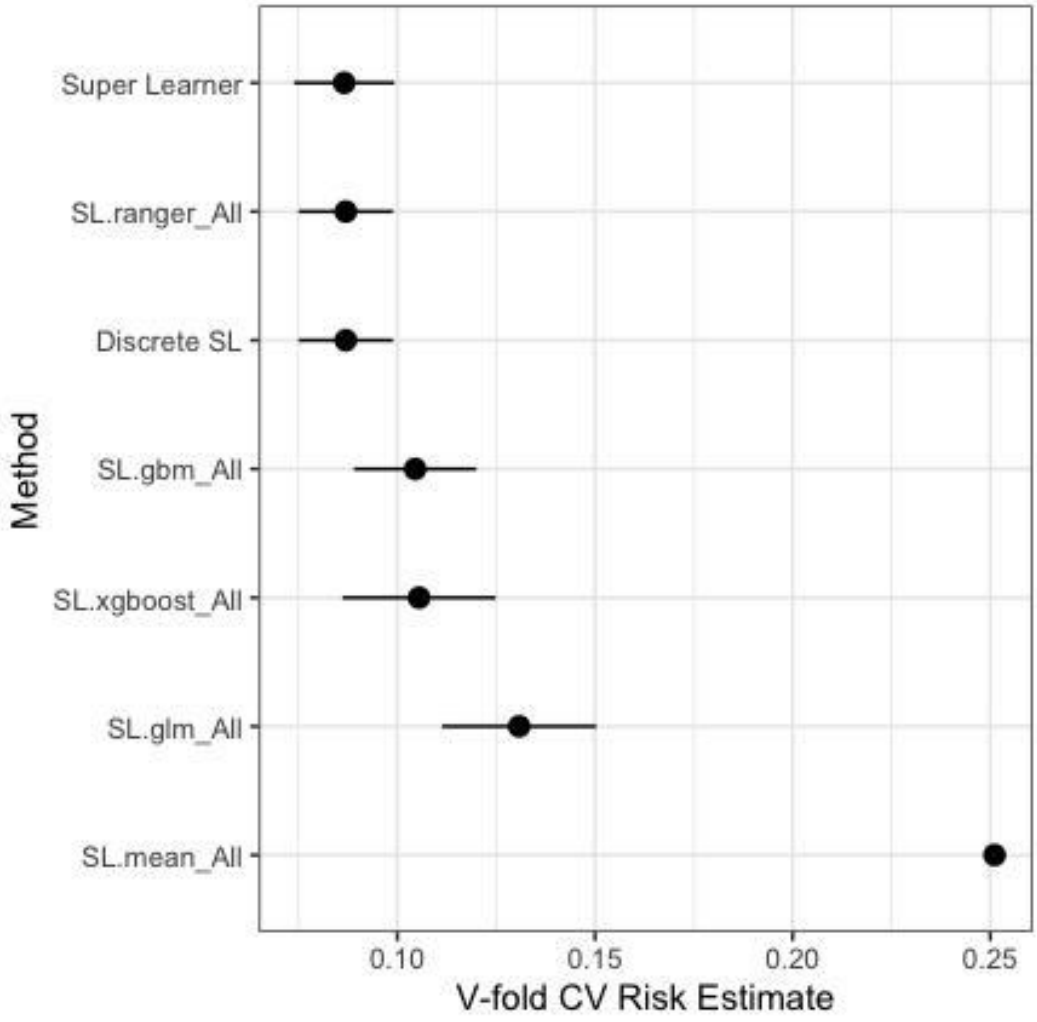
Modeller	Risk	Coefficient
SL.mean_All	0,2513	0,00000
SL.glm_All	0,1308	0,00000
SL.ranger_All	0,0874	0,84328
SL.gbm_All	0,1049	0,00000
SL.xgboost_All	0,1027	0,15672

Ardından 10 tekrar çapraz geçerlilik yapıldı. Çapraz geçerlilik, hem risk değeri için hem de ROC için yapıldı. Çapraz geçerlilik sonuçlarına göre en yüksek başarı oranı modellerin birleştirilmesiyle oluşan SL algoritmasındadır. RO ve dolayısıyla Discrete SL risk değerlendirmesi açısından çok az farkla SL algoritmasından daha düşük bir performans sergilemiştir. Tablo 4-12'de hem risk skorları hem de ROC eğrisi altında kalan alanlar sunulmuştur.

Tablo 4-12: SL algoritması sonucu risk ve AUROC istatistikleri

	Risk				AUROC		
Algoritma	**Ortalama**	**Standart hata**	**Min**	**Maks**	**Ortalama**	**Min**	**Maks**
Super Learner	0,08655	0,00647	0,05365	0,12014	0,95202	0,88333	1,00000
Discrete SL	0,08694	0,00610	0,05452	0,11252	0,95577	0,89167	1,00000
SL.mean_All	0,25109	0,00033	0,25000	0,25398	0,50000	0,50000	0,50000
SL.glm_All	0,13081	0,00995	0,07902	0,16612	0,88463	0,78889	0,95652
SL.ranger_All	0,08694	0,00610	0,05452	0,11252	0,95577	0,89167	1,00000
SL.gbm_All	0,10450	0,00790	0,07058	0,14110	0,93154	0,85417	0,98597
SL.xgboost_All	0,10550	0,00989	0,06172	0,17182	0,93064	0,86389	0,99176

Şekil 4.16'da modellerin Brier risk skorlarının % 95 güven aralıkları görülmektedir. SL ve RO algoritmalarının birbirlerine çok yakın tahmin gücüne sahip oldukları tablodan da görülmektedir.



Şekil 4-16: Modellerin çapraz risklerinin % 95 güven aralıkları

Yapılan tahminde ise AUROC'un % 88'e çıktığı bulunmuştur. Diğer yapılan tahminlerde AUROC'un en fazla % 82,5'e çıktığı görülmekte ve SL algoritması ile daha doğru tahmin yapıldığı bulunmuştur.

4.10. En İyi Modelin Seçilmesi

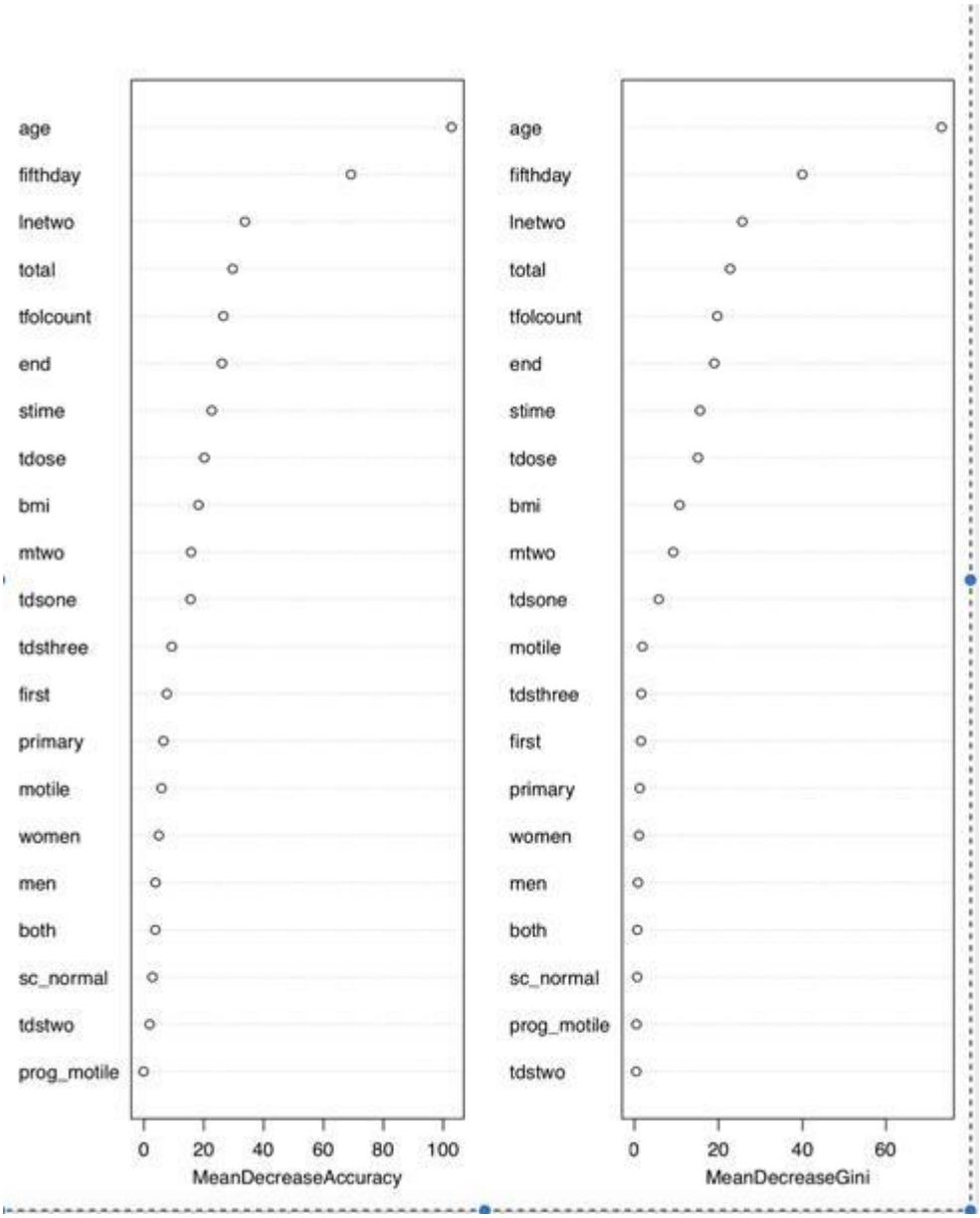
Tablo-4.15'de çalışmada kullanılan modellerin test kümesinin performans ölçütleri verilmiştir. Çalışmada kullanılan en iyi model, F ölçütü ve AUROC'a göre SL algoritması için Model-1 olarak bulunmuştur. Diğer modellerde en yüksek F değeri Model-1'de RO ile 0,666 iken, SL algoritması bu oranı 0,814'e çıkarmıştır. Diğer modellerde en yüksek AUROC Model-1'de 0,825 iken, SL bu oranı 0,88'e çıkartmıştır.

Tablo 4-13: Modellerin Test Verisine Ait Performans Ölçütleri

Modeller	MLA	Sensitivity	Specificity	Precision	Accuracy	F	AUROC
Model-1	**LR**	0,793	0,83	0,547	0,822	0,647	0,812
Model-1	**NB**	0,741	0,741	0,426	0,741	0,541	0,741
Model-2		0,776	0,603	0,334	0,638	0,47	0,69
Model-1	**RO**	0,81	0,839	0,566	0,833	0,666	0,825
Model-2		0,914	0,71	0,449	0,752	0,602	0,81
Model-1	**SVM**	0,655	0,817	0,481	0,784	0,555	0,736
Model-2		0,983	0,477	0,328	0,582	0,492	0,73
Model-1	**NNEt**	0,757	0,821	0,524	0,809	0,619	0,79
Model-2		0,862	0,705	0,431	0,738	0,575	0,784
Model-1	**Bagging**	0,707	0,79	0,466	0,773	0,562	0,749
Model-2		0,862	0,688	0,417	0,723	0,562	0,775
Model-1	**Gbm**	0,707	0,835	0,526	0,809	0,603	0,771
Model-2		0,828	0,768	0,48	0,78	0,608	0,798
Model-1	**Xgboosting**	0,707	0,839	0,533	0,812	0,608	0,773
Model-2		0,862	0,723	0,446	0,752	0,588	0,793
Model-1	**SL**	0,7455	0,8103	0,9382	0,7589	0,814	0,88

4.11. Değişkenlerin Önem Dereceleri

Çalışmada değişkenlerin önem derecesi belirlenirken en yüksek performansa sahip RO modeli sonuçları kabul edilmiştir (Şekil 4.17). Buna göre annenin yaşı, embriyonun transfer günü ve E2 seviyesi en önemli değişkenler olarak belirlenmiştir. Sperm hareketliliği, embriyo skorunun iki olması ve sperm sayısı IVF başarısında en az önem seviyesine sahip olan değişkenler olarak belirlenmiştir.



Şekil 4-17: Rastgele Orman Yöntemine Göre Değişkenlerin Önem Dereceleri

5. TARTIŞMA

Bu çalışmada, IVF tedavisinde embriyo seçimi sonrasında implantasyon başarısını makine öğrenmesi yöntemleri ile karşılaştırılması amaçlanmıştır. Çıktı değişkeni embryo implamantasyonunu başarılı/başarısız olarak gruplandırdığı için çalışma problemi ikili sınıflandırma problemidir. Tedavinin başarısının modellenmesinde embriyolojik, klinik ve demografik değişkenler bir arada kullanılmıştır. Modelleme de ise LR, NB, RO, DVM, YSA,Bagging, GBM, Xgboosting ve SL algoritmaları
kullanılmıştır. Makine öğrenmesi yöntemleri ile IVF tedavisinde başarının modellenmesinde, 939 embryodan elde edilen 17 değişken kullanılmıştır. Literatürde bu konuda yapılmış çalışmalar incelendiğinde en az 4 değişken (Raef & Ferdousi, 2019; Wald et al., 2005), en çok ise 82 (Raef et al., 2020) değişken ile çalışıldığı görülmüştür. Çalışılan değişkenler açısından incelendiğinde, 24 çalışmadan 9 çalışmada başarı modellemesinin embriyolojik, klinik ve demografik veriyi kullanarak yapıldığı görülmüştür (Blank et al., 2019; Guh, Wu, & Weng, 2011; Hassan, Al-Insaif, Hossain, & Kamruzzaman, 2020; La Marca & Sunkara, 2014; Mirroshandel, Ghasemian, & Monji-Azad, 2016; Raef et al., 2020; Ranjini, Suruliandi, & Raja, 2020; Uyar et al., 2015). Bu çalışmada ayrıca Raef vd. (Raef et al., 2020) çalışmasında olduğu gibi 3. Gün ve 5. Gün embriyolarının transferleri bir arada kullanılmıştır. Yapılan çalışmalarda, 5. Gün yapılan transferlerin daha başarılı olduğu söylenmiş ve daha sonraki yapılan çalışmalara transferlerin dahil edilmesi önerilmiştir(Raef et al., 2020; Uyar et al., 2015).

Çalışmada modelleme yapmak için iki ayrı model kullanılmıştır. Model-1 sürekli ve kesikli değişkenlerin bir arada kullanıldığı model, Model-2 ise sürekli değişkenleri kesikli değişken haline getirerek, bütün değişkenlerin kesikli olarak kullanıldığı modeldir. Veri setleri % 70 eğitim, % 30 test veri seti olarak ikiye bölünmüştür. Eğitim kümesine 10 kat çapraz geçerleme 10 tekrar yöntemi kullanılarak eğitim gerçekleşmiştir. Ayrıca veri setinde sınıflandırma problemlerinin tahmininde ciddi bir sıkıntı olan veri dengesizliği problemi bulunmaktadır. Bu problemi giderebilmek için Lojistik Regresyonda SMOTE, RUS, ROS, ROSE örnekleme yöntemleri kullanılmış, en iyi performansı sağlayan SMOTE yöntemi diğer modeller için de kullanılmıştır.

Bu çalışmada, IVF başarısının modellenmesi için dokuz makine öğrenmesi yöntemi kullanılmıştır. Modellerin performansının değerlendirilmesinde AUROC yönteminin başarılı bir performans karşılaştırma ölçütü olduğu bilinmektedir (Güvenir et

al., 2015; Raef et al., 2020). Test veri setinin AUROC Model-1 için % 74,1 (NB) ve %88 (SL) arasında değişmektedir. Model-2 için ise bu alanlar % 69 (NB) ve % 87,6 (LR) arasında değişmektedir. Bunun yanında, sınıf dengesizliği problemi gösteren çalışmalarda F değerinin önemli bir performans ölçütü olduğu literatürde belirtilmektedir. Modeller F değerlerine göre karşılaştırıldığında, Model-1 için iyi performans sergileyen yöntemin % 66 F ile RO olduğu sonrasında ise sırasıyla, LR (% 64,7 F), YSA (% 61,9 F), Xgboostimg (% 60,8 F), Gbm (% 60,3) modellerinin en iyi sonuç verdiği bulunmuştur. Model-2 için ise en iyi tahmini yapan yöntemin % 63,8 F değeri ile NB olarak bulunmuştur. Ardından sırasıyla iyi performansa sahip modeller: LR (% 61,6 F), GBM (% 60,8 F), RO (% 60,2 F)'dir.

Çalışmada, hem sürekli hem kesikli değişkenlerle bir arada modellenen Model-1'in performansı en yüksek bulunmuştur. Bunun yanında, Model-1 için SL ile model kurulmuş ve kurulan modelde RO ve Xgboosting dışındaki modellerin SL'deki ağırlıklarının olmadığı görülmüştür. Kesikli SL'de RO algoritması en iyi performans sergilediği için seçilmiştir. SL'de ise RO'nun ağırlığı ise Xgboosting'den yüksektir. SL ile AUROC alanı %88'e , F ise % 81,4'e çıkartılmıştır.

Literatür incelendiğinde IVF modellemesinde topluluk öğrenmesine dayanan karar ağacı modellerinin son yıllarda kullanıldığı görülmüştür(Hafiz, Nematollahi, Boostani, & Jahromi, 2017; Mirroshandel et al., 2016; Qiu, Li, Dong, Xin, & Tan, 2019). Bu konuda yapılan çalışmalar incelendiğinde RO yönteminin daha önce yapılan beş araştırmada en iyi tahmin gücüne sahip olduğu görülmektedir(Blank et al., 2019; Goyal, Kuchana, & Ayyagari, 2020; Hafiz et al., 2017; Mirroshandel et al., 2016; Raef et al., 2020). Qiu ve arkadaşları (Qiu et al., 2019) Xgboosting, LR, DVM ve RO'yu karşılaştırdıkları çalışmalarında Xgboosting'i en iyi model olarak bulmuşlardır. Ayrıca çalışmamızda, Model-2'de LR'nin de iyi performans gösterdiği bulunmuş ve literatürde makine öğrenmesi kullanan ART çalışmalarından yalnızca birinde kullanıldığı görülmüştür(Spiessens et al., 2016) . Literatürde, LR'nin pek çok modern makine öğrenmesi yöntemine göre daha iyi sonuç verdiği çalışmalar mevcuttur (Baćak & Kennedy, 2018; Faisal et al., 2018). Ayrıca, Model-1 için YSA'nın iyi sonuç verdiği bulunmuştur. Literatürde YSA'nın IVF'de başarının modellendinmesinde pek.çok çalışmada kullanıldığı görülmektedir(Barnett-Itzhaki et al., 2020; Brás de Guimarães et al., 2020; Chen, Cheng, Hsu, & Li, 2009; Durairaj & Nandhakumar, 2013; Goyal et al.,

2020; Hafiz et al., 2017; Hassan et al., 2020; Kaufmann, Eastaugh, Snowden, Smye, & Sharma, 1997; Milewski, Milewska, Wiesak, & Morgan, 2013; Mirroshandel et al., 2016; Nanni L, 2010; Raef et al., 2020; Uyar et al., 2015; Wald et al., 2005). Bu çalışmalardan, beş tanesinde en iyi model olduğu bulunmuştur (Brás de Guimarães et al., 2020; Hassan et al., 2020; Milewski et al., 2013; Raef et al., 2020; Wald et al., 2005). Ayrıca, Model-2'de iyi bir perfornans sergileyen NB, Türkiye'de bir tüp bebek merkezinden alınarak gerçekleştirilen Bener vd. (Uyar et al., 2015) çalışmasında en iyi model olarak bulunmuştur.

Her ne kadar IVF tedavisinin başarı oranı her geçen gün artsa da istenilen oranlara ulaşamamıştır. Literatürde, annenin yaşı, yumurta rezervi IVF başarısını etkileyen önemli değişkenler olduğu bilinse de E2 seviyesi, embriyonun transfer edildiği gün gibi değişenlerin başarı üzerindeki etkisi hala tartışılmaktadır (Prasad, Kumar, Singhal, & Sharma, 2014).

Bu çalışmada, IVF tedavisinde, embriyonun tutunma olasılığını çiftlerin demografik özellikleri, embriyo kalitesi ve COS değişkenler ile makine öğrenmesi yöntemleri ile tahmin etmektir. Çalışmada en iyi tahmini yapan makine öğrenmesi tekniğinin belirlenmesinin yanı sıra değişkenlerin önem derecesi belirlenmiştir.

Ayrıca, çalışmada en önemli değişkenler annenin yaşı, embriyonun transfer günü ve E2 seviyesi olarak belirlenmiştir. Annenin yaşı, IVF başarısında pek çok makine öğrenmesi modelleme çalışmasında en önemli değişken olarak bulunmuştur (Corani, Magli, Giusti, Gianaroli, & Gambardella, 2013; Guh et al., 2011; Hafiz et al., 2017; Hassan et al., 2020; Kaufmann et al., 1997; Kim & Jung, 2003; Spiessens et al., 2016; Uyar et al., 2015; Wald et al., 2005). Bunun dışında IVF çalışmalarında da tarafından da annenin yaşının en önemli değişken olduğu belirtilmiştir (Cimadomo et al., 2018; Ubaldi et al., 2019). Bir sonraki önemli değişken, embriyonun transfer günü olarak bulunmuştur. Yine infertilite çalışmalarında bu konu tartışılmakta, genel kabul 5. Günde yapılan transferlerin daha başarılı olduğu savunulmaktadır (Haas, Meriano, Bassil, Barzilay, & Casper, 2019; Sacha, Dimitriadis, Christou, Souter, & Bormann, 2018; Yang et al., 2018). Sonraki önemli değişkenler, E2 seviyesi ve yumurta sayısıdır. Literatürde bu değişkenlerin başarıyı etkilediği konunun uzmanları tarafından belirtilmiştir(Broer et al., 2013; Iliodromiti & Nelson, 2015; La Marca & Sunkara, 2014).

Bu çalışma, IVF tedavisinde istatistiksel öğrenme algoritmalarının embriyo implamantasyonunu modellenmesi amacıyla yapılmıştır. Bu modelin kullanılmasıyla, tıbbi, maddi ve manevi pek çok zararı söz konusu olan tedavinin, başarı oranının arttırılmasının sağlanması amaçlanmıştır. SL ve RO algoritmalarının, çalışmada implantasyonu tahmin etmek için en iyi modeller olduğu belirlenmiştir. Son olarak, çalışma, IVF uzmanlarının başarıyı tahmin etmelerine ve tedaviden önce hastalara danışmanlık yapmalarına yardımcı olacak bir klinik karar destek sistemi yaklaşımı için bir model önermiştir.

İnfertilite, üreme çağındaki popülasyonu her geçen gün daha fazla etkileyen bir halk sağlığı sorunudur. Konu ile ilgili giderek artan sayıda çalışma yapılmakta ve olayın hem tıbbi ve ekonomik hem de psikolojik ve sosyal boyutları ele alınmaktadır. İnfetilite tedavisinde kullanılan temel tekniklerden biri olan IVF yöntemi ile bu sorunu yaşayan çiftlere bir çözüm ortaya konulmakta ve normalde gebelik şansı olmayan çiftlerin çocuk sahibi olabilmeleri mümkün hale gelmektedir. Ancak IVF ile başarı şansı, tekrarlı denemeler sonucunda bile %49'u geçemediği literatürde belirtilmekte, bu da maliyet, zaman kaybına neden olduğu gibi aynı zamanda ilaç kullanımına bağlı riski, psikolojik ve ailevi sorunları artırmakta ve devlete maliyeti artırmaktadır. IVF tedavisinde sonucu etkileyen çok fazla parametre olması nedeniyle, bu parametrelerin tedavi başarısı üzerine etkisi tam olarak bilinmemektedir. Bu çalışmada ortaya koyulan sonuçlar doğrultusunda IVF te tedavi başarısını etkileyen faktörlerin önemi ve önem sırası belirlenmiştir.

Elde edilen sonuçların klinik anlamlılığı incelendiğinde ise,

Klinisyeninin hastaya özel faktörleri gözönünde bulundurarak bireysel tedavi yöntemlerinin (over stimülsayonu ve laboratuvar uygulamalarında) planlamasına,

Farklı hasta alt grupları için farklı tedavi planlarının geliştirilmesi, optimize edilmesi ve hem ulusal hem uluslararası tedavi kılavuzlarının hazırlanmasına,

Kişiye özel tedavi yöntemlerinin kullanılması sonucu başarı oranlarının yükseltilmesine,

Her hastaya tedavi başlamadan önce tedavinin başarı oranlarının predikte edilmesi ve hastanın bilgilendirilmesine,

Tedavi başarısını tahmin edecek yazılımların geliştirilmesi ve tüp bebek merkezlerinde kullanıma uygun olarak ticari hale dönüştürülmesine,

Tedavi kararının verilmesi ve devam ettirilmesinde klinisyen ve çiftlere gerçek ve kanıta dayalı veri sunulmasına,

Tedavinin kümülatif ve siklus başına maliyet-etkinlik analizlerinin yapılmasına,

Tedavinin ne kadar devam edeceği ve ne zaman sonlandırılması gerektiği konusunda karar alma mekanizmalarına katkı sağlamasına,

Tedaviye yüksek derecede etki eden, ya da daha önce etki derecesi bilinmeyen yeni faktörlerin tedavisi için yeni çalışmaların planlanmasına

Yeni ortaya koyulan faktörlerin alt gruplarda detaylandırılıp araştırılmasına

Embriyo seçim mekanizmalarının geliştirilip uygulanabilmesine,

Yüksek öneme sahip faktörlere gereken dikkatin çeklimesi ve bu konularla ilgili devlet politikası geliştirilebilmesine

ışık tutacağı düşünülmektedir.

Ayrıca, bir halk sağlığı sorunu olan infertilite ve bunun sonucunda gelişen yardımcı üreme teknikleri alanında, çalışmada kullanılan merkez sayısı ve dolayısıyla değişken ve örneklem sayısı arttıralarak daha kapsamlı araştırmalar planlanabilir. Bu araştırmalar, derin öğrenme yöntemleri (deep learning) kullanılarak çözülebilir ve alana daha etkili katkılar sunulabilir.

KAYNAKLAR

Akpınar, H. (2014). *Data Veri Madenciliği - Veri Analizi*. (H. Akpınar, Ed.) (1st ed.). İstanbul: Papatya Bilim.

Alpar, R. (2011). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*. Detay Yayıncılık.

Archer, K. J., & Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Computational Statistics and Data Analysis*, *52*(4), 2249–2260. https://doi.org/10.1016/j.csda.2007.08.015

Ayhan, S., & Erdoğmuş Şenol. (2014). Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, *9*(1), 175–201. https://doi.org/10.17153/eoguiibfd.33265

Baćak, V., & Kennedy, E. H. (2018). Principled Machine Learning Using the Super Learner. *Sociological Methods & Research*, 004912411774730. https://doi.org/10.1177/0049124117747301

Balaban, B., Urman, B., Sertac, A., Alatas, C., Aksoy, S., & Mercan, R. (2000). Blastocyst quality affects the success of blastocyst-stage embryo transfer. *Fertility and Sterility*, *74*(2), 282–287. https://doi.org/10.1016/S0015-0282(00)00645-2

Barnett-Itzhaki, Z., Elbaz, M., Butterman, R., Amar, D., Amitay, M., Racowsky, C., … Machtinger, R. (2020). Machine learning vs. classic statistics for the prediction of IVF outcomes. *Journal of Assisted Reproduction and Genetics*. https://doi.org/10.1007/s10815-020-01908-1

Bilgin, M. (2018). *Makine Öğrenmesi* (2nd ed.). İstanbul: Papatya Yayıncılık Eğitim.

Blank, C., Wildeboer, R. R., DeCroo, I., Tilleman, K., Weyers, B., de Sutter, P., … Schoot, B. C. (2019). Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. *Fertility and Sterility*, *111*(2), 318–326. https://doi.org/10.1016/j.fertnstert.2018.10.030

Brás de Guimarães, B., Martins, L., Metello, J. L., Ferreira, F. L., Ferreira, P., & Fonseca, J. M. (2020). Application of Artificial Intelligence Algorithms to Estimate the Success Rate in Medically Assisted Procreation. *Reproductive Medicine*, *1*(3), 181– 194. https://doi.org/10.3390/reprodmed1030014

Breiman, L. (1996a). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1023/A:1018054314350

Breiman, L. (1996b). Stacked regressions. *Machine Learning*, *24*(1), 49–64. https://doi.org/10.1023/A:1018046112532

Breiman, L. E. O. (2001). Random Forests, 5–32.

Broer, S. L., van Disseldorp, J., Broeze, K. A., Dolleman, M., Opmeer, B. C., Bossuyt, P., … Vladimirov, I. K. (2013). Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: An individual patient data approach. *Human Reproduction Update*, *19*(1), 26–36. https://doi.org/10.1093/humupd/dms041

Cai, Q., Wan, F., Appleby, D., Hu, L., & Zhang, H. (2014). Quality of embryos transferred and progesterone levels are the most important predictors of live birth after fresh embryo transfer: A retrospective cohort study. *Journal of Assisted Reproduction and Genetics*, *31*(2), 185–194. https://doi.org/10.1007/s10815-013-0129-4

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1002/eap.2043

Chen, C. C., Cheng, Y. C., Hsu, C. C., & Li, S. T. (2009). Knowledge discovery on in vitro fertilization clinical data using particle swarm optimization. *Proceedings of the 2009 9th IEEE International Conference on Bioinformatics and BioEngineering, BIBE 2009*, 278–283. https://doi.org/10.1109/BIBE.2009.36

Cimadomo, D., Fabozzi, G., Vaiarelli, A., Ubaldi, N., Ubaldi, F. M., & Rienzi, L. (2018, June 29). Impact of maternal age on oocyte and embryo competence. *Frontiers in Endocrinology*. Frontiers Media S.A. https://doi.org/10.3389/fendo.2018.00327

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational And Psychological Measurement*, *XX*(1), 37–46.

Corani, G., Magli, C., Giusti, A., Gianaroli, L., & Gambardella, L. M. (2013). A Bayesian network model for predicting pregnancy after in vitro fertilization. *Computers in Biology and Medicine*, *43*(11), 1783–1792.

https://doi.org/10.1016/j.compbiomed.2013.07.035

Cummins, J. M., Breen, T. M., Harrison, K. L., Shaw, J. M., Wilson, L. M., & Hennessey, J. F. (1986). A formula for scoring human embryo growth rates in in vitro fertilization: Its value in predicting pregnancy and in comparison with visual estimates of embryo quality. *Journal of In Vitro Fertilization and Embryo Transfer*, *3*(5), 284–295. https://doi.org/10.1007/BF01133388

Desai, N. N. (2002). Morphological evaluation of human embryos and derivation of an embryo quality scoring system specific for day 3 embryos: a preliminary study. *Human Reproduction*, *15*(10), 2190–2196. https://doi.org/10.1093/humrep/15.10.2190

Dhillon, R. K., McLernon, D. J., Smith, P. P., Fishel, S., Dowell, K., Deeks, J. J., … Coomarasamy, A. (2016). Predicting the chance of live birth for women undergoing IVF: A novel pretreatment counselling tool. *Human Reproduction*, *31*(1), 84–92. https://doi.org/10.1093/humrep/dev268

Ding, C., Cao, X. (Jason), & Næss, P. (2018). Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo. *Transportation Research Part A: Policy and Practice*, *110*(March), 107–117. https://doi.org/10.1016/j.tra.2018.02.009

Doğaner, A. (2020). *TOPLULUK ÖĞRENME YÖNTEMLERİ İLE RENAL HÜCRELİ KARSİNOM'UN TAHMİN EDİLMESİ*. İnönü Üniversitesi Sağlık Bilimleri Enstitüsü Doktora Tezi.

Dokras, A., Sargent, I. L., & Barlow, D. H. (1993). Fertilization and early embryology: Human blastocyst grading: An indicator of developmental potential? *Human Reproduction*, *8*(12), 2119–2127. https://doi.org/10.1093/oxfordjournals.humrep.a137993

Durairaj, M., & Nandhakumar, R. (2013). Data Mining Application on IVF Data For The Selection of Influential Parameters on Fertility. *International Journal of Engineering and Advanced Technology*, *2*(6), 262–266. Retrieved from http://www.ijeat.org/attachments/File/v2i6/F2068082613.pdf

Faisal, M., Scally, A., Howes, R., Beatson, K., Richardson, D., & Mohammed, M. A. (2018). A comparison of logistic regression models with alternative machine

learning methods to predict the risk of in-hospital mortality in emergency medical admissions via external validation. *Health Informatics Journal*. https://doi.org/10.1177/1460458218813600

Freund, Y., Schapire, R. E., & Hill, M. (1996). Experiments with a New Boosting Algorithm. *Thirteenth International Conference on ML*.

Friedman, J. H. (2001). Greedy Function Approximation : A Gradient Boosting Machine Author (s): Jerome H . Friedman Source : The Annals of Statistics , Oct ., 2001 , Vol . 29 , No . 5 (Oct ., 2001), pp . 1189-1232 Published by : Institute of Mathematical Statistics Stable UR, *29*(5), 1189–1232.

Gardner, D. K., Lane, M., Stevens, J., Schlenker, T., & Schoolcraft, W. B. (2000). Blastocyst score affects implantation and pregnancy outcome: Towards a single blastocyst transfer. *Fertility and Sterility*, *73*(6), 1155–1158. https://doi.org/10.1016/S0015-0282(00)00518-5

Gardner, D. K., & Schoolcraft, W. B. (1999). Culture and transfer of human blastocysts. *Current Opinion in Obstetrics and Gynecology*. https://doi.org/10.1097/00001703-199906000-00013

Gerris, J., & Evers, J. L. H. (1999). Prevention of twin pregnancy after in vitro fertilization/intracytoplasmic sperm injection based on strict embryo criteria: A prospective randomized trial. *Evidence-Based Obstetrics and Gynecology*, *1*(3), 98– 99. https://doi.org/10.1093/humrep/14.suppl_3.8-a

Goyal, A., Kuchana, M., & Ayyagari, K. P. R. (2020). Machine learning predicts live-birth occurrence before in-vitro fertilization treatment. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-76928-z

Guh, R. S., Wu, T. C. J., & Weng, S. P. (2011). Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. *Expert Systems with Applications*, *38*(4), 4437–4449. https://doi.org/10.1016/j.eswa.2010.09.112

Gürsakal, N. (2018). *Makine Öğrenmesi - - Dora Yayıncılık*. (Dora Yayıncılık, Ed.) (1st ed.). Bursa.

Güvenir, H. A., Misirli, G., Dilbaz, S., Ozdegirmenci, O., Demir, B., & Dilbaz, B. (2015).

Estimating the chance of success in IVF treatment using a ranking algorithm. *Medical and Biological Engineering and Computing*, *53*(9), 911–920. https://doi.org/10.1007/s11517-015-1299-2

Haas, J., Meriano, J., Bassil, R., Barzilay, E., & Casper, R. F. (2019). What is the optimal timing of embryo transfer when there are only one or two embryos at cleavage stage? *Gynecological Endocrinology*, *35*(8), 665–668. https://doi.org/10.1080/09513590.2019.1580259

Hachesu, P. R., Ahmadi, M., Alizadeh, S., & Sadoughi, F. (2013). Use of data mining techniques to determine and predict length of stay of cardiac patients. *Healthcare Informatics Research*, *19*(2), 121–129. https://doi.org/10.4258/hir.2013.19.2.121

Hafiz, P., Nematollahi, M., Boostani, R., & Jahromi, B. N. (2017). Predicting implantation outcome of in vitro fertilization and intracytoplasmic sperm injection using data mining techniques. *International Journal of Fertility and Sterility*, *11*(3), 184–190. https://doi.org/10.22074/ifs.2017.4882

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining_ Concepts and Techniques (2012, Morgan Kaufmann Publishers) - libgen.lc.pdf*. *Data Mining*. https://doi.org/10.1016/b978-0-12-381479-1.00001-0

Hassan, M. R., Al-Insaif, S., Hossain, M. I., & Kamruzzaman, J. (2020). A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*, *32*(7), 2283–2297. https://doi.org/10.1007/s00521-018-3693-9

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning* (Second Edi). Springer US.

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832–844. https://doi.org/10.1109/34.709601

Iliodromiti, S., & Nelson, S. M. (2015, June 6). Ovarian response biomarkers: Physiology and performance. *Current Opinion in Obstetrics and Gynecology*. Lippincott Williams and Wilkins. https://doi.org/10.1097/GCO.0000000000000175

Kantardzic, M. (2020). *Data Mining: Concepts, Models, Methods, and Algorithms*. (E.

Hossain, Ed.), *IIE Transactions* (3rd ed., Vol. 36). Piscataway: IEE Press. https://doi.org/10.1080/07408170490426107

Kaufmann, S. J., Eastaugh, J. L., Snowden, S., Smye, S. W., & Sharma, V. (1997). The application of neural networks in predicting the outcome of in-vitro fertilization. *Human Reproduction*, *12*(7), 1454–1457. https://doi.org/10.1093/humrep/12.7.1454

Kim, I.-C., & Jung, Y.-G. (2003). *Using Bayesian Networks to Analyze Medical Data*. (P. Perner & A. Rosenfeld, Eds.) (Machine le). Retrieved from http://link.springer.com/content/pdf/10.1007/978-3-642-03070-3.pdf

Koç, S. (2019). *Genetik alaninda elde edilen verilerin makine öğrenimi algoritmalari yardimiyla karşilaştirilarak en etkin yöntemin belirlenmesi*. Ondokuz Mayıs Üniversitesi Sağlık Bilimleri Enstitsü.

La Marca, A., & Sunkara, S. K. (2014). Individualization of controlled ovarian stimulation in IVF using ovarian reserve markers: From theory to practice. *Human Reproduction Update*, *20*(1), 124–140. https://doi.org/10.1093/humupd/dmt037

Laan, M. J. van der, & Rose, S. (2011). *Targeted Learning - preface*. *Springer Series in Statistics* (Vol. 27). https://doi.org/10.1007/978-1-4419-9782-1

Lewin, A., Schenker, J. G., Safran, A., Zigelman, N., Avrech, O., Abramov, Y., … Reubinoff, B. E. (1994). Embryo growth rate in vitro as an indicator of embryo quality in IVF cycles. *Journal of Assisted Reproduction and Genetics*, *11*(10), 500–503. https://doi.org/10.1007/BF02216029

Mannila, H. (1996). Data mining: machine learning, statistics, and databases. In *In Proceedings of 8th International Conference on Scientific and Statistical Data Base Management* . *IEEE*. (pp. 2–9). Retrieved from http://link.springer.com/content/pdf/10.1007/978-3-642-03070-3.pdf

McLernon, D. J., Steyerberg, E. W., Te Velde, E. R., Lee, A. J., & Bhattacharya, S. (2016). Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: Population based study of linked cycle data from 113 873 women. *BMJ (Online)*, *355*. https://doi.org/10.1136/bmj.i5735

Mencar, C., Mantero, M., Tarsia, P., Carpagnano, G. E., Barbaro, M. P. F., & Lacedonia, D. (2019). Application of machine learning to predict obstructive sleep apnea

syndrome severity. https://doi.org/10.1177/1460458218824725

Milewski, R., Milewska, A. J., Wiesak, T., & Morgan, A. (2013). Comparison of artificial neural networks and logistic regression analysis in pregnancy prediction using the in vitro fertilization treatment. *Studies in Logic, Grammar and Rhetoric*, *35*(48), 39– 48. https://doi.org/10.2478/slgr-2013-0033

Mirroshandel, S. A., Ghasemian, F., & Monji-Azad, S. (2016). Applying data mining techniques for increasing implantation rate by selecting best sperms for intra-cytoplasmic sperm injection treatment. *Computer Methods and Programs in Biomedicine*, *137*, 215–229. https://doi.org/10.1016/j.cmpb.2016.09.013

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of Mahine Learning, second edition*. The MIT Press Cambridge, Massachusetts London, England.

Nanni L, L. A. and M. C. (2010). A data mining approach for predicting the pregnancy rate in human assisted reproduction. In Brahnam S and Jain LC (Ed.), *Advanced computational intelligence paradigms in healthcare 5* (pp. 97–111). Berlin: Springer Berlin Heidelberg.

Orhunbilge, N. (n.d.). *Çok Değişkenli İstatistik Yöntemler*. İstanbul : İstanbul Üniversitesi Basım ve Yayınevi Müdürlüğü, 2010. Retrieved from http://katalog.istanbul.edu.tr/client/tr_TR/default_tr/search/detailnonmodal/ent:$002f$002fSD_ILS$002f0$002fSD_ILS:422943/ada?qu=İstatistik.&rw=300&ic=true&ps=300&isd=true

Öztemel, E. (2006). *YAPAY SİNİR AĞLARI* (2. Basım). İstanbul: Papatya Yayıncılık Eğitim. Retrieved from www.papatya.gen.tr

Polley, E. C., Hubbard, A. E., & Laan, M. J. Van Der. (2010). Super Learner In Prediction. *The Berkeley Electronic Press*.

Prasad, S., Kumar, Y., Singhal, M., & Sharma, S. (2014). Estradiol level on day 2 and day of trigger: A potential predictor of the IVF-ET success. *Journal of Obstetrics and Gynecology of India*, *64*(3), 202–207. https://doi.org/10.1007/s13224-014-0515-6

Q, X., Q, Y., M, W., B, H., B, Z., Z, L., … L, J. (2021). Individualized Embryo Selection Strategy Developed by Stacking Machine Learning Model for Better In Vitro

Fertilization Outcomes: An Application Study. https://doi.org/10.21203/RS.3.RS-140383/V1

Qiu, J., Li, P., Dong, M., Xin, X., & Tan, J. (2019). Personalized prediction of live birth prior to the first in vitro fertilization treatment: A machine learning method. *Journal of Translational Medicine*, *17*(1), 1–8. https://doi.org/10.1186/s12967-019-2062-5

Raef, B., & Ferdousi, R. (2019). A review of machine learning approaches in assisted reproductive technologies. *Acta Informatica Medica*, *27*(3), 205–211. https://doi.org/10.5455/aim.2019.27.205-211

Raef, B., Maleki, M., & Ferdousi, R. (2020). Computational prediction of implantation outcome after embryo transfer. *Health Informatics Journal*, *26*(3), 1810–1826. https://doi.org/10.1177/1460458219892138

Ranjini, K., Suruliandi, A., & Raja, S. P. (2020). An Ensemble of Heterogeneous Incremental Classifiers for Assisted Reproductive Technology Outcome Prediction. *IEEE Transactions on Computational Social Systems*, 1–11. https://doi.org/10.1109/TCSS.2020.3032640

Richter, K. S., Harris, D. C., Daneshmand, S. T., & Shapiro, B. S. (2001). Quantitative grading of a human blastocyst: Optimal inner cell mass size and shape. *Fertility and Sterility*, *76*(6), 1157–1167. https://doi.org/10.1016/S0015-0282(01)02870-9

Rienzi, L., Ubaldi, F., Iacobelli, M., Romano, S., Minasi, M. G., Ferrero, S., … Greco, E. (2005). Significance of morphological attributes of the early embryo. *Reproductive BioMedicine Online*, *10*(5), 669–681. https://doi.org/10.1016/S1472-6483(10)61676-8

Roseboom, T. J., Vermeiden, J. P. W., Schoute, E., Lens, J. W., & Schats, R. (1995). The probability of pregnancy after embryo transfer is affected by the age of the patient, cause of infertility, number of embryos transferred and the average morphology score, as revealed by multiple logistic regression analysis. *Human Reproduction*, *10*(11), 3035–3041. https://doi.org/10.1093/oxfordjournals.humrep.a135842

Sacha, C. R., Dimitriadis, I., Christou, G., Souter, I., & Bormann, C. L. (2018). The effect of day 2 versus day 3 embryo transfer on early pregnancy outcomes in women with a low yield of fertilized oocytes. *Journal of Assisted Reproduction and Genetics*, *35*(5), 879–884. https://doi.org/10.1007/s10815-018-1157-x

Scott, L. A., & Smith, S. (1998). The successful use of pronuclear embryo transfers the day following oocyte retrieval. *Human Reproduction*, *13*(4), 1003–1013. https://doi.org/10.1093/humrep/13.4.1003

Sharlip, I. D., Jarow, J. P., Belker, A. M., Lipshultz, L. I., Sigman, M., Thomas, A. J., … Sadovsky, R. (2002). Best practice policies for male infertility. *Fertility and Sterility*, *77*(5), 873–882. https://doi.org/10.1016/S0015-0282(02)03105-9

Speirs, A. L., Baker, H. W. G., & Abdullah, N. (1996). Analysis of factors affecting embryo implantation. *Human Reproduction*, *11*(SUPPL. 1), 187–191. https://doi.org/10.1093/humrep/11.suppl_5.187

Spiessens, C., De Neubourg, D., Peeraer, K., D'Hooghe, T., Chen, F., & Debrock, S. (2016). Selecting the embryo with the highest implantation potential using a data mining based prediction model. *Reproductive Biology and Endocrinology*, *14*(1), 1– 13. https://doi.org/10.1186/s12958-016-0145-1

Tantithamthavorn, C., Hassan, A. E., & Matsumoto, K. (2018). The impact of class rebalancing techniques on the performance and interpretation of defect prediction models. *ArXiv*, *46*(11), 1200–1219.

Tesarik, J., & Greco, E. (1999). The probability of abnormal preimplantation development can be predicted by a single static observation on pronuclear stage morphology. *Human Reproduction*, *14*(5), 1318–1323. https://doi.org/10.1093/humrep/14.5.1318

Trimarchi, J. R. (2001). A mathematical model for predicting which embryos to transfer - An illusion of control or a powerful tool? [6] (multiple letters). *Fertility and Sterility*, *76*(6), 1286–1287. https://doi.org/10.1016/S0015-0282(01)02915-6

Tu, J. V. (1996). Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes, *49*(11), 1225–1231.

Ubaldi, F. M., Cimadomo, D., Vaiarelli, A., Fabozzi, G., Venturella, R., Maggiulli, R., … Rienzi, L. (2019). Advanced maternal age in IVF: Still a challenge? The present and the future of its treatment. *Frontiers in Endocrinology*. Frontiers Media S.A. https://doi.org/10.3389/fendo.2019.00094

Uyar, A., Bener, A., & Ciray, H. N. (2015). Predictive Modeling of Implantation

Outcome in an in Vitro Fertilization Setting. *Medical Decision Making*, *35*(6), 714– 725. https://doi.org/10.1177/0272989X14535984

Van Blerkom, J., Bell, H., & Weipz, D. (1990). Cellular and developmental biological aspects of bovine meiotic maturation, fertilization, and preimplantation embryogenesis in vitro. *Journal of Electron Microscopy Technique*, *16*(4), 298–323. https://doi.org/10.1002/jemt.1060160404

Van Der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Statistical Applications in Genetics and Molecular Biology Super Learner Super Learner. *Degruyter.Com*, *6*(1). Retrieved from https://www.degruyter.com/view/j/sagmb.2007.6.issue-1/sagmb.2007.6.1.1309/sagmb.2007.6.1.1309.xml

Van Royen, E., Mangelschots, K., De Neubourg, D., Laureys, I., Ryckaert, G., & Gerris, J. (2001). Calculating the implantation potential of day 3 embryos in women younger than 38 years of age: A new model. *Human Reproduction*, *16*(2), 326–332. https://doi.org/10.1093/humrep/16.2.326

Van Royen, Eric, Mangelschots, K., De Neubourg, D., Valkenburg, M., Van De Meerssche, M. V., Ryckaert, G., … Gerris, J. (1999). Characterization of a top quality embryo, a step towards single-embryo transfer. *Human Reproduction*, *14*(9), 2345–2349. https://doi.org/10.1093/humrep/14.9.2345

Vander Borght, M., & Wyns, C. (2018, December 1). Fertility and infertility: Definition and epidemiology. *Clinical Biochemistry*. Elsevier Inc. https://doi.org/10.1016/j.clinbiochem.2018.03.012

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory* (1st editio). USA: Springer.

Wald, M., Sparks, A. E. T., Sandlow, J., Van-Voorhis, B., Syrop, C. H., & Niederberger, C. S. (2005). Computational models for prediction of IVF/ICSI outcomes with surgically retrieved spermatozoa. *Reproductive BioMedicine Online*, *11*(3), 325– 331. https://doi.org/10.1016/S1472-6483(10)60840-1

Wang, C., Kou, S., & Song, Y. (2019). Identify risk pattern of e-bike riders in China based on machine learning framework. *Entropy*, *21*(11). https://doi.org/10.3390/e21111084

Wang, X., & Zhong, Y. (2003). Statistical learning theory and state of the art in SVM. *Proceedings - 2nd IEEE International Conference on Cognitive Informatics, ICCI 2003*, (2), 55–59. https://doi.org/10.1109/COGINF.2003.1225953

Witten, I. H. (2017). *Data Mining (Fourth Edition). Practical Machine Learning Tools and Techniques*.

Wolpert, D. H. (1992). Stacked Generalization. *Elsevier*, *87545*(505), 1–57. Retrieved from https://www.sciencedirect.com/science/article/pii/S0893608005800231%0Apapers://5e3e5e59-48a2-47c1-b6b1-a778137d3ec1/Paper/p2022

Yang, L., Cai, S., Zhang, S., Kong, X., Gu, Y., Lu, C., … Lin, G. (2018). Single embryo transfer by Day 3 time-lapse selection versus Day 5 conventional morphological selection: A randomized, open-label, non-inferiority trial. *Human Reproduction*, *33*(5), 869–876. https://doi.org/10.1093/humrep/dey047

Yilmaz Akşehirli, Ö., Ankaralı, H., Aydın, D., & Saraçlı, Ö. (2013). Tıbbi Tahminde Alternatif Bir Yaklaşım : Destek Vektör Makineleri. *Turkiye Klinikleri J Biostat*, *5*(1), 19–28.

Zegers-Hochschild, F., Adamson, G. D., Mouzon, J. De, Ishihara, O., Mansour, R., Nygren, K., & Sullivan, E. (2009). International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART. *Fertility and Sterility*, *92*(5), 1520–1524. https://doi.org/10.1016/j.fertnstert.2009.09.009

Zegers-hochschild, F., Nygren, K., Adamson, G. D., Mouzon, J. De, Lancaster, P., Mansour, R., & Sullivan, E. (2006). The ICMART glossary on ART terminology *, *21*(8), 1968–1970. https://doi.org/10.1093/humrep/del171

Zhang, C., & Ma, Y. (2012). *Ensemble Machine Learning*. (C. Zhang & Y. Ma, Eds.), *Ensemble Machine Learning*. Springer New York Dordrecht Heidelberg London. https://doi.org/10.1007/978-1-4419-9326-7

Zhu, X., & Goldberg, A. B. (2009). *Introduction to semi-supervised learning*. *Synthesis Lectures on Artificial Intelligence and Machine Learning* (Vol. 6). https://doi.org/10.2200/S00196ED1V01Y200906AIM006

Ziebe, S., Petersen, K., Lindenberg, S., Andersen, A. G., Gabrielsen, A., & Andersen, A. N. (1997). Embryo morphology or cleavage stage: How to select the best embryos for transfer after in-vitro fertilization. *Human Reproduction*, *12*(7), 1545–1549. https://doi.org/10.1093/humrep/12.7.1545

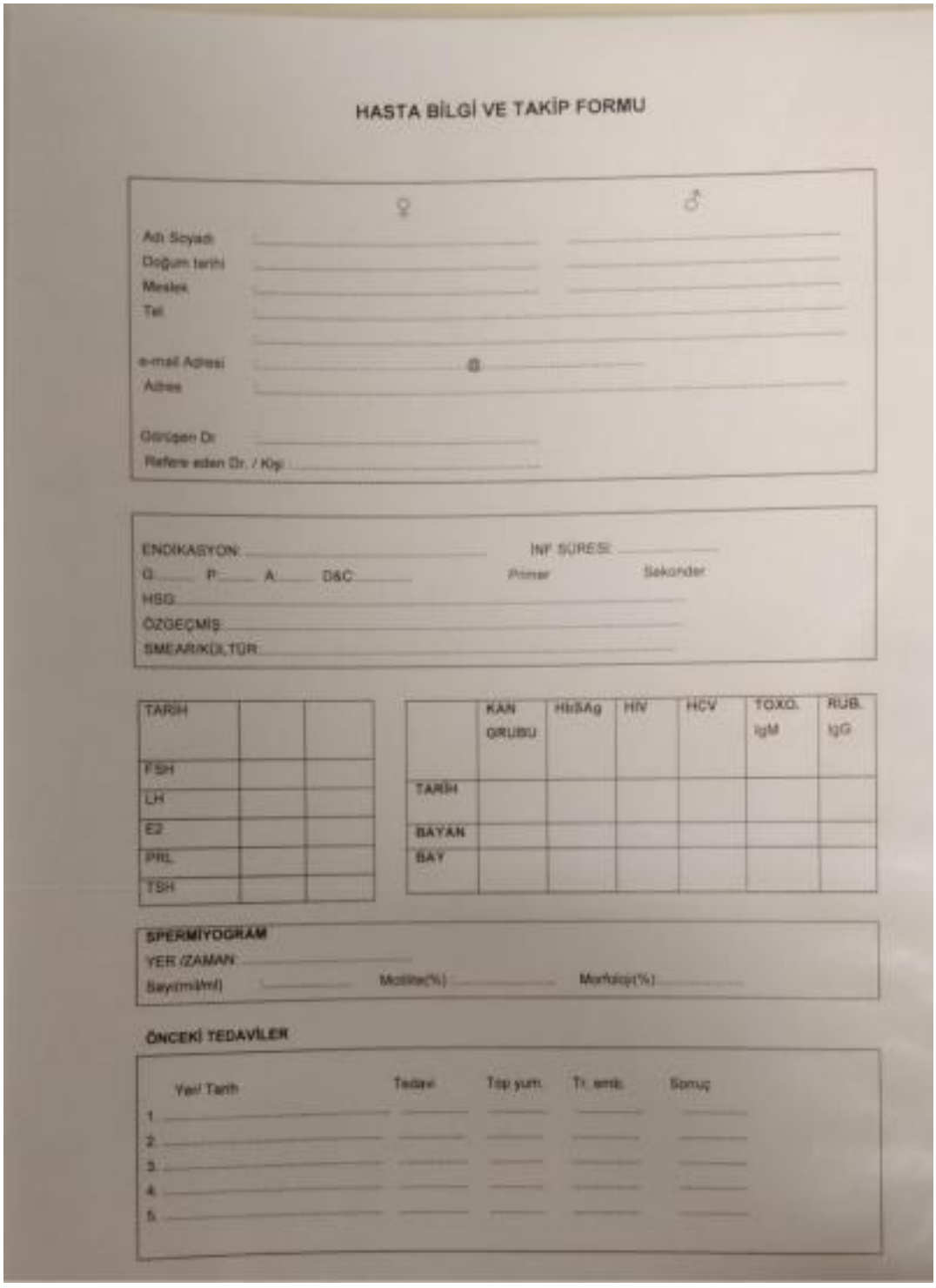
FORMLAR

Ek-1-Korelasyon Analizi

Correlations

	age	bmi	primary	women	men	both	stime	tdose	lnetwo	tfolcount	end	first	total	mtwo	unsferdaysco	ss_ej	sc_normal	motile	prog_motile
age	1																		
bmi	.104**	1																	
primary	-.248**	-.132**	1																
women	.179**	.029	-.074*	1															
men	-.230**	-.045	.055	-.636**	1														
both	.123**	.025	-.011	-.280**	-.520**	1													
stime	-.014	.102**	-.022	.072*	-.146**	.072*	1												
tdose	.318**	.242**	-.139**	.184**	-.215**	.106**	.300**	1											
lnetwo	-.207**	-.047	.130**	-.065*	.011	.025	.082*	-.154**	1										
tfolcount	-.439**	.016	.136**	-.221**	.265**	-.117**	.058	-.291**	.482**	1									
end	-.107**	-.111**	-.076*	-.092**	.164**	-.093**	.039	-.041	.070*	.136**	1								
first	-.212**	-.053	.160**	.069*	-.050	-.013	-.026	-.138**	.130**	.082*	.058	1							
total	-.446**	-.014	.146**	-.242**	.289**	-.145**	.020	-.238**	.506**	.733**	.119**	.081*	1						
mtwo	-.404**	-.010	.162**	-.224**	.277**	-.152**	.028	-.222**	.532**	.667**	.106**	.135**	.890**	1					
unsferdaysco	.021	.025	-.016	.043	-.001	-.023	.010	.021	.090**	.055	.006	-.080*	.043	.035	1				
ss_ej	.049	.012	-.037	.091**	-.133**	.056	-.030	-.013	.009	-.007	-.025	.137**	-.052	-.065*	.026	1			
sc_normal	.088**	-.008	.013	.370**	-.393**	.065*	.037	.031	.058	-.038	-.033	.075*	-.068*	-.060	.014	.177**	1		
motile	-.008	-.005	.022	.129**	-.088**	-.038	.005	-.017	.115**	.065*	.019	.060	-.016	-.013	-.005	-.001	.183**	1	
prog_motile	.004	.047	.073*	.023	-.053	.043	.022	.079*	.059	-.024	-.058	-.048	.022	.020	.014	.102**	.264**	.060	1

Ek-2-Hasta Bilgi Formları



HASTA BİLGİ VE TAKİP FORMU

♀ ♂

Adı Soyadı
Doğum tarihi
Meslek
Tel
e-mail Adresi @
Adres

Gönderen Dr.
Refere eden Dr. / Kişi

ENDİKASYON: INF SÜRESİ:
G.... P.... A.... D&C.... Primer Sekonder
HSG:
ÖZGEÇMİŞ:
SMEAR/KÜLTÜR:

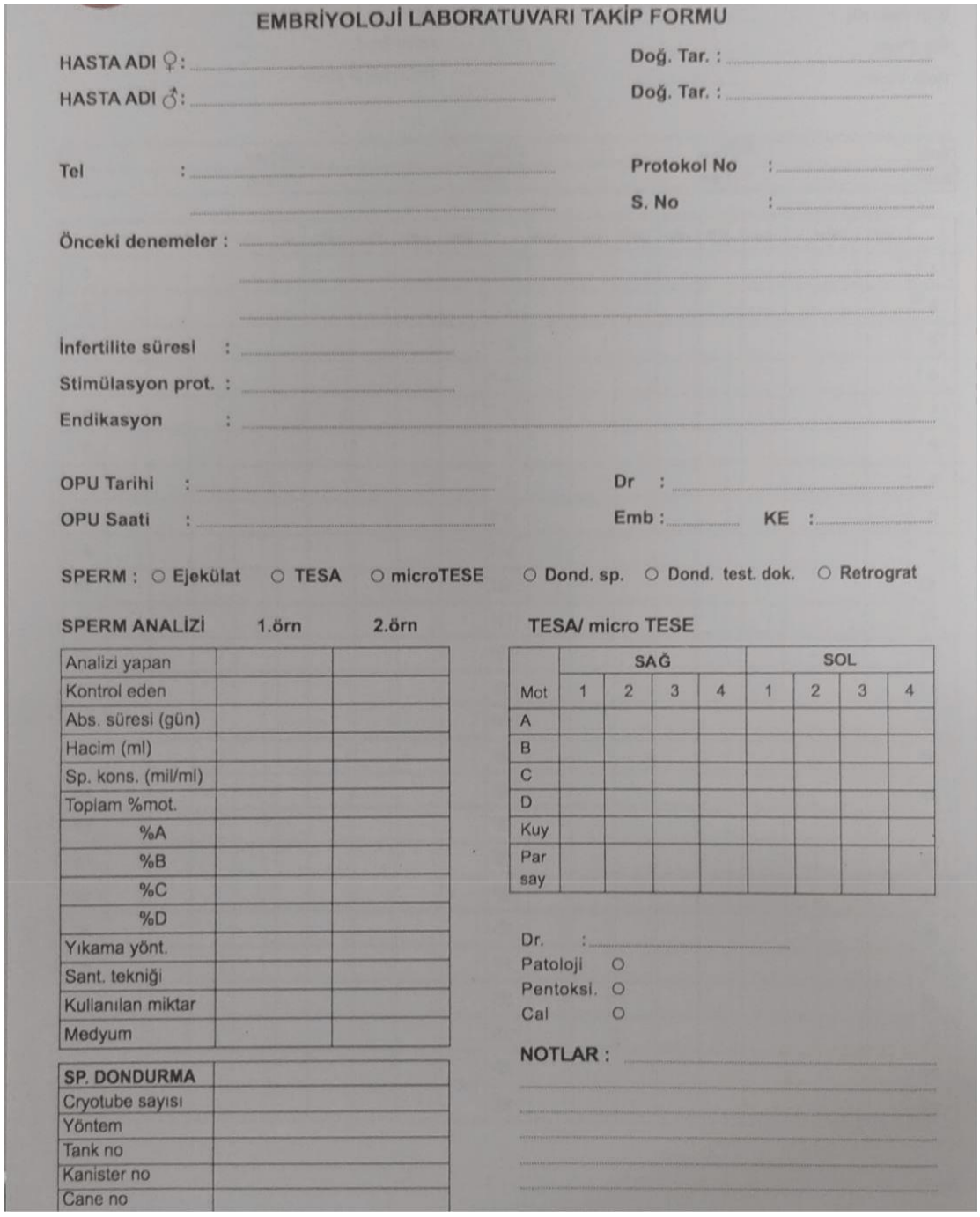
TARİH		
FSH		
LH		
E2		
PRL		
TSH		

	KAN GRUBU	HbSAg	HIV	HCV	TOXO. IgM	RUB. IgG
TARİH						
BAYAN						
BAY						

SPERMİYOGRAM
YER /ZAMAN
Say.(mil/ml) Motilite(%) Morfoloji(%)

ÖNCEKİ TEDAVİLER

	Yer/ Tarih	Tedavi	Top yum.	Tr. emb.	Sonuç
1.					
2.					
3.					
4.					
5.					



EMBRİYOLOJİ LABORATUVARI TAKİP FORMU

HASTA ADI ♀: ..... Doğ. Tar. : .....

HASTA ADI ♂: ..... Doğ. Tar. : .....

Tel : ..... Protokol No : .....

..... S. No : .....

Önceki denemeler : .....

.....

.....

İnfertilite süresi : .....

Stimülasyon prot. : .....

Endikasyon : .....

OPU Tarihi : ..... Dr : .....

OPU Saati : ..... Emb : ..... KE : .....

SPERM : ○ Ejekülat ○ TESA ○ microTESE ○ Dond. sp. ○ Dond. test. dok. ○ Retrograt

SPERM ANALİZİ 1.örn 2.örn

Analizi yapan		
Kontrol eden		
Abs. süresi (gün)		
Hacim (ml)		
Sp. kons. (mil/ml)		
Toplam %mot.		
%A		
%B		
%C		
%D		
Yıkama yönt.		
Sant. tekniği		
Kullanılan miktar		
Medyum		

SP. DONDURMA	
Cryotube sayısı	
Yöntem	
Tank no	
Kanister no	
Cane no	

TESA/ micro TESE

	SAĞ				SOL			
Mot	1	2	3	4	1	2	3	4
A								
B								
C								
D								
Kuy								
Par say								

Dr. : .....

Patoloji ○

Pentoksi. ○

Cal ○

NOTLAR : .....

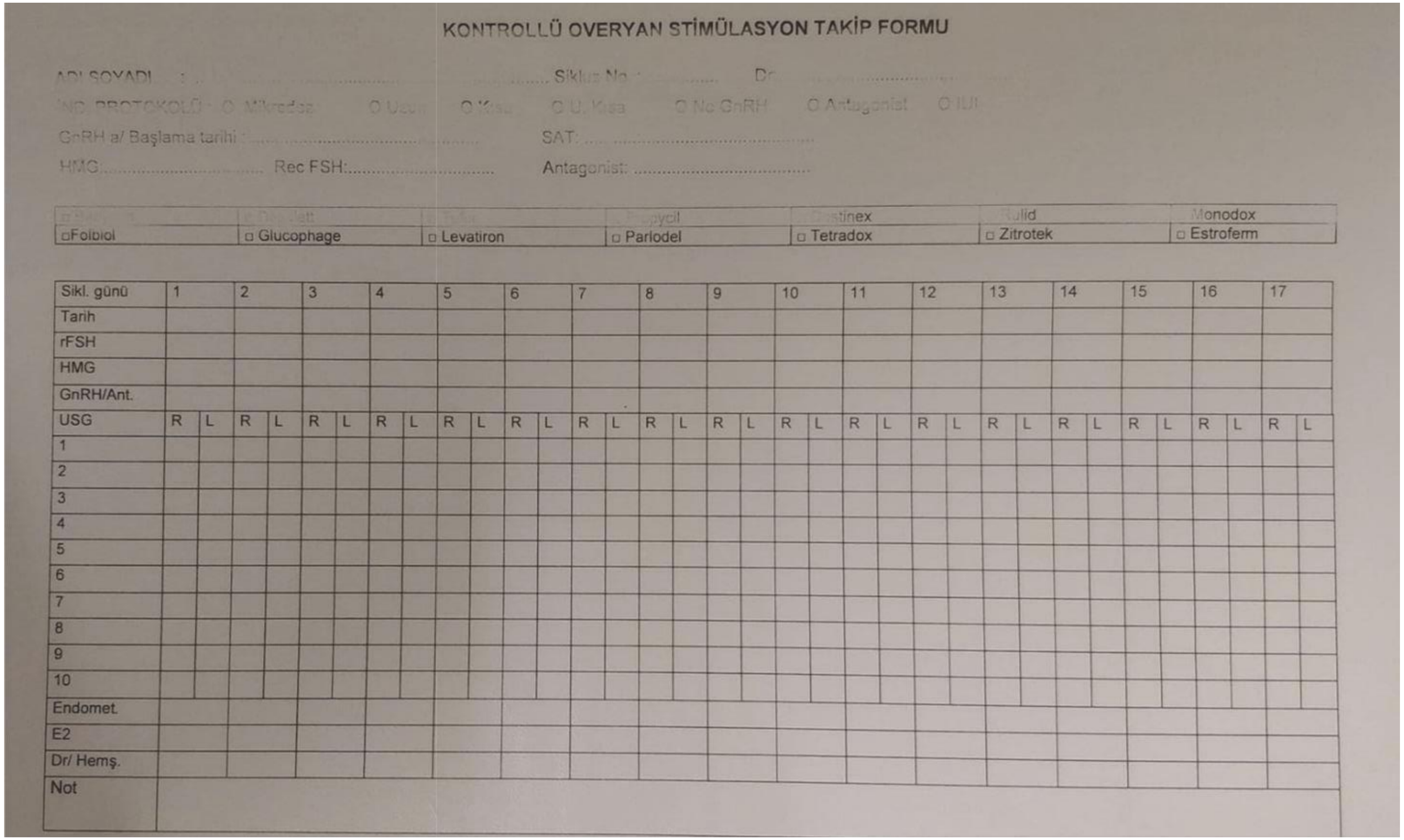
.....

.....

.....

.....

.....



KONTROLLÜ OVERYAN STİMÜLASYON TAKİP FORMU

ADI SOYADI : .... Siklus No : .... Dr ....

NO. PROTOKOLÜ : O Mikrodoz O Uzun O Kısa O U. Kısa O No GnRH O Antagonist O IUI

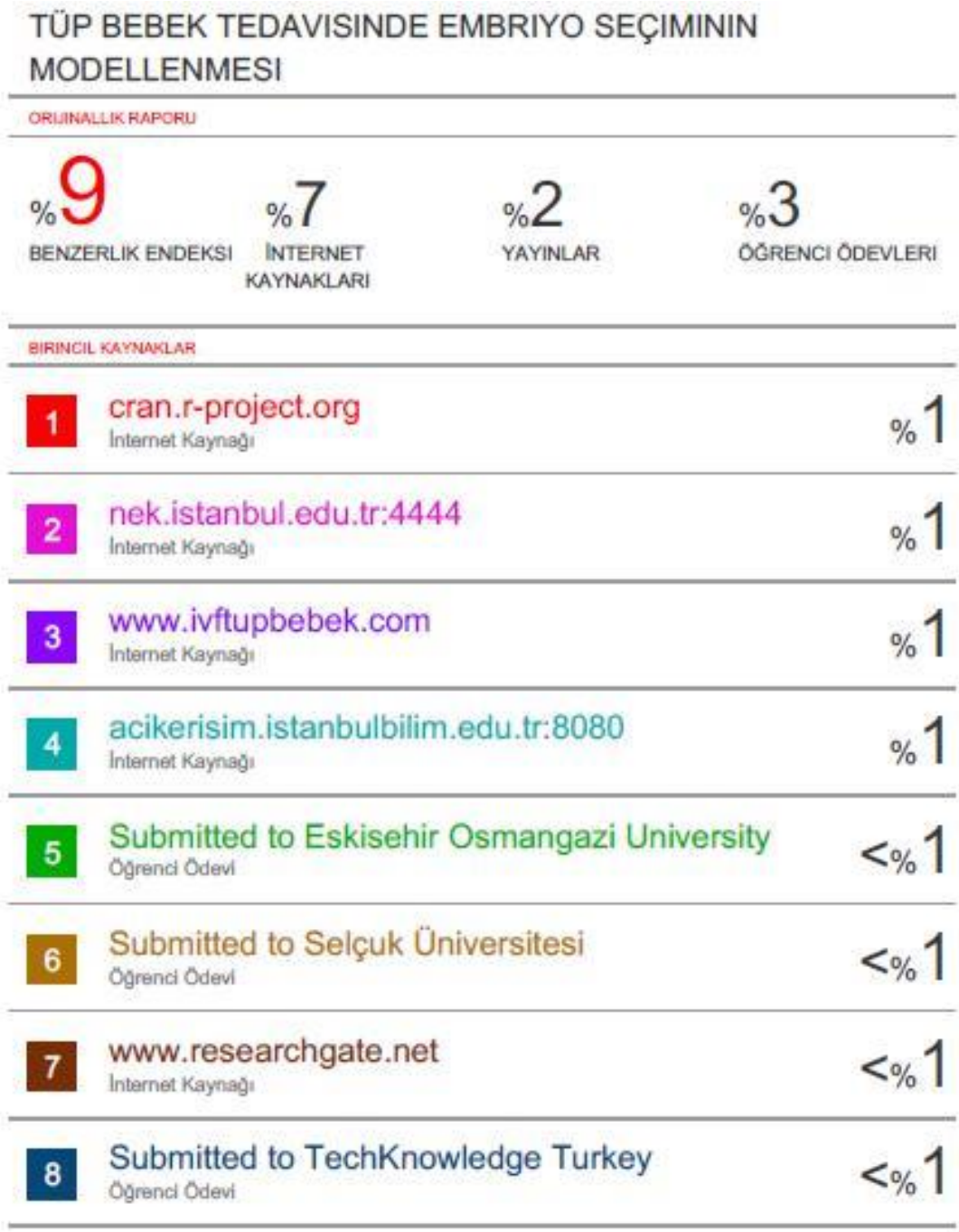
GnRH a/ Başlama tarihi : .... SAT: ....

HMG:.................. Rec FSH:.................. Antagonist: ....

				...stinex	...lid	Monodox
□Folbiol	□ Glucophage	□ Levatiron	□ Parlodel	□ Tetradox	□ Zitrotek	□ Estroferm

Sikl. günü	1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17	
Tarih																																		
rFSH																																		
HMG																																		
GnRH/Ant.																																		
USG	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L
1																																		
2																																		
3																																		
4																																		
5																																		
6																																		
7																																		
8																																		
9																																		
10																																		
Endomet.																																		
E2																																		
Dr/ Hemş.																																		
Not																																		

İNTİHAL RAPORU İLK SAYFASI



TÜP BEBEK TEDAVISINDE EMBRIYO SEÇIMININ MODELLENMESI

ORIJINALLIK RAPORU

%9	%7	%2	%3
BENZERLIK ENDEKSI	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCI ÖDEVLERI

BIRINCIL KAYNAKLAR

1	cran.r-project.org İnternet Kaynağı	%1
2	nek.istanbul.edu.tr:4444 İnternet Kaynağı	%1
3	www.ivftupbebek.com İnternet Kaynağı	%1
4	acikerisim.istanbulbilim.edu.tr:8080 İnternet Kaynağı	%1
5	Submitted to Eskisehir Osmangazi University Öğrenci Ödevi	<%1
6	Submitted to Selçuk Üniversitesi Öğrenci Ödevi	<%1
7	www.researchgate.net İnternet Kaynağı	<%1
8	Submitted to TechKnowledge Turkey Öğrenci Ödevi	<%1

Yayınları/Tebligleri Sertifikaları/Ödülleri

Uluslararası hakemli dergilerde yayınlanan makaleler

Yiğit, P. Kedikli, E., Yılmaz, E., Hayran,O. (2021), Ekonomik Kalkınma ve İşbirliği Örgütü Ülkelerinde Bulaşıcı Olmayan Hastalıklar Risk Faktörlerinin GINI Katsayısı ve Sağlık Düzeyi ile İlişkisi, Turkiye Klinikleri J Health Sci, Doi: 10.5336/healthsci.2020-77423

Yiğit, P., Ataç, Ö., & Aydın, S. (2020). Trends in the incidence rate of genitourinary system cancers in Turkey: 2004–2015. *Turkish Journal of Urology*, *46*(3), 196.

Guven, S., Yigit, P., Tuncel, A., Karabulut, İ., Sahin, S., Kilic, O., ... & Tefik, T. (2020). Retrograde intrarenal surgery of renal stones: a critical multi-aspect evaluation of the outcomes by the Turkish Academy of Urology Prospective Study Group (ACUP Study). *World journal of urology*.2020.

Ozcelik M., Yigit P. (2020) Türkiye Sağlık Sistemi Verimliliğinin İncelenmesi. *Cukurova Medical Journal*, *45*(3), 1-1

Erdem, S. S., Yiğitbaşı, T., Yigit, P., & Emekli, N. (2019). Compliance of medical biochemistry education in medical schools with national core education program 2014. Turkish Journal of Biochemistry, 44(5), 578-584.

Erdem, S. S., Obeidin, V. A., Yigitbasi, T., Tumer, S. S., & Yigit, P. (2018). Verteporfin mediated sequence dependent combination therapy against ovarian cancer cell line. Journal of Photochemistry and Photobiology B: Biology, 183, 266-274.

Büyükuslu, N., Ovalı, S., Altuntaş, Ş. L., Batırel, S., Yiğit, P., & Garipağaoğlu, M. (2017). Supplementation of docosahexaenoic acid (DHA)/Eicosapentaenoic acid (EPA) in a ratio of 1/1.3 during the last trimester of pregnancy results in EPA accumulation in cord blood. Prostaglandins, Leukotrienes and Essential Fatty Acids, 125, 32-36.

Buyukuslu Nihal, Esin Kubra, Hızlı Hilal, Sunal Nihal Yiğit Pakize Garipağaoğlu Muazzez (2014), Clothing preference affects vitamin D status of young females. Nutrition Research, 688 - 693.

Uluslararası bilimsel toplantılarda sunulan ve bildiri kitabında (Proceeding) basılan bildiriler.

Guldemir, H. H., Büyükuslu, N., Çakıcı, Ç., Yiğit, P., Dilsiz, P., & Özdemir, E. M. (2019). OR22: Short Term Effect of Omega Fatty Acids on Intestinal Satiety Hormones and c-Fos Protein Levels in Hypothalamus. Clinical Nutrition, 38, s12.

Esin, K., Şanlier, N., Adal, E., Batırel, S., Ülfer, G., & Yiğit, P. (2017). MON-P056: The Relationship of the Serum Irisin Levels and Anthropometric Measurements of Obese and Non-Obesechildren. Clinical Nutrition, 36, s200.

Ulusal hakemli dergilerde yayınlanan makaleler

Yiğit, P., & Kumru, S. (2019). Türkiye'de 2003-2016 Yılları Arasında Temel Sağlık Göstergelerinin Joinpoint Regresyon Yöntemi ile Analizi. Turkiye Klinikleri Journal of Biostatistics, 11(1).

Aydın S., Yiğit P., Demir M., Güler H. (2013), "Tanı ilişkili gruplama verileri çerçevesinde Türkiye'de ürogenital kanserlere bakış" Yeni Üroloji Dergisi, 8(2), 72-78.

Esin K., Durmaz Ö.,Gökçay E.G., Garipağaoğlu G., Yiğit P.(2015)," Anne Sütü İle Beslenen 0-6 Aylık Bebeklerin Dışkılama Özellikleri", Türkiye Klinikleri Pediatri Dergisi, 10.5336/pediatr.2015-47270

Ulusal Kongrelerde Sunulan Bildiriler

Genç P., Yiğit, P." Sağlık Sektöründe Hizmet Kalite Algısının Değerlendirilmesi: Bir Devlet Hastanesi Örneği", 2. Uluslararası 12. Ulusal Sağlık ve Hastane İdaresi Kongresi, Muğla, Eylül 2018.

Yiğit P., Baytören E., Ersoy Ö. (2014**), "**Sağlıkta Dönüşüm Programı ile Türkiye'nin Sağlık Göstergelerindeki Gelişiminin Çok Değişkenli Tekniklerle Analizi", 8. Sağlık ve Hastane İdaresi Kongresi, Kıbrıs, s. 52-66

Yiğit P., Eren B., Çatalca H. (2013), "OECD VE Dünya Sağlık Örgütü Avrupa Bölgesi Ülkelerinin Sağlık, Eğitim ve Ekonomik Göstergelerinin Karşılaştırılması ve Türkiye'nin Durumu",
7. Sağlık ve Hastane İdaresi Kongresi, Konya,s. 891-907.

Ersoy Ö., Baytören E., Yiğit P. (2012), "Oecd Ülkelerinin Sosyal Ve Ekonomik Göstergeleri Arasindaki Ilişkinin Kanonik Korelasyon Ile Belirlenmesi". 13. Uluslararası Ekonometri, Yöneylem Araştırması ve İstatistik Konferansı.

Özel İlgi Alanları (Hobileri):
Sinema