



TARSUS
ÜNİVERSİTESİ

T.C.

TARSUS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

SEFALOMETRİK RADYOGRAFİLERDEN MAKİNE ÖĞRENMESİ
YÖNTEMLERİ KULLANARAK BÜYÜME EVRESİ TESPİTİ

FAZİLET TÜRKOĞLU

YÜKSEK LİSANS

TARSUS – 2025

T.C.
TARSUS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

SEFALOMETRİK RADYOGRAFİLERDEN MAKİNE ÖĞRENMESİ
YÖNTEMLERİ KULLANARAK BÜYÜME EVRESİ TESPİTİ

FAZİLET TÜRKOĞLU

Danışman: Doç. Dr. Şule YÜCELBAŞ

YÜKSEK LİSANS

TARSUS – 2025

Tarsus Üniversitesi Lisansüstü Eğitim Enstitüsü Müdürlüğüne;

Bu çalışma, jüri tarafından Bilgisayar Mühendisliği Ana Bilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan: Doç. Dr. Şule YÜCELBAŞ
(Danışman)

Üye: Doç. Dr. İrem ERSÖZ KAYA

Üye: Dr. Öğr. Üyesi Betül UZBAŞ

Yukarıdaki Jüri kararı Lisansüstü Eğitim Enstitüsü Yönetim Kurulu'nun/..../2025 tarih ve/.....sayılı kararıyla onaylanmıştır.

ONAY

Yukarıdaki imzaların, adı geçen öğretim elemanlarına ait olduklarını onaylarım.
.../.../2025

Doç. Dr. Ayşe ERGİN ÜNAL
Enstitü Müdürü

NOT: Bu tezde kullanılan ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunu'ndaki hükümlere tabidir.

ETİK BEYANI

Tarsus Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde ve ortaya çıkan sonuçlarda herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 15 / 01 / 2025

İMZA

Fazilet TÜRKOĞLU

ÖZET

SEFALOMETRİK RADYOGRAFİLERDEN MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANARAK BÜYÜME EVRESİ TESPİTİ

FAZİLET TÜRKOĞLU

Yüksek Lisans, Bilgisayar Mühendisliği Ana Bilim Dalı

Danışman: Doç. Dr. Şule YÜCELBAŞ

Ocak 2025, 69 sayfa

Büyüme ve gelişme evreleri sefalometrik radyografi görüntüleri üzerinde tanımlanan servikal vertebra aşamaları ile belirlenebilmektedir. Geleneksel olarak Servikal Vertebra Maturasyon (SVM) evresinin doğru bir şekilde tespit edilmesi, zaman alıcı olabilmekte ve aynı zamanda da uzmanlık gerektirmektedir. Günümüzde yapay zekâ sistemleri, çeşitli algoritmalar kullanarak birçok alanda tahmin ve sınıflandırma gibi farklı görevleri yerine getirebilme kabiliyetine sahiptir. Bu sistemler yüksek doğrulukla teşhis yapabilmekte ve bu şekilde uzmanların işlerini kolaylaştırmakta aynı zamanda hata oranını da en aza indirmektedir. Bu amaçla bu tez çalışmasında sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılmasıyla birlikte makine öğrenmesi algoritmaları kullanılarak SVM evresinin tespiti hedeflenmiştir. Çalışma kapsamında kullanılan veri seti Selçuk Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalı'na başvuran hastaların kayıtlarından elde edilmiştir. Bu kayıtlar kullanılarak oluşturulan veri kümesi, SVM'nin 3 farklı evresine ait toplamda 7-17 yaş aralığında 120 hastadan alınan sefalometrik radyografilerden oluşmaktadır. Çalışmada kullanılan veri setini oluşturmak için evreleri belirlemede önemli olan noktalar uzman diş hekimi tarafından önerilmiş ve bireylerin sefalometrik radyografileri üzerinden 18 adet özellik otomatik olarak tespit edilmiştir (toplamda 120*18 örnek). İlk aşamada, belirlenen 18 özellikli veri seti kullanılıp 5 kat çapraz doğrulama yapılarak, 7 farklı makine öğrenmesi algoritmasıyla sınıflama işlemleri gerçekleştirilmiştir. Bu sınıflandırma sonucunda en iyi doğruluk oranı %82.5 olarak elde edilmiştir. Daha sonra veri setinden 30 adet örnek ayrılarak geriye kalan örnekler Varyasyonel Otomatik Kodlayıcı (VAE) yöntemiyle çoğaltılmıştır. Bu veriler üzerinde sayısal veri standardizasyonu gerçekleştirilmiş ve sonrasında Lasso yöntemi kullanılarak 7 adet etkin özellik belirlenmiştir. Yeni veri seti için de aynı çapraz doğrulama tekniği uygulanmış ve algoritmaların eğitim parametreleri kaydedilmiştir. Elde edilen modellere daha önce ayrılan test veri seti sunulmuş ve nihai performans parametreleri not edilmiştir. Sonuç olarak XGBoost algoritmasıyla tüm metriklerde en iyi sonuçlar elde edilmiş ve %90 doğruluk oranına ulaşılmıştır. Böylece literatüre hem özgün hem de başarılı bir bakış açısı kazandırılmıştır.

Anahtar kelimeler: Sefalometrik analiz, SVM sınıflama, Varyasyonel Otomatik Kodlayıcı (VAE), Lasso, XGBoost.

ABSTRACT

GROWTH STAGE DETECTION USING MACHINE LEARNING METHODS FROM CEPHALOMETRIC RADIOGRAPHS

FAZİLET TÜRKOĞLU

Master Thesis, Department of Computer Engineering

Supervisor: Assoc. Prof. Dr. Şule YÜCELBAŞ

January 2025, 69 pages

Growth and development stages can be determined through the cervical vertebra stages defined on cephalometric radiographic images. Traditionally, accurately identifying the Cervical Vertebra Maturation (CVM) stage can be time-consuming and requires expertise. Today, artificial intelligence systems, utilizing various algorithms, have the capability to perform various tasks such as prediction and classification across multiple fields. These systems can make diagnoses with high accuracy, thus facilitating the work of specialists while minimizing the error rate. The aim of this thesis is to analyze the cervical vertebrae observed in cephalometric radiographs and to detect the stages of CVM using machine learning algorithms. The dataset used in this study was obtained from the records of patients who applied to the Department of Orthodontics, Faculty of Dentistry, Selçuk University. The dataset, created using these records, consists of cephalometric radiographs from 120 patients aged between 7 and 17 years, corresponding to the three different stages of CVM. The key points for determining the stages were suggested by an expert dentist, and 18 features were automatically extracted from the individuals' cephalometric radiographs (a total of 120×18 samples). In the first phase, a classification process was carried out using the dataset with 18 features, applying 5-fold cross-validation and 7 different machine learning algorithms. The best accuracy obtained from this classification was 82.5%. Then, 30 samples were separated from the dataset, and the remaining samples were augmented using the Variational Autoencoder (VAE) method. Afterward, numerical data standardization was performed, and 7 significant features were selected using the Lasso method. The same cross-validation technique was applied to the new dataset, and the training parameters of the algorithms were recorded. The previously separated test dataset was presented to the trained models, and the final performance metrics were noted. As a result, the XGBoost algorithm achieved the best results across all metrics, reaching an accuracy of 90%. Thus, a unique and successful perspective was contributed to the literature.

Keywords: Cephalometric analysis, CVM classification, Variational Autoencoder (VAE), Lasso, XGBoost.

ÖNSÖZ

Günümüzde teknolojik ve bilimsel ilerlemeler, sağlık bilimleri ve ilgili disiplinlerde daha hızlı, doğru, kişiselleştirilmiş tanı ve tedavi yöntemlerini mümkün kılmaktadır. Özellikle büyüme ve gelişme süreçlerinin doğru bir şekilde değerlendirilmesi başarılı tedavi planlarının temelini oluştururken aynı zamanda da yeni teknolojilerin sağlık alanına entegrasyonunu daha da gerekli kılmaktadır. Bu bağlamda SVM evresinin tespiti gibi süreçlerde bu teknolojiler, zamandan tasarruf sağlarken daha doğru ve etkili sonuçlar elde edilmesine olanak tanımaktadır.

Bu yüksek lisans tezi kapsamında, sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılması ve makine öğrenmesi algoritmaları kullanılarak SVM evresinin tespit edilmesi amaçlanmaktadır. Bu amaçla bu tez çalışmasıyla bu alanda karşılaşılan zorluklara çözüm sunmak hedeflenmiştir.

Yüksek lisans ders dönemimden tez yazım sürecine kadar geçen süreçte birlikte çalışma şansı bulduğum, tez konumun belirlenmesinden çalışma sürecimin her aşamasına kadar bilgi ve deneyimleriyle bana yol gösteren, tecrübelerini ve değerli görüşlerini paylaşarak bu yolda gelişimime büyük katkı sağlayan, çalışma sürecim boyunca motivasyon sağlayan yaklaşımları ve sürekli desteğiyle her zaman yanımda olan, bilimsel bakış açısı ve yol gösterici yorumlarıyla çalışmamın gelişimine katkıda bulunan saygıdeğer danışman hocam Sayın Doç. Dr. Şule YÜCELBAŞ' a en derin şükranlarımı sunar ve çok teşekkür ederim. Bu değerli rehberlik ve destek için her zaman minnettar olacağım. Bunun yanında, tez çalışmam için veri alımında desteklerini esirgemeyen değerli uzman diş hekimimiz Sayın Doç. Dr. Hatice KÖK hocama da teşekkürlerimi sunarım.

Ayrıca beni bugünlere getiren, ne olursa olsun her halimle yanımda olan, desteğini ve sevgisini hiç esirgemeyen sevgili anneme, babama ve canım kardeşlerime tüm kalbimle sonsuz minnetlerimi sunuyorum.

İÇİNDEKİLER

Sayfa

ÖZET	i
ABSTRACT.....	ii
ÖNSÖZ	iii
İÇİNDEKİLER	iv
KISALTMALAR	vi
SİMGELER.....	vi
TABLolar LİSTESİ.....	vii
ŞEKİLLER LİSTESİ	viii
GİRİŞ.....	1

BÖLÜM I

KAYNAK ARAŞTIRMASI.....	4
1.1. Büyüme Evresi Tespiti Üzerine Yapılmış Makine Öğrenmesi Çalışmaları	5

BÖLÜM II

MATERYAL VE YÖNTEM.....	10
2.1. Büyüme Gelişmenin Belirlenmesinde Kullanılan Yöntemler.....	11
2.1.1. Radyografik Yöntemler	11
2.1.1.1. Panoramik Radyografiler	11
2.1.1.2. El-Bilek Radyografileri.....	12
2.1.1.3. Sefalometrik Radyografiler.....	12
2.1.1.4. Servikal Vertebra	13
2.2. Veri seti	15
2.3. Yöntem.....	17
2.3.1. Ön İşlemler	17
2.3.2. Veri Çoğaltma	17
2.3.2.1. Varyasyonel Otomatik Kodlayıcılar	17

2.3.2.2. Kayıp Fonksiyonu.....	19
2.3.2.3. Kodlama Sürecinde Yeniden Parametrelendirme Hilesi	20
2.3.3. Veri Standardizasyonu.....	22
2.3.4. Özellik Seçimi	22
2.3.4.1. Lasso Özellik Seçme Yöntemi.....	23
2.4. Makine Öğrenmesi	24
2.5. Kullanılan Algoritmalar	24
2.6. İstatistiksel Değerlendirme Ölçütleri	27
BÖLÜM III	
BULGULAR VE TARTIŞMA	30
3.1. Ham Veriyle Yapılan Sınıflama İşlemleri.....	30
3.2. VAE Yöntemi ile Örnek Çoğaltma ve Özellik Seçme İşlemleri Sonrası Yapılan Sınıflama İşlemleri	31
SONUÇ VE ÖNERİLER.....	47
KAYNAKÇA	49
ÖZGEÇMİŞ	56

KISALTMALAR

AAv	: Ağırlıklandırılmış Averaj
DVM	: Destek Vektör Makinesi
EYKA	: K-En Yakın Komşu Algoritması
FN	: Yanlış Negatif
FP	: Yanlış Pozitif
KL	: Kullback-Leibler
LR	: Lojistik Regresyon
MAv	: Makro Averaj
N	: Toplam Örnek Sayısı
RO	: Rastgele Orman
SVM	: Servikal Vertebra Maturasyon
TN	: Doğru Negatif
TP	: Doğru Pozitif
VAE	: Varyasyonel Otomatik Kodlayıcı
XGBoost	: Ekstra Gradyan Artırma
YSA	: Yapay Sinir Ağı

SİMGELER

x	: Veri Giriş Örneği
x'	: Model Tarafından Yeniden Üretilen Veri Örneği
μ	: Ortalama Değer
σ	: Standart Sapma
N(0,1)	: Standart Normal Dağılım
Loss_{ori}	: Modelin Temel Kayıp Fonksiyonu
λ	: Regularizasyon Parametresi
w_i	: Modelin i'nci Ağırlık Parametresi
$\sum w_i$: Modeldeki Tüm Ağırlıkların Mutlak Değerlerinin Toplamı

TABLolar LİSTESİ

Sayfa

Tablo 2.1. Tez Çalışması İçin Oluşturulan Veri Setine Ait Özellikler	16
Tablo 3.1. Elde Edilen 18 Özellik ile Yapılan Ortalama Sınıflandırma Sonuçları	30
Tablo 3.2. Elde Edilen Birinci Kat Sonuçları	32
Tablo 3.3. Elde Edilen İkinci Kat Sonuçları	34
Tablo 3.4. Elde Edilen Üçüncü Kat Sonuçları	36
Tablo 3.5. Elde Edilen Dördüncü Kat Sonuçları.....	38
Tablo 3.6. Elde Edilen Beşinci Kat Sonuçları.....	40
Tablo 3.7. Başlangıçta Ayrılan 30 Örneğin Eğitilmiş XGBoost Modeliyle Sınıflandırılma Sonuçları	43

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1. Çalışmaya Ait İş Akış Şeması	10
Şekil 2.2. Baccetti ve Ark. Tarafından Yenilenmiş Olan SVM Aşamaları ve Radyografik Görünümü	14
Şekil 2.3. SVM Evresi Tepitinde Kullanılan Sefalometrik Referans Noktaları	14
Şekil 2.4. Büyüme ve Gelişme Tespitinde Dönemler	15
Şekil 2.5. VAE ile Standart Otomatik Kodlayıcılar Arasındaki Fark	18
Şekil 2.6. Yeniden Parametrelendirme Hilesi Uygulamasının Yapıyı Nasıl Dönüştürdüğünün Gösterimi.....	20
Şekil 2.7. VAE Modeline Ait Sözde Kod.....	21
Şekil 2.8. XGBoost Algoritmasına Ait Sözde Kod	27
Şekil 3.1. Tüm Sınıflandırıcıların Doğruluk Oranları	31
Şekil 3.2. Birinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi	33
Şekil 3.3. Birinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi.....	33
Şekil 3.4. Birinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları.....	34
Şekil 3.5. İkinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi	35
Şekil 3.6. İkinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi.....	35
Şekil 3.7. İkinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları.....	36
Şekil 3.8. Üçüncü Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi	37
Şekil 3.9. Üçüncü Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi.....	37
Şekil 3.10. Üçüncü Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları...	38

Şekil 3.11. Dördüncü Kat Sonuçlarına Göre LR Sınıflayıcısına Ait Karmaşıklık Matrisi	39
Şekil 3.12. Dördüncü Kat Sonuçlarına Göre LR Sınıflayıcısına Ait ROC Eğrisi.....	39
Şekil 3.13. Dördüncü Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları	40
Şekil 3.14. Beşinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi	41
Şekil 3.15. Beşinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi.....	41
Şekil 3.16. Beşinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları.....	42
Şekil 3.17. Tüm Sınıflandırıcıların 5-Kat Çapraz Doğrulama Sonucundaki Ortalama Performans Çıktıları.....	43
Şekil 3.18. XGBoost'un Başlangıçta Ayrılan 30 örnekli Veri Setini Sınıflamasına Ait Karmaşıklık Matrisi	44
Şekil 3.19. XGBoost'un Başlangıçta Ayrılan 30 Örnekli Veri Setini Sınıflamasına Ait ROC Eğrisi.....	44

GİRİŞ

Büyüme ve gelişme evreleri, birçok tedavinin planlamasında belirleyici bir faktördür. Çocuk sağlığı, adli tıp, kemik ve eklem sağlığı, hormon bozuklukları ve ağız-diş sağlığı gibi farklı tıbbi alanlarda bu evreler önemli rol oynamaktadır (Kök vd.,2020). Bu evrelerin doğru belirlenmesi, önemli alanlarda bireyin yaş, gelişim ve sağlık durumlarıyla ilgili doğru değerlendirmeler yapılmasına yardımcı olmaktadır. Bireylerin büyüme ve gelişmesinde kronolojik yaş, adet görme, ses değişiklikleri, boy uzunluğundaki artış ve kemik olgunlaşması gibi parametreler büyüme ve gelişme sürecinin takibi için önemlidir (Franchi vd.,2000; Hägg ve Taranger,1982; İzgi ve Kök,2020). Ortodontik tedaviler için de büyüme ve gelişme evrelerinin belirlenmesi, bu tedavilerin etkinliği ve başarısı açısından kritik öneme sahiptir (İzgi ve Kök,2020). Çünkü belirlenen doğru evrelere göre kişiye özel tedavi planlamasının yapılması, tedavinin etkinliğini artırılması ve uzun vadeli başarının elde edilmesi açısından avantaj sağlamaktadır. Hangi yaşta hangi tedavi yönteminin uygulanacağı, büyüme evresine göre farklılık gösterebilmektedir. Bu sebeple alanında uzman hekimler, verilerden elde edilen bilgileri kullanarak kişiye özel tedavi planları oluşturabilmektedir.

Büyüme ve gelişme evresi tespiti sayesinde, bir bireyin büyüme ve gelişme sürecinin hangi aşamada olduğu belirlenebilmektedir (İzgi ve Kök,2020). Büyüme-gelişim evreleri sefalometrik radyografi görüntüleri üzerinde tanımlanan servikal omur aşamaları ile tespit edilebilmektedir (Fudalej ve Bollen,2010). Ayrıca ortodontide bu evrelerin tespiti için panoramik ve el bileği radyografileri de tercih edilmektedir (Fishman,1982; Flores-Mir vd.,2004). Öte yandan sefalometrik radyografilerle servikal omurların analizi büyüme-gelişim evrelerini belirleme konusunda daha kullanışlı bir yol olarak görülmekte ve aynı zamanda radyasyona maruz kalma riskini de azaltmaktadır (Baccetti vd.,2005; Flores-Mir vd.,2006). Büyüme-gelişim ve kemik yaşı belirlemede el-bilek radyografilerinin kullanılmasının ek radyasyon dozu gerektirmesi nedeniyle, ortodontik tedavi sırasında rutin olarak çekilen sefalometrik radyografilerin bu amaçla kullanılması önerilmektedir (Gençay ve Kasımoğlu,2016).

Servikal vertebra, omurganın boyun bölgesinde bulunan omurlarına verilen isimdir (İzgi ve Kök,2020). Servikal Vertebra Maturasyonu (SVM) ortodonti ve büyüme-gelişim çalışmalarında sıklıkla başvuru alan, bireylerin iskeletsel büyüme süreçlerinin takibi ve analiz edilmesi amacıyla kullanılan son derece önemli bir değerlendirme aracıdır (Hassel ve Farman,1995). Bu araç, genellikle çocuklarda ve

gençlerde boyun omurlarının kemiksel gelişimini belirleyerek omurganın olgunlaşma durumunu takip etmek için kullanılmaktadır.

Büyüme-gelişme evrelerinin doğru belirlenmesi ortodontik tedavi planlamasında kritik bir rol oynarken bu sürecin güvenilirliğini arttırmak amacıyla da çeşitli yöntemler geliştirilmiştir. Bu bağlamda, Lamparski'nin servikal omurların büyüme-gelişme sürecindeki rolünü açıklaması, bu alandaki önemli bir dönüm noktası olmuştur. Büyüme-gelişmenin servikal omurlarından tespit edilebileceğini ilk kez Lamparski dile getirmiştir. Bu hipotezden yola çıkılarak, daha sonra bu yöntemin güvenilirliği ve geçerliliği üzerine odaklanılmıştır (Grave ve Townsend,2003). Lamparski'nin çalışması el-bileği yöntemiyle birleştirilerek ve farklı coğrafi bölgelerde yapılan araştırmalarla desteklenerek, büyüme evrelerinin belirlenmesinde kullanılıp doğrulanmıştır (Grave ve Townsend,2003; Lamparski,1972). Yapılan bu çalışmalar büyüme-gelişme evresi tespitinde farklı metodolojilerin geliştirilmesi, kullanılan yöntemlerin doğruluğunun ve güvenilirliğinin değerlendirilmesi konularında bilimsel altyapıyı geliştirmiş ve birçok klinik uygulamaya katkı sağlamıştır. Ancak büyüme ve gelişme evresi tespiti alanında hala karşılaşılan bazı önemli eksiklikler ve boşluklar bulunmaktadır. Bu eksiklikler büyüme ve gelişme süreçlerini etkileyen birçok faktörün etkileşimiyle ortaya çıkmakta ve bu eksiklikler alanın daha kapsamlı bir şekilde gelişmesini sınırlamaktadır. Bu eksiklikler arasında standart olmayan değerlendirme yöntemleri, veri yetersizliği, teknolojik altyapı ve disiplinler arası yaklaşımların eksikliği konuları yer almaktadır. Mevcut yöntemlerin standardizasyonunun eksikliği, farklı uzmanlar arasında sonuçların karşılaştırılmasını zorlaştırabilmekte ve bu durum da tedavi kararlarının tutarlılığını etkileyebilmektedir. Bunun yanında, bu faktörlerin tamamını içeren geniş kapsamlı veri setleri genellikle eksiktir. Teknolojik altyapı açısından gelişmiş görüntüleme ve veri analizi teknolojilerine de erişim kısıtlı olabilmektedir. Bununla birlikte disiplinler arası bir yaklaşımın eksikliği büyüme ve gelişme evresi tespiti alanında kapsamlı bir değerlendirmenin yapılmasını, daha etkili ve doğru stratejilerin geliştirilmesini engelleyebilmektedir. Görüldüğü üzere, belirtilen boşlukların giderilmesi için standartlaştırılmış yöntemlerin oluşturulması, daha geniş kapsamlı veri toplama ve analiz yöntemlerinin benimsenmesi, uzmanlar arasında daha fazla iş birliği ve bilgi paylaşımı teşvik edilmesi gibi adımlar atılmalıdır. Geleneksel olarak büyüme ve gelişme evresi tespiti zaman alıcı olabilmekte ve aynı zamanda uzmanlık gerektirmektedir. Makine öğrenimi büyüme ve gelişme evrelerini otomatik ve hızlı bir şekilde belirleyerek zaman ve maliyet tasarrufu sağlayabilmektedir. Geliştirilecek teknolojiler

hastaların daha iyi ve kişiye özel tedavi imkânlarına erişimini artırabilirken aynı zamanda sağlık profesyonellerine de daha doğru ve etkili bir çalışma ortamı sunabilecektir (İzgi ve Kök,2020). Bu mevcut eksikliklerin üstesinden gelinerek büyüme ve gelişme evresi tespiti alanında daha etkili ve güvenilir bir değerlendirme süreci sağlanabilecektir.

Bu tez çalışmasında, sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılmasıyla alana özgü olarak seçilmiş makine öğrenme algoritmaları kullanılarak SVM evresi tespiti gerçekleştirilmiştir. Bu işlemler esnasında literatür açısından önemli kazanımlar sağlayacak hassas özellik çıkarma ve işleme adımları izlenmiştir. Bu sayede ortodontik tedavi süreçlerinin ve planlamalarının daha etkin bir şekilde gerçekleştirilmesine katkıda bulunulması amaçlanmıştır. Çalışma adımlarında sefalometrik radyografilerde servikal vertebra analizi üzerine odaklanılarak, sınıflandırma alanında etkin olarak kullanılan makine öğrenmesi algoritmalarının kullanımıyla SVM evresinin yüksek doğrulukla tespit edilmesi üzerinde durulmuştur. Veri setinin oluşturulması için ilk aşamada, alanında uzman diş hekimi Doç. Dr. Hatice KÖK tarafından radyografiler üzerinde önerilen önemli noktalar belirlenmiş ve bu noktalar kullanılarak önemli özellikler otomatik olarak hesaplanmıştır. Daha sonra veri çoğaltma işlemi gerçekleştirilerek veri setinin çeşitliliği artırılmış ve modelin daha geniş veri setleri ile eğitilmesi sağlanmıştır. Bunun yanı sıra, Lasso yöntemiyle özellik seçimi yapılmış ve modelin performansını arttıran daha etkili özellikler belirlenmiştir. Bu yöntemlerin etkinliği istatistiksel performans kriterleri ile değerlendirilmiş ve sınıflama sonuçlarına olan etkisi test edilmiştir. Çalışmanın ilerleyen aşamalarında kullanılan algoritmaların performansı ayrıntılı olarak incelenmiş ve ulaşılan çıktılar sunulmuştur. Son kısımda da elde edilen bulguların tartışılması ve çalışmanın sınırlılıkları ile gelecekteki araştırmalar için önerilerde bulunulmuştur.

BÖLÜM I

KAYNAK ARAŞTIRMASI

Bu bölümde, büyüme evresinin tespitinde etkin bir şekilde kullanılan makine öğrenmesi yöntemleriyle gerçekleştirilen sefalometrik radyografilerdeki servikal vertebra analizi üzerine yapılan araştırmaların özetleri ve elde edilen başarı sonuçları sunulmaktadır. SVM, ortodonti ve büyüme-gelişim çalışmalarında sıkça başvurulan, bireylerin iskeletsel büyüme süreçlerinin takibi ve analiz edilmesi amacıyla kullanılan son derece önemli bir değerlendirme aracıdır (Gu ve McNamara Jr,2007). SVM analizi servikal omurların belirli yapısal özelliklerini ve bu omurların zamanla geçirdiği değişiklikleri inceleyerek bireylerin büyüme potansiyellerini ve iskeletsel gelişim süreçlerinin hangi aşamasında bulduklarını belirlemek için uygulanmaktadır. SVM evrelerinin otomatik sınıflandırılması ortodonti alanında önemli bir potansiyele sahiptir. Bu alandaki makine öğrenmesi çalışmaları servikal vertebra görüntülerinin hızlı ve doğru bir şekilde analiz edilmesini sağlayarak ortodontik tedavi planlamalarını daha etkili hale getirmiştir. Bu bağlamda SVM evrelerinin otomatik sınıflandırılması hem zaman tasarrufu sağlar hem de insan hatasını minimize eder ve böylece tedavi sürecinin doğruluğu ve güvenilirliği artmaktadır. SVM özellikle ergenlik döneminde gerçekleşen hızlı büyüme evrelerinin zamanlamasını ve bu evrelerin yoğunluğunu doğru bir şekilde belirlemek amacıyla kullanılmaktadır (Grave ve Townsend,2003). Bu analiz yöntemi daha çok servikal omurların röntgen görüntüleri üzerinden yapılmakta olup omurların şekli, boyutu, konkavite varlığı ve diğer morfolojik değişiklik özellikleri göz önünde bulundurularak gerçekleştirilir (Oflaz ve Görgülü,2017). Ortodontik tedavi planlamasında oldukça önemli bir yere sahip olan bu yöntem, iskeletsel büyüme-gelişme ile ilgili olarak daha bilinçli ve isabetli kararlar alınmasına olanak tanır ve bu sayede tedavi sürecinin zamanlaması ve uygulanacak tedavi yöntemleri en verimli şekilde belirlenebilir. SVM analizi, genellikle altı farklı evrede değerlendirilir ve bu evreler bireyin iskeletsel olgunluk seviyesini ve büyüme potansiyelini yansıtır (El-Askary vd.,2012; Robbins,1991). Bu kapsamda servikal vertebra incelenmesiyle elde edilen veriler, büyüme-gelişim süreçlerinin takibi ve ergenlik döneminin başlangıç ve bitişi gibi kritik dönemlerin belirlenmesi için önemli bir bilgi kaynağı sağlamaktadır.

1.1. Büyüme Evresi Tespiti Üzerine Yapılmış Makine Öğrenmesi Çalışmaları

Büyüme-gelişme evrelerinin doğru bir şekilde tespit edilmesi hem klinik uygulamalar hem de araştırmalar için kritik bir öneme sahiptir. Yapay zekâ, çeşitli algoritmalar kullanarak birçok yaşam alanında tahmin ve sınıflandırma gibi çeşitli aktiviteleri gerçekleştirme yeteneğine sahiptir. Bu alanda makine öğrenmesi tekniklerinin kullanılmaya başlanması bu sürecin doğruluğunu ve etkinliğini artırma potansiyelini de beraberinde getirmiştir. Makine öğrenmesi yöntemleri büyük veri setlerini analiz edebilme ve karmaşık desenleri tanıyabilme yeteneği sayesinde bireylerin büyüme-gelişme evrelerini daha hızlı ve doğru bir şekilde belirleme imkânı sunmaktadır.

Lamparski, büyüme-gelişme ve kemik yaşının tespitinde el-bilek radyografilerinin yanı sıra sefalometrik radyografilerdeki servikal vertebralardan da faydalanılabileceğini, ilk kez 1972 yılında yaptığı yüksek lisans tezinde açıklamıştır (Lamparski ve Nanda,2002). Bu tezden yola çıkarak, servikal vertebraların radyografik analizine dayalı olarak iskeletsel olgunlaşma düzeyini değerlendirme yöntemini geliştirmiştir. Bu yöntemle Lamparski, servikal vertebralar kullanılarak gerçekleştirilen büyüme-gelişim ve yaş tespitinin el-bilek radyografileri kadar kararlı ve tutarlı olduğunu savunmuştur (Lamparski ve Nanda,2002). Yaptığı çalışma, büyüme-gelişme süreçlerini daha doğru bir şekilde izlemek ve değerlendirmek için ortodonti alanına önemli bir yenilik getirmiştir. Çalışmaya göre, veri kümesi içerisindeki 500 hastadan seçilen 10-15 yaş arası 69 erkek ve 72 kız hastayı seçerek C2'den C6'ya kadar olan beş vertebrayı ele almış ve vertebra olgunlaşma farklılıklarını 6 evreye ayırmıştır. Lamparski'nin tanımladığı SVM aşamalarına göre, birinci aşamada tüm vertebra gövdelerinin alt kenarları düz olmakla birlikte üst kenarları arka kısımdan ön kısma doğru eğimlidir. İkinci aşamada servikal vertebranın alt kenarındaki konkavite belirginleşmiş ve vertebra gövdelerinin ön kısmındaki yükseklikler artmıştır. Üçüncü aşamada üçüncü vertebranın alt kenarında konkavite daha belirgin hale gelmiştir. Dördüncü aşamada C3'ten C6'ya kadar olan tüm servikal vertebraların gövdeleri dikdörtgen biçimini almıştır ve dördüncü vertebranın alt kenarında konkavite oluşmuştur; C5 ve C6'da ise yeni oluşan konkaviteler gözlemlenir. Beşinci aşamada vertebra gövdeleri kareye benzeyen bir şekil almaya başlar ve vertebralar arasındaki boşluklar azalmaktadır. Son aşama olan altıncı aşamada ise alt kenardaki konkaviteler

tüm servikal vertebralarda belirginleşir. Vertebra gövdelerinin dikey yüksekliği artmaktadır ve konkaviteler daha derin hale gelmektedir (Lamparski ve Nanda,2002).

Bu çalışmadan sonra bu yöntemin güvenilirliği ve geçerliliği üzerine odaklanılmıştır (Grave ve Townsend,2003). Lamparski'nin çalışması el bileği yöntemiyle birleştirilerek ve farklı coğrafi bölgelerde yapılan araştırmalarla desteklenerek, etkinliği büyüme-gelişme evrelerinin belirlenmesinde doğrulanmıştır (Grave ve Townsend,2003; Lamparski,1972). Hassel ve Farman yaptıkları çalışmalarında iskelet olgunlaşmasının sürekli bir süreç olduğunu belirterek Lamparski'nin yöntemini değiştirip servikal vertebra olgunlaşmasını altı aşamada sınıflandıran bir yöntem geliştirmişlerdir (Hassel ve Farman,1995). Bu çalışma sonucunda kemik yaşının tespitinde servikal vertebraların analizinin el-bilek kemikleriyle aynı derecede doğru ve tutarlı bir yöntem olduğunu bildirmişlerdir. Franchi ve ark. 2000 yılında 34 kişilik bir grupta Lamparski'nin orijinal yönteminin geçerliliğini test etmiştir (Franchi vd.,2000). Baccetti ve ark., 2002 yılında 30 hastadan oluşan bir grup üzerinde yaptıkları çalışmada, 6 aşamalı servikal vertebra analizine alternatif olarak, tiroidi koruyan yaka ile engellenmeyen 2., 3. ve 4. servikal vertebrayı kapsayan 5 aşamalı bir sınıflama önerisi sunmuşlardır (Baccetti vd.,2002). Aşamalarda 1. ve 2. evrelerin ayırt edilmesinin zor olduğunu ve bu evrelerin tek bir aşama olarak birleştirilmesinin uygun olabileceğini ifade etmişlerdir. Daha sonra 2005 yılında kendi yöntemlerini yeniden düzenleyerek C2, C3 ve C4'ün net bir şekilde görüldüğü lateral sefalogramları değerlendirmiş ve servikal vertebralar temelinde büyüme-gelişim evrelerini 6 aşamada incelemişlerdir (Baccetti vd.,2005). Bu yöntemle SVM ile iskeletsel maturasyon arasındaki bağlantının daha doğru bir şekilde tespit edilebildiğini ve bu metodun daha kullanışlı olduğunu belirtmişlerdir. Ayrıca bu modifikasyon ile SVM yönteminin daha geniş bir uygulama alanına sahip olmasını sağlamışlardır. Bu yaklaşıma göre, bireylerin büyüme-gelişimlerinin belirlenmesi için sefalometrik radyografilerde ikinci (CS2), üçüncü (CS3) ve dördüncü (CS4) servikal vertebra segmentlerinin değerlendirildiği ve altı aşamaya ayrıldığı yer almaktadır. Bu evrelerin başlangıcı servikal vertebra birinci aşama (CS1) olarak bilinen ilk evredir ve ergen gelişiminin başladığı dönemi ifade etmektedir. Burada ilk evrede C2, C3 ve C4 ise yayvan dörtgen şeklindedir ve üst kenarları görüldüğü üzere öne doğru eğimlidir. İkinci aşama (CS2) ergen büyümesinin hızlandığı aşamadır. Burada C2 ve C3'ün alt kenarlarında bir iç bükeylik başlarken C3 ve C4 dikdörtgen bir görünüm kazanmaya başlar. Ergen büyüme artışının yaşandığı değişim evresi CS3 olarak tanımlanır. Bir

önceki evreye göre kenarlardaki iç bükeylik C2 ve C3'te belirginleşir. C4'ün alt kenarında ise oyulmalar başlar, C3 ve C4 dikdörtgen şekillerini alır. CS4 aşamasına ulaşıldığında ergenlik büyümesinin önemli ölçüde yavaşladığı fark edilir. C2, C3 ve C4'ün alt kenarlarında çukurluklar belirginleşirken C3 ve C4 kare şeklinde görünmeye başlar. Ergenlik büyümesinin büyük ölçüde dikkate alınmadığı dönem olgunluk evresi CS5'tir. Bu dönemde C2, C3 ve C4'ün alt kenarlarında bulunan oyuklar daha belirgin hale gelirken C3 ve C4 kare şekline dönüşür. CS6 aşaması, ergen büyümesinin bitişini işaret eden tamamlanma evresidir. Bu aşamada büyüme gözlenmez. C2, C3 ve C4'ün alt kenarlarında bulunan çukurluklar belirgin bir şekilde derinleşirken C3 ve C4 kare formuna girer.

Yapılan bu araştırmalar, puberte (ergenlik) dönemindeki iskeletsel gelişimi incelemek için servikal vertebraların analiz edilmesinin el-bilek kemiklerinin değerlendirilmesiyle büyük oranda uyumlu olduğunu ve bu tekniğin büyüme sürecini izlemede güvenilir bir gösterge olduğunu ortaya koymuştur.

Bu alanda gerçekleştirilen başka bir çalışmada Yapay Sinir Ağı (YSA) kullanılarak servikal vertebralar üzerinden büyüme-gelişim dönemleri ve cinsiyetin tespit edilmesi hedeflenmiştir. Çalışmada kullanılan veri kümesi yaşları 8 ile 17 arasında olan 419 hastanın sefalometrik ve el-bilek radyografilerini içermektedir. Sefalometrik radyografiler her biri farklı bir büyüme-gelişim aşamasını temsil eden altı ayrı servikal vertebra evresine göre gruplandırıldı. Servikal vertebra verileriyle yapılan büyüme-gelişim dönemleri belirlemede en iyi sonuç %94,27 ve cinsiyet belirlemede ise %89,50 doğruluk elde edilmiştir (Kök vd.,2021). Aynı araştırmacılar tarafından yapılan benzer bir çalışmada da çeşitli servikal vertebra oranları kullanılarak eğitilmiş YSA ve Naive Bayes modellerinin büyüme-gelişim dönemlerini belirlemedeki başarılarını değerlendirmeyi amaçlamışlardır. Sefalometrik radyografileri çekilmiş 8 ile 17 yaşları arasındaki 360 birey üzerinde gerçekleştirilen bu çalışmada, sefalometrik radyografilerin değerlendirilmesi sonucunda büyüme-gelişim dönemleri 6 vertebral evreye ayrılmıştır. Her evre bir grup olarak ele alınmış ve her grup 30 kız ve 30 erkekte oluşmuştur. 28 servikal vertebra oranı on yatay ve 13 dikey ölçümle elde edilerek dört farklı kombinasyonda birleştirilmiştir ve her biri eğitim ve test olarak ikiye ayrılmıştır aynı zamanda aşırı öğrenmeyi önlemek için eğitim aşamasında 5-kat çapraz doğrulama tekniği kullanılmıştır. Sonuçta en yüksek belirleme başarı oranı YSA 3 modelinde %95 elde edilirken, en düşük başarı oranı NBM 4 modelinde %50 bulunmuştur (Kök vd.,2020). Yine aynı araştırmacılar tarafından aynı amaç

doğrultusunda sunulan benzer başka bir çalışmada da servikal vertebra aşamalarını belirlemek için yedi yapay zekâ algoritmasının sonuçlarını karşılaştırmışlardır. Çalışma kapsamında kullanılan veri kümesinde, yaşları 8 ile 17 arasında değişen 300 bireyden elde edilen sefalometrik radyografi görüntüsü üzerinde yapılan analizde karar ağacı, DVM ve EYKA algoritmaları en yüksek doğruluk oranlarını sağlamıştır; YSA ise genel olarak en istikrarlı sonuçları göstermiştir. İkinci, üçüncü ve dördüncü servikal vertebrada on dokuz referans noktası belirlenmiş ve bu noktalardan yirmi farklı lineer ölçüm yapılmıştır. Elde edilen sonuca göre, karar ağacı algoritması CVS1 (%97.1) ve CVS2 (%90.5), DVM CVS3 (%73.2) ve CVS4 (%58.5), EYKA ise CVS5 (%60.9) ve CVS6 (%78.7) aşamalarını belirlemede en yüksek doğruluk oranlarına ulaşmıştır (Kök vd.,2019).

Gerçekleştirilen başka bir çalışmada da araştırmacılar 8 ile 25 yaş arasındaki bireylerde cinsiyet, kronolojik yaş, diş olgunlaşması ve servikal vertebral olgunlaşma arasındaki ilişkiyi değerlendirmeyi amaçlamışlardır. 500 hastanın dijital panoramik ve lateral sefalometrik radyografileri kullanılarak Baccetti'nin dental olgunlaşma ile ilgili modifiye Demirjian sınıflandırma sistemiyle servikal vertebra olgunlaşması incelenmiş ve Kolmogorov-Smirnov testi dâhil çeşitli istatistiksel analizler uygulanmıştır (Mağat,2024). Bu çalışma sonucunda, cinsiyet demografik parametresine göre dental olgunlaşma düzeyleri değişmezken, servikal vertebra olgunlaşmasında cinsiyete göre farklılık gözlenmiştir. Ayrıca, kronolojik yaş ile dental ve servikal vertebral olgunlaşma arasında güçlü bir pozitif korelasyon tespit edilmiştir.

Diğer bir çalışmada da SVM analizi için farklı makine öğrenimi algoritmalarını kullanarak beş ayrı model geliştirilmiş ve bu modellerin performansları detaylı bir şekilde karşılaştırılmıştır. 647 dijital lateral sefalometrik radyografi kullanılarak yapılan analizde elde edilen sonuçlara göre, SVM aşama sınıflandırmasında en iyi doğruluk YSA modeli ile %92.6 olarak sağlanmıştır (Amasya vd.,2020).

Bu alanda yapılan başka bir çalışmaya göre, servikal vertebralardaki şekil değişiklikleri hastanın büyüme-gelişim aşamasını belirlemek için kullanılmıştır ve genellikle 4 veya 6 aşamaya ayrılarak analiz edilmiştir (Ofraz ve Görgülü,2017). Ayrıca ortodontik tedavi zamanlamasının belirlenmesinde servikal vertebra yönteminin ek radyografiye ihtiyaç duymadan mevcut sefalometrik radyografilerle yapılabileceği ve bu sayede tedavi sürecinin daha etkili yönetilebileceği ifade edilmiştir.

Bir diğer çalışmada 7 ile 21 yaş aralığındaki hastalardan elde edilen 100 lateral sefalogramı kullanarak SVM analizinde özel bir yazılımın manuel analizle

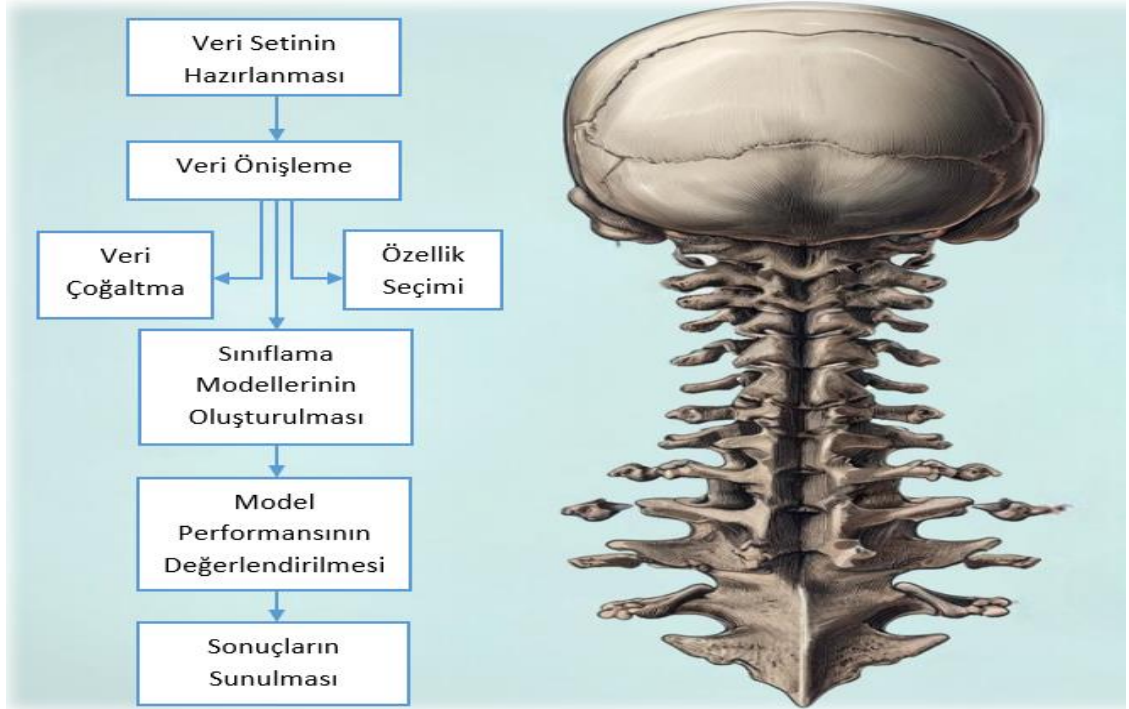
karşılaştırıldığında uygulanabilirliğini değerlendirmişlerdir. Elde edilen sonuca göre yazılımın manuel yönteme göre %94-%93 uyum sağladığını ancak manuel analizlerin daha hızlı olduğunu göstermiştir (Padalino vd.,2014).



BÖLÜM II

MATERYAL VE YÖNTEM

Bu çalışmanın temel amacı, sefalometrik radyografilerde gözlemlenen servikal vertebraların analiz edilmesi ve alana özgü olarak seçilmiş makine öğrenme algoritmaları kullanılarak servikal vertebralardaki olgunlaşma aşamalarının tespit edilmesidir. Bu sayede ortodontik tedavi süreçlerinin ve planlamalarının daha etkin bir şekilde gerçekleştirilmesine katkıda bulunulmak hedeflenmiştir. Bu bağlamda materyal ve yöntem bölümünde, sefalometrik radyografiler kullanılarak gerçekleştirilen büyüme evresi tespitinde kullanılan yöntemlerin sistematik bir analizi yapılmıştır. Makine öğrenmesine genel bir bakış yapılmış olmakta birlikte verilerin işlenmesi, sınıflandırılması ve sonuçların değerlendirilmesi süreçlerinde izlenen adımlar detaylı bir şekilde açıklanmaktadır. Bu süreçte kullanılan adımların bir özeti Şekil 2.1'de sunulan akış diyagramında gösterilmiştir.



Şekil 2.1. Çalışmaya Ait İş Akış Şeması

Kaynak: Yazar tarafından oluşturulmuştur

2.1. Büyüme Gelişiminin Belirlenmesinde Kullanılan Yöntemler

Büyüme-gelişimin izlenmesi özellikle ortodontik tedavi ve iskeletsel gelişim süreçlerinin yönetiminde kritik bir öneme sahiptir (Özer vd.,2006). Bu süreçlerin belirlenmesinde çeşitli yöntemler kullanılmakta olup en yaygın olanlar arasında radyografik teknikler öne çıkmaktadır. Panoramik, sefalometrik ve el-bilek radyografileri gibi yöntemler iskeletsel ve dental yapıların gelişim evrelerini detaylı bir şekilde analiz etme imkânı sunmaktadır. Bu yöntemlerin büyüme-gelişim takibi ve yaş belirleme açısından sağladığı kesinlik ve güvenilirlik onları diğer tekniklerden daha çok tercih edilir kılmaktadır.

2.1.1. Radyografik Yöntemler

Büyüme-gelişimin belirlenmesinde kullanılan radyolojik yöntemler; panoramik, sefalometrik ve el-bilek radyografileri olarak gruplandırılmaktadır (Carty,2002; Greulich ve Pyle,1959; Hassel ve Farman,1995; Lamparski ve Nanda,2002; Tanner vd., 1997; White ve Pharoah,2013). Radyografik yöntemlerle elde edilen görüntüler, dişler ve kemiklerin her dönemdeki değişimlerini kolayca izlemeye olanak tanıdığı için bu yöntem diğer büyüme-gelişim ve yaş belirleme tekniklerine göre daha fazla tercih edilmektedir (Akkoyun,2013).

2.1.1.1. Panoramik Radyografiler

Panoramik radyografi ağız ve çene bölgesinin geniş bir görüntüsünü sağlayarak bu bölgenin ayrıntılı olarak izlenmesine olanak tanıyan bir görüntüleme yöntemi olarak ifade edilmektedir (White ve Pharoah,2013) . Bu yöntem dişler, çene kemikleri, sinüsler ve çevre dokuların bir bütün olarak tek bir film veya dijital görüntü üzerinde izlenmesine olanak tanımaktadır. Demirjian ve ark. 1973 yılında panoramik radyografiler üzerinde gerçekleştirdikleri çalışmayı güncelleyerek dental matürasyonu dişlerin kök formasyonuna göre 8 evreye ayırarak sınıflandırmış ve bu teknik Demirjian metodu şeklinde literatürde yer bulmuştur (Demirjian ve Goldstein,1976). Bu yöntemin geçerliliği farklı toplumlarda test edilmiştir ancak sonuçların ülkeler arasında ve aynı ülkenin farklı bölgelerinde bile değişkenlik göstermesi nedeniyle sınırlı kullanım alanı bulunmaktadır (Jayaraman vd.,2013; Nur vd.,2012). Dental matürasyonun büyüme-gelişim döneminin sadece ilk aşamalarını kapsamaması ve diş sürme zamanının çeşitli

faktörlerden etkilenmesi nedeniyle günümüzde ortodontide yaygın olarak tercih edilmemektedir (Erverdi,1988; Sierra,1987; Ülgen,2010).

2.1.1.2. El-Bilek Radyografileri

X-ışınının keşfedilmesinden dört ay sonra 1896 yılında Londra'da Sydney Rowland tarafından ilk el-bilek radyografisi çekilmiştir (Singer,1980). Yine aynı yıl içerisinde iskelet gelişimini el-bilek radyografileri üzerinden inceleyen ilk kişi ise Ranke olmuştur (Ünver,2013). W.Greulich-Pyle el-bilek radyografilerini inceleyerek bir atlas oluşturmuş ve bireylerin kemik yaşını bu atlasla karşılaştırarak belirlemiştir (Greulich ve Pyle,1959). 1962'de Tanner ve ark. kemik yaşını belirlemek için TW2 yöntemini geliştirmiş bu yöntem daha sonra TW3 olarak güncellenmiştir (Carty,2002; Tanner vd.,1997). Fishman el-bilek radyografilerinde anatomik noktaları kullanarak kemik gelişim süreçlerini değerlendiren bir yöntem geliştirmiştir (Bica ve Mariana,2010; Fishman,1982). Bu tekniğin olumsuz yanları arasında el-bilek bölgesindeki karmaşık referans noktalarının pubertal büyüme yaşını sınırlı bir şekilde belirlemesi, hatalı yorumlara neden olabilmesi ve ek radyasyon dozuna ihtiyaç duyulması yer almaktadır.

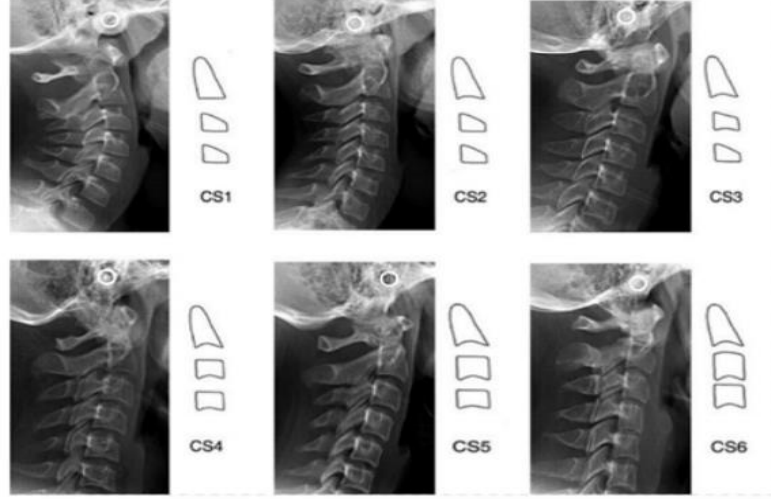
2.1.1.3. Sefalometrik Radyografiler

Sefalometrik radyografiler, baş ve yüz bölgesinin ayrıntılı bir şekilde incelenmesine olanak tanıyan bir görüntüleme yöntemidir. Bu teknik dişlerin, çene kemiklerinin ve çevre yapılarının iki boyutlu bir projeksiyonunu sunarak ortodontik değerlendirme, büyüme-gelişim analizi ve tedavi planlaması için kritik veriler sağlamaktadır. Literatür incelendiğinde Ruel ve Bench, servikal vertebraların büyüme-gelişimini ilk kez inceleyerek bu büyüme-gelişimin sefalometrik radyografilerle ölçülebileceğini belirtmişlerdir (Bench,1963). Bu araştırmacılar belirli bir vertebra ile büyüme-gelişim tespiti yapmamıştır (Bench,1963). Lamparski el-bilek radyografilerine alternatif olarak sefalometrik radyografilerdeki servikal vertebralardan büyüme-gelişim ve kemik yaşı tespit edilebileceğini ortaya koymuştur (Lamparski,1972). Lamparski, servikal vertebraların büyüme-gelişim dönemlerini belirlemek için altı aşamalı bir standart seri geliştirmiş ve bu yöntemin el-bilek radyografileri kadar kararlı ve tutarlı olduğunu belirtmiştir. Bu bulgular, sefalometrik radyografilerin sadece diş ve çene

yapılarını değil aynı zamanda servikal vertebra gelişimini de kapsamlı bir şekilde değerlendirmek için etkili bir araç olduğunu ortaya koymaktadır.

2.1.1.4. Servikal Vertebra

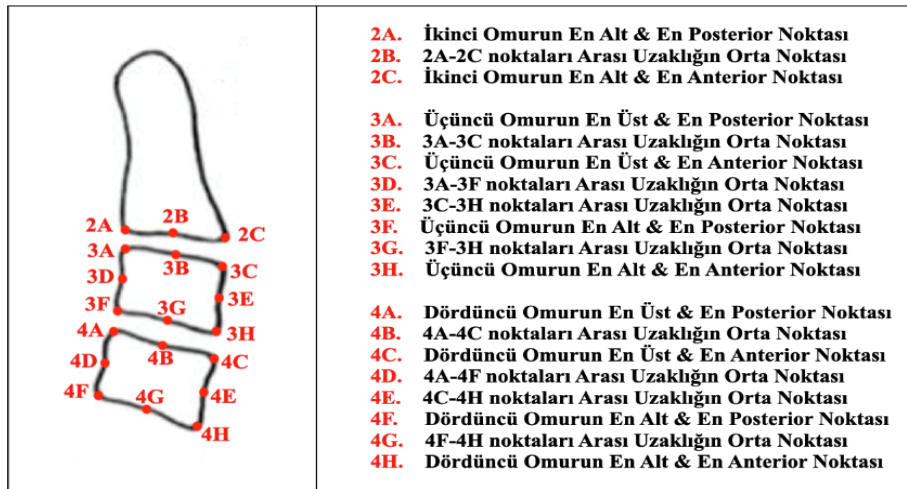
Servikal vertebra boyun bölgesindeki omurlar olup büyüme-gelişim süreçlerinin izlenmesinde önemli rol oynamaktadır. Özellikle sefalometrik radyografilerde C2'den C6'ya kadar olan servikal vertebra gelişim evreleri analiz edilerek bireylerin büyüme-gelişim döneminin ve iskelet yaşının tahmin edilmesi mümkündür (Chatziagianni ve Halazonetis,2009; Franchi vd.,2000; Garcia-Fernandez vd.,1998; Hellsing,1991; Kucukkeles vd.,1999; Mito vd.,2002; O'Reilly ve Yanniello,1988; Román vd.,2002). Bu vertebra büyüme atılım dönemlerini belirlemede ve ortodontik tedavi planlamasında kritik veriler sağlamaktadır. Genellikle tahmin işlemi için 4 veya 5 servikal omurun değerlendirilmesi tercih edilirken bazı araştırmacılar bu omurları 6 farklı aşamaya ayırarak olgunlaşma sürecini daha detaylı bir şekilde incelemişlerdir. Bu yaklaşım servikal vertebra büyüme-gelişim evrelerini daha ayrıntılı analiz ederek bireylerin büyüme-gelişim süreçlerini daha hassas bir şekilde değerlendirmeyi amaçlamaktadır. Literatür taramalarına göre birçok araştırmacı SVM değişikliklerini temel alarak farklı yöntemler geliştirmiş ve bu yöntemlerin geçerliliği üzerine çalışmışlardır. Çalışmalar sonucunda bu araştırmacılar servikal vertebra olgunlaşmasını değerlendirmek için Şekil 2.2' de gösterilen altı evreden oluşan bir yaklaşımı tanıtmışlardır (Baccetti vd.,2005). Bu yaklaşıma göre ergenlik büyümesinin büyük ölçüde dikkate alınmadığı dönem olgunluk evresi CS5'tir. Bu dönemde C2, C3 ve C4'ün alt kenarlarında bulunan oyuklar daha belirgin hale gelirken C3 ve C4 kare şekline dönüşür. Sefalometrik radyografiler aracılığıyla servikal vertebra büyüme-gelişim süreçlerinin değerlendirilebileceğini ilk kez Ruel ve Bench ortaya koymuş olsa da kemik yaşı ve büyüme-gelişimin tespitinin yapılabileceği ilk olarak Lamparski tarafından rapor edilmiştir (Bench,1963; Lamparski ve Nanda,2002). Lamparski'nin geliştirdiği altı aşamalı bir standart seri servikal vertebra gelişim aşamalarını daha iyi anlamak ve doğru bir şekilde değerlendirmek için önemli bir referans kaynağı olmuştur. Baccetti ve ark. birinci ve ikinci evrelerin belirlenmesindeki zorluk nedeniyle servikal vertebra gelişim dönemlerini önce 5 evreye daha sonra bu çalışmayı modifiye ederek 6 evreye sınıflamıştır (Baccetti vd.,2005).



Şekil 2.2. Baccetti ve Ark. Tarafından Yenilenmiş Olan SVM Aşamaları ve Radyografik Görünümü

Kaynak: (Baccetti vd.,2005)

Bu tez çalışmasında, sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılmasıyla bazı seçilmiş makine öğrenme algoritmaları kullanılarak servikal vertebralardaki olgunlaşma aşamalarının tespitinin yapılması hedeflenmiştir. Bu sayede ortodontik tedavi süreçlerinin ve planlamalarının daha etkin bir şekilde gerçekleştirilmesine katkıda bulunulması amaçlanmıştır. Bu hedef için kullanılmış radyografik verilerde referans alınan omurlara ait sefalometrik noktalar Şekil 2.3' de belirtilmiştir.



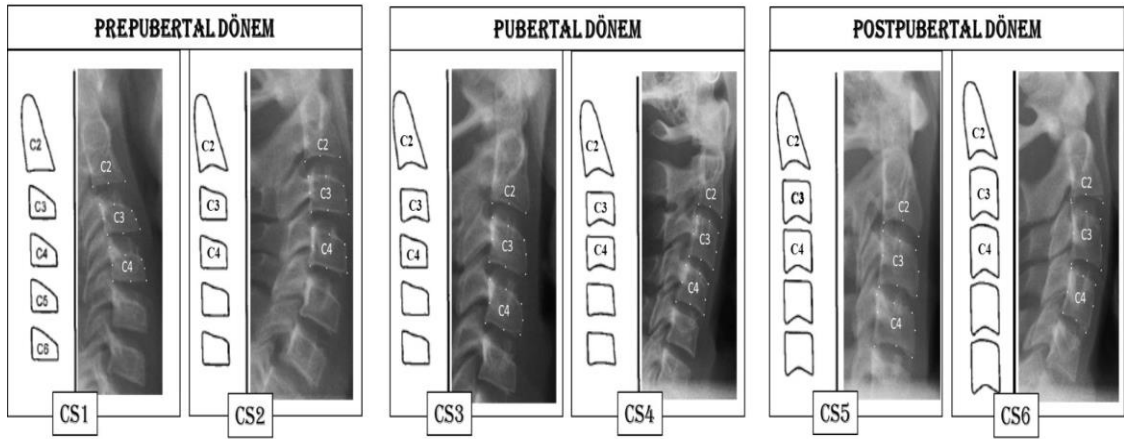
Şekil 2.3. SVM Evresi Tepitinde Kullanılan Sefalometrik Referans Noktaları

Kaynak: Yazar tarafından oluşturulmuştur

2.2. Veri seti

Çalışma kapsamında kullanılan veri kümesi Selçuk Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalı'na başvuran hastalardan sefalometrik radyografileri çekilerek ve 2023/22 sayılı etik kurul onayı alınarak oluşturulmuştur. 120 tane özgün veriden ve 6 sınıftan oluşan bu veri setindeki her bir örnek bir sefalometrik radyografi görüntüsünü temsil etmektedir. Sefalometrik radyografiler gelişim seviyesine bağlı olarak altı farklı kategoride gruplandırılmıştır. Sefalometrik radyografi görüntülerindeki noktaların her biri X ve Y koordinat değerlerini içermektedir. Çalışma kapsamında kullanılan veri setini oluşturmak için literatür çalışmalarından da faydalanarak büyüme ve gelişme tespiti için önemli 18 özellik uzman bir diş hekim tarafından bireylerin sefalometrik radyografileri üzerinden önerilmiştir.

Tez çalışmamız için kullanılan veri seti, servikal vertebral olgunlaşma dönemini temsil eden altı ayrı gruptan oluşmaktaydı. Bu çalışma kapsamında, büyüme-gelişme evrelerinin daha yüksek başarıyla tanımlanabilmesi amacıyla mevcut veri seti, üç farklı kategoride sınıflandırılan bir sınıflandırma problemi olarak ele alınmıştır. SVM evresinin belirlenmesi için uzman diş hekimi Doç. Dr. Hatice KÖK tarafından bireylerin sefalometrik radyografileri üzerinden önerilen 18 özellik, titizlikle otomatik olarak hesaplanıp kaydedilmiştir. Şekil 2.3' de, belirtilen 18 özellik elde edilirken her bir evrede faydalanılan noktalar gösterilmektedir. Çalışmada kullanılan veri setinde Şekil 2.4' te de belirtilen CS1-CS2 Prepubertal Dönem, CS3-CS4 Pubertal Dönem ve CS5-CS6 ise Postpubertal Dönem olarak etiketlenmiştir.



Şekil 2.4. Büyüme ve Gelişme Tespitinde Dönemler

Kaynak: (Franchi et al., 2000)

Çalışma kapsamında oluşturulan veri setine ait özellikler Şekil 2.3 ve Şekil 2.4 temelinde planlanmıştır. Şekil 2.4 incelenirse, evre değişimlerinde belirtilen noktalar arasındaki açı, uzaklık, eğim ve yükseklik değerlerinde değişimlerin olduğu net bir şekilde görülmektedir. Bu değişimler baz alınarak oluşturulan özelliklere ait ayrıntılar Tablo 2.1’ de sunulmaktadır. Uzman diş hekimi Doç. Dr. Hatice KÖK tarafından önemli olduğu belirtilen bu noktalar genel olarak tüm diş hekimlerinin manuel büyüme gelişimi tahmininde kullanmış oldukları yönergeleri temel almaktadır. Tablo 2.1’ de görüldüğü üzere, uzman diş hekimi tarafından belirlenen bu önemli noktalar üzerinde gerekli açı, yükseklik ve Öklid uzaklığı hesaplamaları otomatik yapıldıktan sonra tez çalışması kapsamında kullanılacak olan 18 özellikli veri seti oluşturulmuştur.

Tablo 2.1. Tez Çalışması İçin Oluşturulan Veri Setine Ait Özellikler

Özellik No	Özellik Adı	İlgili Koordinat Değerleri
1	2A, 2B ve 2C noktaları arasındaki açı	2A-x,2A-y, 2B-x,2B-y, 2C-x,2C-y
2	2B noktasının, 2A ve 2C arasındaki doğruya olan mesafesi	2A-x,2A-y, 2B-x,2B-y, 2C-x,2C-y
3	3G, 3F ve 3E noktaları arasındaki açı	3E-x,3E-y, 3F-x,3F-y, 3G-x,3G-y
4	3F noktasının, 3G ve 3E arasındaki doğruya olan mesafesi	3E-x,3E-y, 3F-x,3F-y, 3G-x,3G-y
5	3G, 3A ve 3C noktaları arasındaki açı	3A-x,3A-y, 3C-x,3C-y, 3G-x,3G-y
6	3A, 3C ve 3E noktaları arasındaki açı	3A-x,3A-y, 3C-x,3C-y, 3E-x,3E-y
7	3A ile 3G arası uzaklık ve 3C ile 3E arası uzaklık oranı	3A-x,3A-y, 3C-x,3C-y, 3E-x,3E-y, 3G-x,3G-y
8	3A ile 3G arası uzaklık ve 3G ile 3E arası uzaklık oranı	3A-x,3A-y, 3E-x,3E-y, 3G-x,3G-y
9	3C ile 3F arası uzaklık ve 3G ile 3E arası uzaklık oranı	3C-x,3C-y, 3E-x,3E-y, 3F-x,3F-y, 3G-x,3G-y
10	3A ile 3C arası uzaklık ve 3G ile 3E arası uzaklık oranı	3A-x,3A-y, 3C-x,3C-y, 3E-x,3E-y, 3G-x,3G-y
11	4G, 4F ve 4E noktaları arasındaki açı	4E-x,4E-y, 4F-x,4F-y, 4G-x,4G-y
12	4F noktasının, 4G ve 4E arasındaki doğruya olan mesafesi	4E-x,4E-y, 4F-x,4F-y, 4G-x,4G-y
13	4G, 4A ve 4C noktaları arasındaki açı	4A-x,4A-y, 4C-x,4C-y, 4G-x,4G-y
14	4A, 4C ve 4E noktaları arasındaki açı	4A-x,4A-y, 4C-x,4C-y, 4E-x,4E-y
15	4A ile 4G arası uzaklık ve 4C ile 4E arası uzaklık oranı	4A-x,4A-y, 4C-x,4C-y, 4E-x,4E-y, 4G-x,4G-y
16	4A ile 4G arası uzaklık ve 4G ile 4E arası uzaklık oranı	4A-x,4A-y, 4E-x,4E-y, 4G-x,4G-y
17	4C ile 4F arası uzaklık ve 4G ile 4E arası uzaklık oranı	4C-x,4C-y, 4E-x,4E-y, 4F-x,4F-y, 4G-x,4G-y
18	4A ile 4C arası uzaklık ve 4G ile 4E arası uzaklık oranı	4A-x,4A-y, 4C-x,4C-y, 4E-x,4E-y, 4G-x,4G-y

Kaynak: Yazar tarafından oluşturulmuştur

2.3. Yöntem

Bu tez çalışmasında sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılmasıyla birlikte makine öğrenmesi algoritmaları kullanılarak SVM evresinin tespiti hedeflenmiştir.

2.3.1. Ön İşlemler

Sefalometrik radyografiler üzerinde farklı ön işlem yöntemleri kullanılmıştır. Ön işlemler aşamasında veri setinin analize hazır hale getirilmesi için çeşitli adımlar izlenmiştir. Bu işlemlerin yapılmasındaki ana amaç, makine öğrenmesi modelinin performansını optimize ederek daha doğru ve güvenilir sonuçlar elde etmektir.

2.3.2. Veri Çoğaltma

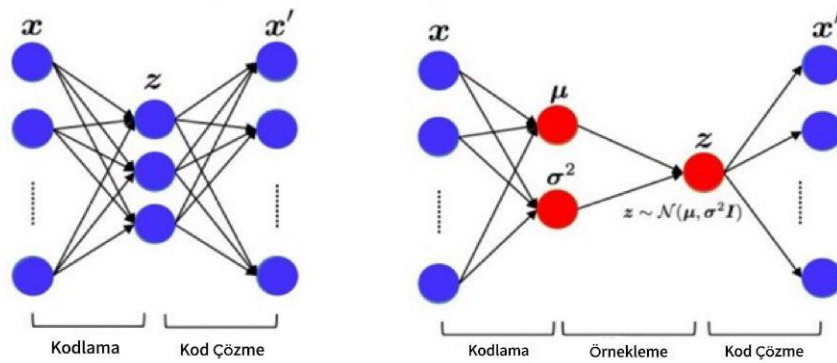
İlk olarak veri setinin çeşitliliğini arttırmak ve modelin genelleme kapasitesini iyileştirmek amacıyla veri çoğaltma tekniği uygulanmıştır. Bu sayede sınırlı sayıda örnekle sınırlı kalmadan modelin farklı senaryoları öğrenmesi sağlanmıştır. Uygulanan veri çoğaltma yöntemi modelin eğitim sürecinde çeşitli varyasyonları tanımaya olanak tanıyarak aynı zamanda da aşırı uyum göstermeden daha geniş bir veri kümesinde tutarlı sonuçlar üretmesini sağlamıştır.

2.3.2.1. Varyasyonel Otomatik Kodlayıcılar

Makine öğrenimi ve yapay zekâ alanındaki ilerlemeler veri üretimi ve temsili üzerinde daha fazla etki yaratmaya başlamıştır. Bu gelişmelerden biri olan Varyasyonel Otomatik Kodlayıcılar (VAE) aslında standart otomatik kodlayıcının bir çeşidi olarak öne çıkar. VAE, 2013 yılında Google ve Qualcomm'da Diederik P. Kingma ve Max Welling tarafından geliştirilip önerilmiştir (Kingma,2013). VAE, veri temsillerini öğrenmek ve yenilikçi bir şekilde yeni veri örneklerini oluşturmak amacıyla kullanılan önemli bir modeldir (Kingma,2013). VAE özellikle yeni ve orijinal veri üretimi, veri artırma, boyut azaltma ve üretici modelleme gibi alanlarda kullanılır. Bu kısımda VAE'yi daha iyi anlayabilmek adına öncelikle standart otomatik kodlayıcının nasıl çalıştığına değinildi ve ardından varyasyonel otomatik kodlayıcının özellikleri açıklanmıştır.

Standart otomatik kodlayıcılar, verilen girdi verisini yeniden üretmeye çalışan bir YSA türüdür. Bu süreçte kullanılan girdi verileri etiketsiz olduğundan standart otomatik kodlayıcılar gözetimsiz öğrenme modeli olarak sınıflandırılır (Bengio vd.,2013; Lanham,2021). Girdi olarak aldığı veriyi ilk önce bir kod haline dönüştürüp ardından bu koddan yeniden girdi verisini elde etmeyi hedefler. Bu noktada yapısının aslında iki ana bölümden oluştuğu belirgin bir şekilde ortaya çıkmaktadır. Girdi verisini alarak bunu bir kod haline dönüştüren kısım kodlayıcı (encoder) bölümünü temsil ederken; bu dönüştürülmüş kod bilgisini kullanarak girdi verisini yeniden elde etmeye çalışan kısım ise kod çözücü (decoder) bölümünü oluşturur. Bu iki kısım standart otomatik kodlayıcının veriyi sıkıştırarak minimum kayıpla yeniden oluşturmasını sağlamak üzere yapılandırılmıştır.

En temel haliyle standart otomatik kodlayıcıda, her bir giriş verisi kodlayıcı katmanından geçip işlenerek belirsiz bir vektöre (latent vector) dönüştürülür ve bu vektör kod olarak adlandırılır. Oluşturulan bu kod, daha sonra kod çözücü katmanına iletilir ve girdi verisinin yeniden oluşturulması sağlanır (Hinton ve Salakhutdinov,2006; Lanham,2021). VAE’de ise kodlayıcıdan geçen girdi doğrudan bir kod olarak değil, bir olasılık dağılımı şeklinde kodlanıp kullanılır. Yani kodlayıcının çıktısı belirli bir koddan ziyade ortalama ve varyans değerleri olarak verilir. Bu ortalama ve varyans değerleri üzerinden örnekleme yapılarak kod üretilir ve bu koda kod çözücü tarafından çözülebilir (Kingma,2013; Lanham,2021; Rezende vd.,2014). Bu yönüyle VAE, standart otomatik kodlayıcılardan ayrılmaktadır. Şekil 2.5’te VAE ile standart otomatik kodlayıcıların genel yapılarını gösteren bir şema bulunmaktadır.



Şekil 2.5. VAE ile Standart Otomatik Kodlayıcılar Arasındaki Fark

Kaynak: (Lanham,2021)

Şekil 2.5.'de VAE ile standart otomatik kodlayıcı arasındaki temel farklılıklar ortaya konulmuştur. VAE'ler standart otomatik kodlayıcılardan olasılıksal yaklaşımı, gizli uzayın temsili ve veri üretimi konularındaki farklılıklarıyla ayrılır (Kingma,2013; Rezende vd.,2014). Şekil 2.5'e göre standart otomatik kodlayıcıda girdi (x), bir koda (z) dönüştürülerek sıkıştırılır ve daha sonra tekrar eski formuna (x') geri getirilir. VAE'de ise girdi (x) üzerinden ortalama (μ) ve varyans (σ^2) değerleri elde edilir. Bu değerler kullanılarak kod (z) oluşturulmak üzere örnekleme (sampling) işlemi gerçekleştirilir. Ardından elde edilen bu kod yeniden çözülerek orijinal verilere (x') ulaşılır (Kingma,2013).

2.3.2.2. Kayıp Fonksiyonu

Standart otomatik kodlayıcıların eğitimi sırasında kullanılan kayıp fonksiyonu yeniden oluşturma kaybı (reconstruction loss) dır. Bu kayıp fonksiyonu modele sağlanan girdi ile modelin ürettiği çıktı arasındaki farkı ifade eder. Bu farkı hesaplarken genellikle ortalama kare hatası yöntemi tercih edilir. Buna karşın VAE'de bu kayba ilave olarak düzenleme kaybı (regularization loss) da kullanılır. Bu kayıp fonksiyonuna başka bir isim olarak belirsizlik kaybı (latent loss) da denilmektedir. VAE'de toplam kayıp kodlayıcının çıktısında elde edilen ortalama ve varyans değerleriyle oluşan dağılımın normal dağılımdan sapma miktarını ifade eder. Bu kaybı hesaplamak için iki olasılık dağılımı arasındaki farklılığı ölçen Kullback Leibler (KL) uzaklık metriği kullanılır. Sonuç olarak VAE'de toplam kayıp denklem 1'de gösterildiği gibi yeniden oluşturma kaybı ile düzenleme kaybının toplamı şeklinde hesaplanarak elde edilir (Kingma,2013; Lanham,2021). Ve bu özellik VAE'yi standart otomatik kodlayıcılardan ayıran başka temel farklardan birisidir.

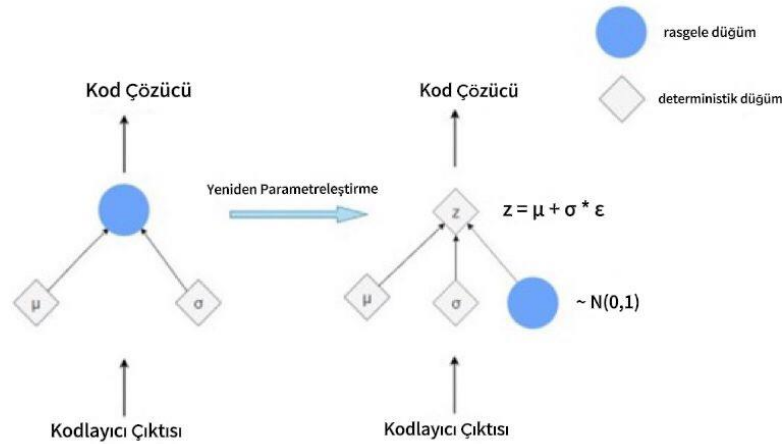
$$\text{Kayıp fonksiyonu} = \|x - x'\|_2 + KL[N(\mu_x, \sigma_x), N(0,1)] \quad (1)$$

Bu Denklem 1'deki kayıp fonksiyonunda x modelin aldığı giriş çerçevelerini x' ise modelin ürettiği çerçeveleri temsil etmektedir. Kayıp fonksiyonun ilk terimi olan $\|x - x'\|_2$, yeniden oluşturma kaybını ifade etmektedir. Burada x ve x' arası öklid uzaklığı ile hesaplanmaktadır. Kayıp fonksiyonun ikinci terimi olan $KL[N(\mu_x, \sigma_x), N(0,1)]$ ifadesi ise düzenleme kaybını ifade etmektedir. Bu terimde modelin kodlama sürecinde

hesapladığı μ ve σ değerlerine dayanan normal dağılım elde edilir. Ardından bu dağılım ile standart normal dağılım arasındaki Kullback-Leibler uzaklığı hesaplanır.

2.3.2.3. Kodlama Sürecinde Yeniden Parametrelendirme Hilesi

YSA modelleri geri yayılım (backpropagation) yöntemi kullanılarak eğitilir ancak VAE’de kodlama işlemi sırasında ortalama ve varyans değerlerinden kod örnekleme işlemi bu klasik geri yayılım yöntemine uygun değildir. Bu durum modelin eğitim sürecinde doğrudan ortalama ve varyans değerleriyle kod üretmenin geri yayılım için engel oluşturmaya yol açar. VAE’yi geri yayılım yöntemi kullanarak eğitebilmek için bir ϵ değeri kullanılır. Bu ϵ değeri, normal dağılımdan rastgele elde edilen bir sayı olarak tanımlanır. Kod (z) üretimi doğrudan μ ve σ değerlerinden yapılmak yerine, “ $z = \mu + \sigma * \epsilon$ ” formülü ile gerçekleştirilir. Geri yayılım işlemi sırasında da ϵ değeri sabit tutulur ve modelin eğitimi bu şekilde gerçekleştirilir. Bu yönteme yeniden parametrelendirme hilesi adı verilir. Bu işlemin ana ve tek amacı sistemin geri yayılım yöntemi ile eğitilebilmesini mümkün kılmaktır (Kingma,2013; Lanham,2021).



Şekil 2.6. Yeniden Parametrelendirme Hilesi Uygulamasının Yapıyı Nasıl Dönüştürdüğünün Gösterimi

Kaynak: (WikiDocs,2023)

Şekil 2.6’da yeniden parametrelendirme işleminin yapıya olan etkisi ve dönüşüm süreci sunulmaktadır.

```

Girdi: Gözlemlenen veri kümesi  $D = \{x^{(1)}, \dots, x^{(N)}\}$ .
Yığın boyutu ( $n$ ), Öğrenme oranı ( $\alpha$ ), Örnekleme boyutu ( $M$ ) ve Parametrelerin ( $\theta$  ve  $\phi$ ) Rastgele Başlatılması

tekrar et
  D'den rastgele bir yığın seç  $X = \{x^{(1)}, \dots, x^{(n)}\}$ .
  her  $x^{(i)} \in X$  için yap
     $N(0, I)$  dağılımından  $\{\varepsilon^{(1)}, \dots, \varepsilon^{(M)}\}$  gürültüsünü üret.
     $\mu, \sigma \leftarrow \text{Encoder}(x^{(i)}; \phi)$ .

    1  $\leftarrow$  1'den  $M$ 'ye kadar yap
       $z^{(l)} \leftarrow \mu + \sigma \odot \varepsilon^{(l)}$ .
    bitir

     $L^{(i)}_{\text{yeniden_olusturma}} \leftarrow (1/2) \sum_{m=1}^M \|x^{(i)} - \text{Decoder}(z^{(l)}; \theta)\|^2$ .
     $L^{(i)}_{\text{KL}} \leftarrow -(1/2) \sum_{j=1}^{\text{dim}} (1 + \log \sigma_j^2 - \mu_j^2 - \sigma_j^2)$ .
  bitir

 $\theta \leftarrow \theta - \alpha \nabla_{\theta} \sum_{n=1}^n (L^{(i)}_{\text{yeniden_olusturma}})$ .
 $\phi \leftarrow \phi - \alpha \nabla_{\phi} \sum_{n=1}^n (L^{(i)}_{\text{yeniden_olusturma}} + L^{(i)}_{\text{KL}})$ .

yakınsama sağlanana kadar devam et.

```

Şekil 2.7. VAE Modeline Ait Sözde Kod

Kaynak: (Ding,2022)

Şekil 2.7’de yer alan sözde kod VAE modelinin temel çalışma prensiplerini adım adım açıklayan bir yapıdır. İlk olarak modelin öğrenmesi için kullanılan veri kümesi bir yığın halinde alınıyor. Algoritmanın başlangıcında parametreler rastgele başlatılıp modelin performansını doğrudan etkileyen yığın boyutu (n) ve öğrenme oranı (α) gibi hiperparametreler optimize edilir. Bu hiperparametrelerin doğru seçimi modelin eğitim sürecinin verimliliği ve doğruluğu açısından kritik öneme sahiptir. Yığın boyutu her iterasyonda işlenecek veri miktarını belirler. Yığın boyutunun küçük tutulması modelin eğitim sürecinde daha sık güncellenmesine ancak daha fazla iterasyon gereksinimine neden olur. Öğrenme oranı model parametrelerinin her iterasyonda ne kadar büyük adımlarla güncelleneceğini belirleyen hiperparametredir. Küçük bir öğrenme oranı modelin yavaş ama istikrarlı bir şekilde öğrenmesini sağlarken, büyük bir öğrenme oranı ise modelin hızlı öğrenmesini ancak potansiyel olarak dengesiz sonuçlar elde etmesini sağlayabilir. Dolayısıyla öğrenme oranı modelin eğitim sürecinin kararlılığı ve doğruluğu açısından oldukça önemli bir faktördür. Burada yer alan örnekleme boyutu

(M) ise örnekleme işlemi için kullanılan bir hiperparametredir. Algoritmanın başlangıç aşamasından sonra her veri örneği için kodlayıcı kısmı veriyi alarak ortalama (μ) ve varyans (σ) değerlerini üretir. Bu değerler veri örneğinin latent (gizli) uzayındaki bir temsili olan z vektörünü elde etmek için kullanılır. Latent uzayda bu temsiller standart normal dağılımdan çekilen rastgele gürültü (ϵ) ile oluşturulur. Daha sonra z değerleri kod çözücüyeye verilerek orijinal verinin yeniden oluşturulması sağlanır. Yeniden oluşturulan verinin orijinal veriye ne kadar yakın olduğunu ölçen yeniden oluşturma kaybı ve latent uzaydaki temsillerin normal dağılıma ne kadar yakın olduğunu ölçen Kullback-Leibler uzaklığı hesaplanır. Bu iki kayıp fonksiyonu birlikte optimize edilerek model hem girdiyi doğru bir şekilde yeniden oluşturmayı hem de latent uzayda düzenli ve anlamlı bir dağılım oluşturmayı öğrenmiş olur. Son olarak, model parametreleri belirli bir öğrenme oranına göre güncellenmeye devam eder ve bu süreç parametrelerin belirli bir kararlılığa ulaşarak kayıp fonksiyonunun minimuma yaklaşmasıyla sonlanır (Ding,2022; Kingma,2013; Lanham,2021). Bu modelin eğitim sürecinin tamamlandığını ve daha fazla optimizasyona ihtiyaç kalmadığını gösterir.

2.3.3. Veri Standardizasyonu

Veri çoğaltma aşamasının ardından verilerin birbirleriyle karşılaştırılabilir hale getirilmesi ve model eğitim sürecinde sapmaların önlenmesi amacıyla sayısal veri standardizasyonu gerçekleştirilmiştir. Böylece modelin öğrenme süreci boyunca her bir özelliğin eşit ağırlıkta değerlendirilmesi sağlanmıştır. Standardizasyon işlemi her bir radyografi noktasının X ve Y koordinat değerlerinin belirlenen bir referans aralığında normalize edilmesini içermektedir. Bu adım modelin daha tutarlı ve stabil sonuçlar üretmesi için kritik bir rol oynamıştır.

2.3.4. Özellik Seçimi

Özellik seçimi makine öğrenmesi modellerinin performansını arttırmak ve aşırı uyum riskini azaltmak için kritik bir adımdır. Gereksiz veya düşük etkili özellikleri elimine ederek modelin daha hızlı, verimli ve doğruluğu yüksek bir şekilde çalışmasını sağlamaktadır. Bu süreç yalnızca en anlamlı ve katkı sağlayan verilerin modele dahil edilmesiyle modelin genelleme kapasitesini iyileştirir ve daha güvenilir sonuçlar üretmesine olanak tanımaktadır. Özellik seçimi aşamasında Lasso yöntemi kullanılarak büyüme-gelişim tespitinde önemli olan 7 etkin özellik belirlenmiştir. Bu özelliklerin

seçimiyle birlikte modelin karmaşıklığının azaltılması ve daha yüksek doğruluk oranlarına ulaşılması hedeflenmiştir.

2.3.4.1. Lasso Özellik Seçme Yöntemi

Makine öğrenmesi alanında regularizasyon terimi aşırı uyum olarak bilinen yaygın sorunu önlemek için kullanılan çeşitli teknikleri ifade eder. Aşırı uyum, bir modelin eğitim verilerine fazla uyum sağlaması sonucu hem verinin temel kalıplarını hem de veriye özgü gürültü ve anomalileri öğrenmesidir (Emeç ve Özcanhan,2023). Bu durum modelin eğitim verilerinde yüksek performans göstermesine rağmen yeni verilere genelleme yapma yeteneğini azaltmaktadır. Regularizasyon bu sorunun önüne geçerek modelin daha genel ve çeşitli veri kümelerinde tutarlı performans sergilemesini sağlar. Regularizasyon eğitim süreci sırasında modele kısıtlamalar getirerek aşırı karmaşıklığı önler ve bu sayede aşırı uyum riskini azaltmaktadır. L1 regularizasyonu tekniğini kullanan Lasso yöntemi, makine öğrenmesi ve istatistikte kullanılan popüler bir özellik seçme yöntemidir (Fonti ve Belitser,2017; Hastie,2009; Ranstam ve Cook,2018).Bu yöntemin asıl amacı modelin karmaşıklığını azaltarak daha yüksek performans elde etmeyi amaçlamaktadır. Lasso, lineer regresyon modellerine bir ceza terimi ekleyerek çalışmaktadır (Fonti ve Belitser,2017; Hastie,2009; Ranstam ve Cook,2018). Bu ceza terimi, modeldeki tüm özelliklerin katsayılarının toplam mutlak değerine dayanmaktadır. Lasso, bu ceza terimini minimize etmeye çalışırken bazı özelliklerin katsayılarını sıfıra indirmekte ve bu durum da o özelliklerin modelden tamamen çıkarılmasına neden olmaktadır. Bu sayede sadece daha önemli ve modelin doğruluğuna daha fazla katkı sağlayan özellikler seçilmektedir. Bu yöntem matematiksel olarak Denklem 2' de şu şekilde gösterilmektedir:

$$\text{Loss}_{L1} = \text{Loss}_{\text{ori}} + \lambda \sum |w_i| \quad (2)$$

Burada, Loss_{ori} modelin tahmin ettiği değerler ile gerçek değerler arasındaki farkı ölçen orijinal kayıp fonksiyonunu temsil etmekte ve modelin doğruluğunu belirlemektedir. λ regularizasyonun gücünü belirleyen bir hiperparametredir ve katsayıların küçültülmesine yönelik cezanın şiddetini kontrol etmektedir; büyük λ değerleri daha fazla regularizasyon uygulayarak katsayıları küçültmeye zorlamaktadır. $\sum |w_i|$ terimi ise modelin katsayılarının mutlak değerlerinin toplamını ifade etmekte ve modelin

kompleksliğini azaltmaya yönelik bir ceza mekanizması sunmaktadır. Bu terim modelin ağırlıklarını küçültmeye teşvik ederek gereksiz veya düşük etkili özellikleri elemeye yardımcı olmakta ve modelin daha basit bir şekilde yorumlanabilir olmasını sağlamaktadır.

2.4. Makine Öğrenmesi

Makine öğrenmesi, bilgisayarların verilerden öğrenmesini ve bu verilerden elde edilen bilgiyle tahminlerde bulunmanın yanı sıra sınıflandırma, desen tanıma, karar alma ve özellik çıkartma gibi çeşitli veri analizi görevlerini de yerine getirmesini sağlayan bir yapay zekâ alanıdır (Mahesh,2020; Zhou,2021). Makine öğrenmesi algoritmaları temelde veri setleri üzerinde analiz yaparak verilerdeki desenleri ve ilişkileri öğrenir. Bu öğrenme süreci algoritmaların belirli görevleri veya problemleri çözmek için model oluşturmasını sağlar. Model eğitildikten algoritma yeni ve görülmemiş veriler üzerinde tahminler yapabilir veya kararlar alabilir. Makine öğrenmesi sürekli olarak verilerle beslenerek ve öğrenme sürecini tekrarlayarak performansını iyileştirir. Makine öğrenmesi, çeşitli öğrenme türlerini kapsar: denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme (Mahesh,2020; Zhou,2021). Denetimli öğrenme etiketlenmiş veriler kullanarak modelin eğitilmesini sağlar. Denetimsiz öğrenme ise etiketlenmemiş verilerden desenleri ve yapıları keşfetmeye odaklanır. Pekiştirmeli öğrenme de bir model çevresiyle etkileşimde bulunarak ve aldığı ödüller veya cezalar doğrultusunda öğrenmesini gerçekleştirir.

2.5. Kullanılan Algoritmalar

Yapay zekânın kullanım amacı sistemlerin etkinliğini arttırmak için daha doğru ve tarafsız kararlar alınmasına yardımcı olmaktır. Bu sebeple bu çalışmada sınıflandırma alanında yaygın olarak kullanılan yedi yapay zekâ algoritması seçilerek ve karşılaştırılarak analiz edildi. Bu algoritmalar arasında YSA, Destek Vektör Makineleri (DVM), K-En Yakın Komşu (EYKA), Torbalama (Bagging), Rastgele Orman (RO), Lojistik Regresyon (LR), Ekstra Gradyan Artırma (XGBoost) algoritmaları bulunmaktadır.

YSA algoritması insan beyninin çalışma prensibini taklit eden aynı zamanda da karmaşık veri yapılarını işlemek ve öğrenmek için tasarlanmış bir makine öğrenimi algoritmasıdır (Öztemel,2003). Bu algoritma sinir hücrelerinden oluşan katmanlar arası

bağlantılarla veriyi işleyerek çeşitli görevleri başarıyla gerçekleştirebilmektedir. Ayrıca, her bir öğrenici modelinin farklı bir özellik alt kümesi üzerinde eğitilerek modeller arasında çeşitlilik oluşturmakta ve topluluk öğrenme modelinin daha güçlü ve genelleştirilebilir olmasını sağlamaktadır.

DVM algoritması genellikle büyük veri kümelerinin olduğu sınıflandırma problemlerinde tercih edilen bir ayırım yöntemidir (Zhou,2021). Veri noktalarını bir hiper düzlemde optimal bir şekilde ayırarak sınıflandırma yapar veya regresyon analizi gerçekleştirmektedir.

EYKA algoritması mesafeye dayalı denetimli bir öğrenme yöntemi olup hem sınıflandırma hem de regresyon problemlerinde kullanılan bir tekniktir. Bu algoritmanın temel prensibi veriler arasındaki mesafeleri ve benzerlikleri kullanarak sınıflandırma yapmaktır. Çalışma mantığı, yeni bir veriyi sınıflandırırken veya tahmin ederken bu veriye en yakın k tane komşunun etiketlerine veya değerlerine bakılmaktadır. Sonrasında veri kendisine en yakın bu k komşunun hangi sınıfına daha çok benziyorsa o komşunun sınıfına etiketlenmektedir (Zhou,2021).

Torbalama algoritması makine öğrenmesinde sınıflandırma ve regresyon problemlerini çözmek için yaygın olarak kullanılan etkili bir topluluk öğrenme yöntemidir. Torbalama çalışma mantığı birden fazla örneklemin rastgele seçilerek alt kümeler oluşturulması, bu alt kümeler üzerinde birden fazla modelin eğitilmesi ve elde edilen tahminlerin bir araya getirilmesi yoluyla daha iyi sonuçlar elde edilmesi amacı üzerine kurgulanmıştır (Zhou,2021). Böylece farklı alt kümeler üzerinde eğitilen modellerin birleştirilmesi tek bir modelin kullanılmasından daha iyi genelleme sağlayacaktır. Birden fazla modelin bir araya getirilmesi, her bir modelin tahminlerindeki varyansı azaltmakta ve ayrıca karmaşıklığı da azaltarak aşırı uyum riskini düşürmektedir. Torbalama algoritması özellikle karmaşık ve değişken veri kümeleri üzerinde iyi performans gösteren etkili bir kolektif yöntemdir.

RO algoritmasında birçok karar ağacı oluşturulmakta ve bu ağaçlar bir araya getirilerek bir karar ormanı tasarlanmaktadır (Zhou,2021). Karar ağaçları oluşturulurken her seviyede en fazla kullanılan öznitelik tüm ağaçlar için yapılan hesaplamalar sonucunda belirlenmekte ve ağaca eklenmektedir. Her bir karar ağacı düğümlerinde rastgele seçilen özniteliklerle en iyi kararları belirleyerek ve tüm düğümler dallara ayrılarak oluşturulmaktadır. Algoritmanın çalışması için her bir düğümde kullanılacak değişkenlerin sayısı ve oluşturulacak ağaçların sayısı kullanıcı tarafından belirlenmektedir. Ağaçlar gerçek veri setinden yer değiştirmeli olarak türetilen alt veri

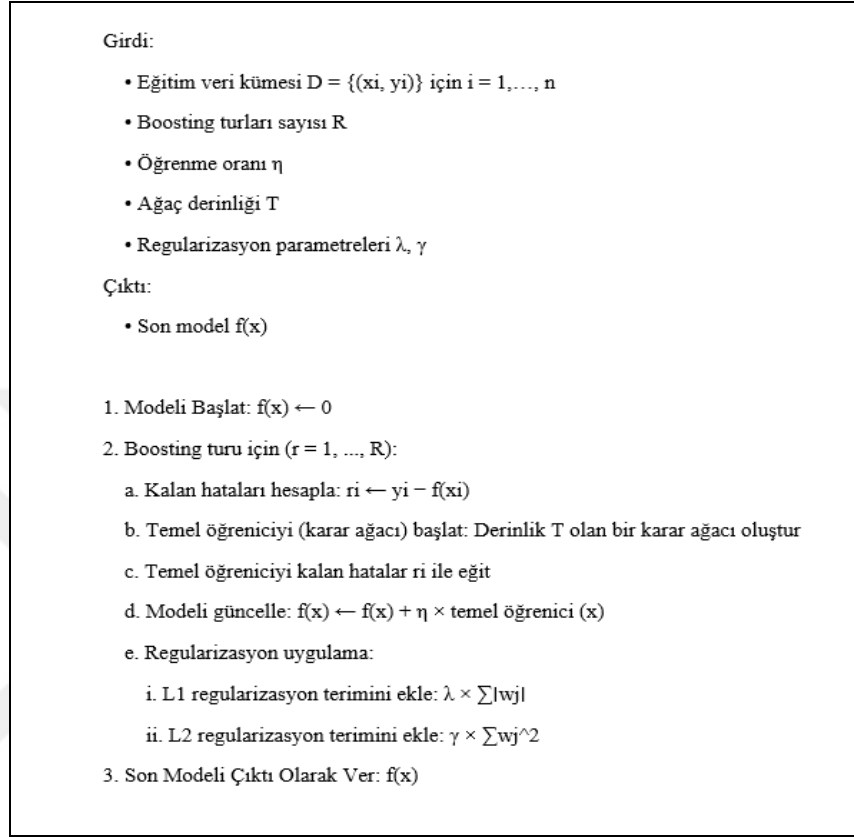
setleriyle oluşturulmaktadır. Bu yöntem RO algoritmasının diğer algoritmalara kıyasla daha hızlı sonuçlar üretmesini sağlamaktadır.

LR algoritması sınıflandırma problemlerinde kullanılan bir denetimli öğrenme yöntemidir. Temel amacı bir veri noktasının belirli bir sınıfa ait olma olasılığını tahmin etmektir (Zhou,2021). LR'nin çalışma prensibi veri kümesindeki her örneğin bir veya daha fazla öznelik ile temsil edilmesine dayanmaktadır. Bu özneliklerin ağırlıkları ve bir eşik değeri hesaplanmakta ve bu hesaplamalar örneklerin doğru sınıfa atanmasında kullanılmaktadır. Temel hedef en doğru sınıflandırmayı yapmak için en uygun ağırlıklar ve eşik değerini bulmaktır. Sınıflandırma sürecinde algoritma bağımsız değişkenler ile bağımlı değişken arasındaki ilişkiyi incelemekte ve buna göre karar vermektedir.

XGBoost algoritması karar ağaçları ile çalışan, hız, performans ve bellek kullanımı açısından optimize edilmiş, aşırı uyumu önleyici özelliklere sahip gelişmiş bir aşamalı artırma algoritmasıdır. Sınıflandırma ve regresyon problemlerinde oldukça etkili sonuçlar veren bu makine öğrenme yöntemi özellikle hızlı hesaplama yeteneğiyle öne çıkmaktadır. Boosting tekniği, mevcut ağaçların hatalarını düzeltmek amacıyla yeni ağaçlar eklemektedir. Temel olarak zayıf öğreniciler olarak adlandırılan birçok karar ağacı oluşturulmakta ve her ağaç bir önceki ağacın hatalarını düzelterek daha güçlü bir model oluşturmaktadır. XGBoost algoritması, L1 ve L2 regularizasyonunu kullanarak aşırı uyumu engellemekte, böylece eksik verilerde daha iyi performans göstererek daha hızlı sonuç üretmektedir (Chen ve Guestrin,2016). Bu süreçlerin daha iyi anlaşılabilmesi için XGBoost algoritmasının temel adımlarını gösteren bir sözde kod Şekil 2.8'de verilmiştir. XGBoost algoritması ağaç tabanlı öğrenme yöntemleri ile güçlü bir model oluşturmaktadır.

Şekil 2.8'deki sözde kodda görüldüğü gibi algoritma önce modelin başlangıç tahminini sıfır olarak başlatır. Ardından boosting turları boyunca her bir gözlem için tahmin edilen değer ile gerçek değer arasındaki hatalar hesaplanır. Ayrıca burada yer alan boosting turları ifadesi modelin eğitim sürecindeki her bir aşamayı ifade etmektedir. Bu hatalar modelin bir sonraki tahminini oluşturacak olan temel öğreniciyi yani karar ağacını eğitmek için kullanılır. Her boosting turunda karar ağacı önceki modelin hatalarını düzeltmeye yönelik öğrenir ve modelin tahminleri güncellenir. Ayrıca modelin aşırı uyum yapmasını engellemek amacıyla L1 ve L2 regularizasyon terimleri eklenir. L1 regularizasyonu modelin ağırlıklarının mutlak değerlerinin toplamını; L2 ise ağırlıkların karelerinin toplamını içerir. Bu adımlar her boosting turunda tekrarlanır ve süreç sonunda boosting turlarının tamamlanmasıyla elde edilen

son model $f(x)$ çıktı olarak verilir (Chen ve Guestrin,2016). Bu şekilde XGBoost algoritması hataları azaltarak modelin doğruluğunu artırır ve aşırı uyum yapmadan genelleme gücünü korur.



Şekil 2.8. XGBoost Algoritmasına Ait Sözde Kod

Kaynak:(Chen ve Guestrin,2016)

2.6. İstatistiksel Değerlendirme Ölçütleri

Çalışma sonucunda elde edilen sınıflandırma sonucunu ve kullanılan algoritmaların performansını değerlendirmek için çeşitli yaklaşımlar kullanılmaktadır. Bu çalışma kapsamında modelin performansını değerlendirmek için doğruluk, kesinlik, duyarlılık, F1-Skor, Makro Averaj (MAv), Ağırlıklandırılmış Averaj (AAv) ve ROC eğrisi altında kalan alan (AUC) metrikleri kullanılmıştır. Veri kümesine belirlenen algoritmalar uygulandığında sınıflandırma sonucunda dört muhtemel sonuç vardır. Bunlar: doğru pozitif (True Positive-TP), yanlış pozitif (False Positive-FP), doğru negatif (True Negative-TN) ve yanlış negatif (False Negative-FN) tahmin değerleridir.

Denklem 3'te verilen doğruluk, bir sınıflandırma modelinin doğru tahmin ettiği örneklerin oranını ifade eden bir performans metriğidir (Japkowicz ve Shah,2011).

$$\text{Doğruluk} = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

Denklem 4'te verilen kesinlik, sınıflandırma modelinin pozitif olarak tahmin ettiği örneklerin gerçekten pozitif olma oranını ifade eden metriktir (Japkowicz ve Shah,2011).

$$\text{Kesinlik} = TP / (TP + FP) \quad (4)$$

Denklem 5'te verilen duyarlılık, modelin gerçekten pozitif olan örnekleri doğru bir şekilde tahmin etme oranını gösteren metriktir (Japkowicz ve Shah,2011).

$$\text{Duyarlılık} = TP / (TP + FN) \quad (5)$$

Denklem 6'da verilen F1-Skor, hem kesinlik hem de duyarlılık arasında bir denge sağlar (Japkowicz ve Shah,2011). Modelin hem doğru pozitifleri doğru bir şekilde tanınması (yüksek kesinlik), hem de gerçek pozitiflerin kaçırılmaması (yüksek duyarlılık) önemlidir. F1-Skor bu iki metriğin birleşik bir değerlendirmesini sunmaktadır.

$$\text{F1-Skor} = 2 * (\text{kesinlik} * \text{duyarlılık}) / (\text{kesinlik} + \text{duyarlılık}) \quad (6)$$

MAV, sınıflandırma problemlerinde performans metriklerini hesaplamak için kullanılan bir yöntemdir. Bu yöntem, özellikle çok sınıflı problemlerde her bir sınıfın performansını bağımsız olarak değerlendirir ve tüm sınıfların ortalamasını alarak genel bir performans ölçümü sunar (Opitz,2024). Öncelikle her sınıf için kesinlik, hassasiyet veya F1-Skor gibi performans ölçütleri hesaplanır. MAV, her sınıfın metrik değerini sınıfın veri setindeki örnek sayısından bağımsız olarak aynı ağırlıkta dikkate almaktadır. Bu sayede veri setindeki küçük sınıfların performansı göz ardı edilmez ve tüm sınıflar eşit öneme sahip olur. Her sınıf için hesaplanan performans ölçütleri toplanır ve aritmetik ortalama alınarak modelin genel performansı belirlenir. MAV metriğinin hesaplama formülü Denklem 7'de verilmiştir.

$$\text{MAV} = (1/N) \sum_{i=1}^N \text{Metrik}_i \quad (7)$$

AAv, performans metriklerini sınıf örnek sayısına göre ağırlıklandırarak veri setindeki örnek dağılımını dikkate alır. Denklem 7'de gösterildiği üzere, her sınıfın metriği o sınıfın veri setindeki ağırlığıyla çarpılır ve bu değerler tüm sınıflar için toplanır. Sonrasında bu toplam değer sınıf ağırlıklarının toplamına bölünür. Böylece her sınıfın metriği veri setindeki ağırlığına göre değerlendirilmiş olur (Opitz,2024). Denklem 8'de AAv metriğinin hesaplanma formülünü sunulmaktadır.

$$AAv = \frac{\sum_{i=1}^N (\text{Metrik}_i \times \text{Sınıf Ağırlığı}_i)}{\sum_{i=1}^N \text{Sınıf Ağırlığı}_i} \quad (8)$$

ROC eğri grafiği, y eksenindeki TPR (duyarlılık) değerinin x eksenindeki FPR (özgüllük) değerine oranının çizilmesiyle elde edilmektedir. Grafik üzerinde ROC eğrisinin altında kalan alan AUC değerini ifade eder (Japkowicz ve Shah,2011). Bu değer 1'e ne kadar yakınsa model de o kadar başarılıdır demektir. ROC eğrisinin altında kalan alanın değerinin 0'dan 1'e doğru yaklaşması demek pozitif değerlerin negatif değerlerden başarılı bir şekilde ayrıldığı anlamına gelmektedir.

BÖLÜM III

BULGULAR VE TARTIŞMA

Bu bölümde çalışmada kullanılan makine öğrenmesi algoritmalarının SVM evrelerini tespit etmedeki başarısı üzerine elde edilen sonuçlar sistematik bir şekilde sunulmuş ve bu sonuçların doğruluğu ile güvenilirliği kapsamlı bir şekilde araştırılmıştır. Çalışma iki aşamada gerçekleştirilmiş olup tüm ayrıntılar bu bölümde belirtilmiştir.

3.1. Ham Veriyle Yapılan Sınıflama İşlemleri

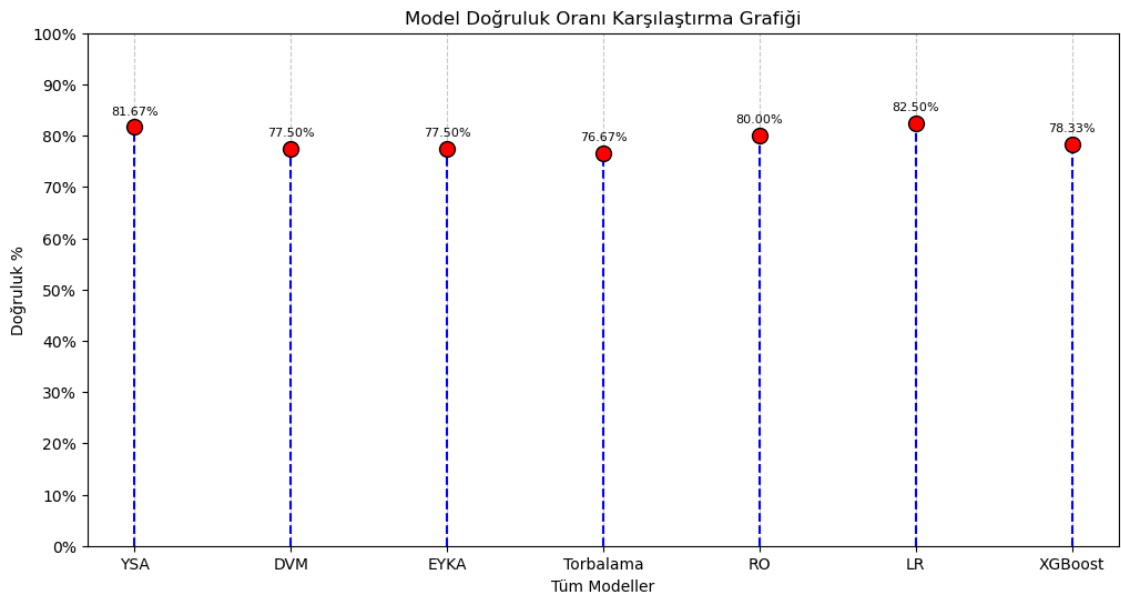
Oluşturulan çalışma modeli kapsamında büyüme ve gelişme tespitinde önemli olan 18 özellik belirlendikten sonra verileri temel alarak sınıflandırma modelleri belirlenmiştir. Üç sınıfa ayrılan bu veri seti YSA, DVM, EYKA, Torbalama, RO, LR, XGBoost olmak üzere toplamda 7 farklı sınıflayıcı ile değerlendirilmiştir. 18 özelliğe sahip 120 veriden oluşan veri seti kullanılarak ve 5-kat çapraz doğrulama yöntemi uygulanarak bir analiz gerçekleştirilmiştir. Bu analize ait performans çıktılarının modeller açısından ortalama (%) değerleri Tablo 3.1’de görülmektedir.

Tablo 3.1. Elde Edilen 18 Özellik ile Yapılan Ortalama Sınıflandırma Sonuçları

Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %
YSA	81.67	81.72	81.67	81.69
DVM	77.50	79.56	77.50	77.94
EYKA	77.50	78.87	77.50	77.89
Torbalama	76.67	77.75	76.67	76.86
RO	80	80	80	80
LR	82.50	82.27	82.50	82.33
XGBoost	78.33	78.31	78.33	78.31

Tablo 3.1’de 18 özellik kullanılarak gerçekleştirilen sınıflandırma sonuçları sunulmaktadır. Elde edilen sonuçlara göre, LR sınıflayıcısı en yüksek performansı sergileyerek %82.5 doğruluk oranına ulaşmıştır. LR’nin diğer metriklerde de yüksek performans göstermesi en etkili sınıflayıcı olduğunu göstermektedir. Bu sonuçlar LR’nin veri seti ile uyumlu olduğunu ve yüksek doğrulukla tahminler sağladığını ortaya koymaktadır. YSA ve RO, %81.67 ve %80 doğruluk oranlarıyla LR’nin ardından

gelmektedirler. Her iki algoritmanın F1-Skor oranı %80'nin üzerinde olup bu durum performanslarının LR'den daha düşük olduğunu ancak yine de iyi seviyede olduklarını göstermektedir. XGBoost algoritması %78.33 doğruluk oranı ile orta seviyede performans göstermiştir. DVM, EYKA ve Torbalama algoritmaları ise diğerlerine göre daha düşük performans sergilemişlerdir. Bu sonuçlar bu iki algoritmanın diğerlerine kıyasla veri setindeki sınıflama görevinde daha az etkili olduğunu göstermektedir. Bu sonuçlara göre en yüksek performansı sergileyen algoritma LR olmuştur. Tüm sınıflandırıcıların performansları toplu bir şekilde Şekil 3.1'de gösterilmiştir. Şekil 3.1 tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak sunmaktadır.



Şekil 3.1. Tüm Sınıflandırıcıların Doğruluk Oranları

3.2. VAE Yöntemi ile Örnek Çoğaltma ve Özellik Seçme İşlemleri Sonrası Yapılan Sınıflama İşlemleri

Birinci aşamadaki değerlendirmeler neticesinde başarı oranını arttırmak ve modelin genel performansını iyileştirmek amacıyla örnek çoğaltma işlemi gerçekleştirilmiştir. Veri setindeki her sınıfın eşit sayıda veri içermesi ve sınıf dengesizliğinin olmaması nedeniyle SMOTE yöntemleri kullanılmamıştır. Bu sebeple veri setinin büyüklüğünü ve çeşitliliğini arttırmak için VAE yöntemi tercih edilmiştir.

Öncelikle mevcut veri setinden 30 örnek alınarak, nihai test seti olarak ayrılmıştır. Daha sonra VAE yöntemi kullanılarak kalan veri setinin büyüklüğü 90 örnekten 180 örneğe çıkarılmıştır. Veri çoğaltma işleminin tamamlanmasının ardından

elde edilen veri seti üzerinde sayısal veri standardizasyonu yapılmıştır. Ardından Lasso yöntemi kullanılarak en etkin 7 özellik belirlenmiştir. Bu özellikler şunlardır:

- “2B noktasının, 2A ve 2C arasındaki doğruya olan mesafesi”,
- “3G, 3F ve 3E noktaları arasındaki açı”,
- “3G, 3A ve 3C noktaları arasındaki açı”,
- “3A, 3C ve 3E noktaları arasındaki açı”,
- “4F noktasının, 4G ve 4E arasındaki doğruya olan mesafesi”,
- “4G, 4A ve 4C noktaları arasındaki açı”,
- “4C ile 4F arası uzaklık ve 4G ile 4E arası uzaklık oranı”.

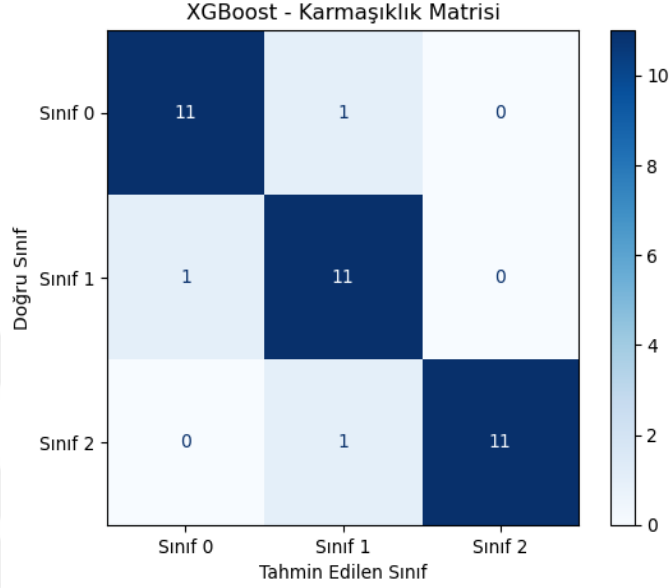
Çoğaltılan, standardizasyona sokulan ve özellik seçimi yapılan yeni veri seti 5-kat çapraz doğrulama yöntemi ile rastgele bölünmüş ve her bir kat aynı sınıflayıcılarla bağımsız olarak değerlendirilmiştir. Bu yöntem modelin performansını daha kapsamlı bir şekilde analiz etmek ve doğrulamak amacıyla uygulanmıştır. Bu aşamada doğrulama veri seti (1 kat=36 örnek) ve eğitim veri seti (4 kat=144 örnek) oluşturulmuştur. Bu bağlamda veri seti 5-kat çapraz doğrulama yöntemiyle ayrıldıktan sonra her bir kat için elde edilen sonuçlar ve her bir sınıflayıcının performansını ayrıntılı olarak gösteren analizler Tablo 3.2’ de sunulmuştur.

Tablo 3.2. Elde Edilen Birinci Kat Sonuçları

Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
YSA	83.33	83.41	83.33	83	83	83
DVM	80.56	80.32	80.56	79.76	79.76	79.76
EYKA	69.44	74.35	69.44	68.26	68.26	68.26
Torbalama	86.11	86.03	86.11	85.98	85.98	85.98
RO	86.11	86.03	86.11	85.98	85.98	85.98
LR	80.56	80.32	80.56	79.76	79.76	79.76
XGBoost	91.67	92.09	91.67	91.77	91.77	91.77

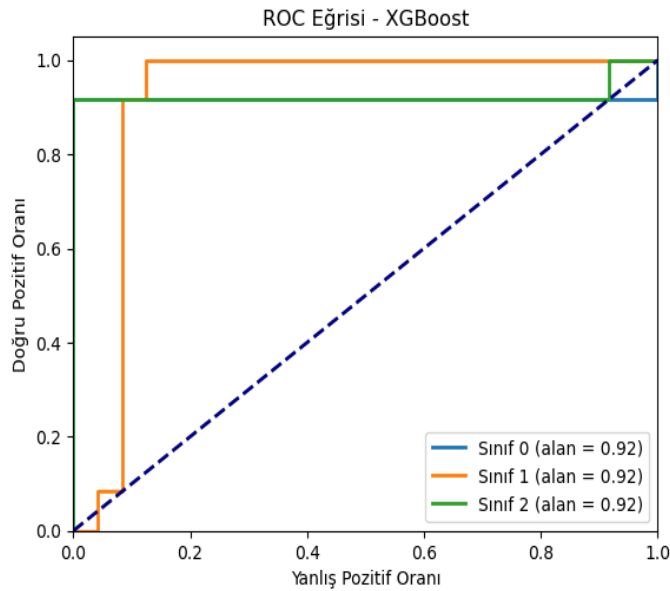
Tablo 3.2’ de, 5-kat çapraz doğrulama yöntemi ile elde edilen birinci kat sonuçları gösterilmektedir. Sonuçlara göre, XGBoost sınıflayıcısı %91.67 doğruluk oranı ile en yüksek performansı göstererek diğer sınıflayıcılardan belirgin şekilde üstün olmuştur. Diğer sınıflayıcılar arasında Torbalama ve RO oldukça iyi sonuçlar elde

etmiştir. Burada özellikle EYKA daha düşük performans göstermiştir. Bu bulgular, veri çoğaltma ve model değerlendirme süreçlerinin model performansını belirgin bir şekilde iyileştirdiğini ve XGBoost' un bu yöntemler arasında en yüksek başarıyı sağladığını göstermektedir. XGBoost' un etkili bir şekilde gerçekleştirdiği sınıflama işlemlerine dair detayları içeren karmaşıklık matrisi Şekil 3.2' de sunulmaktadır.



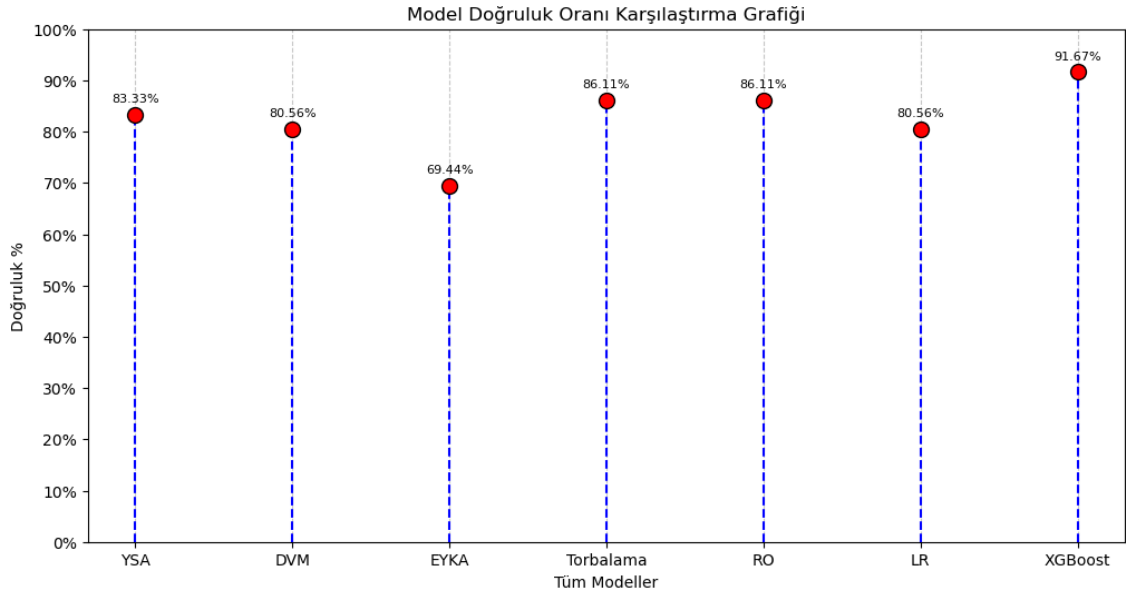
Şekil 3.2. Birinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi

XGBoost modelinin sınıflama performansını daha ayrıntılı bir şekilde değerlendirmek için ROC eğrisi Şekil 3.3' te gösterilmiştir.



Şekil 3.3. Birinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi

Şekil 3.4, tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır.



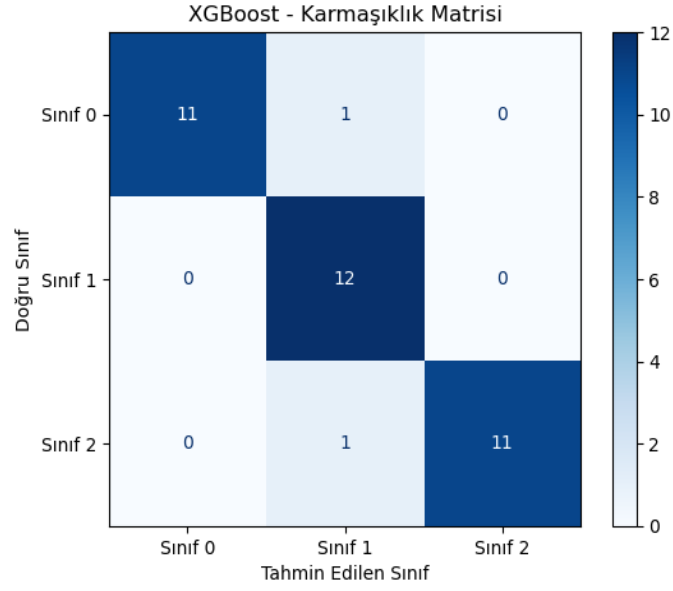
Şekil 3.4. Birinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları

Tablo 3.3' te, 5-kat çapraz doğrulama yöntemi ile elde edilen ikinci kat sonuçları gösterilmektedir.

Tablo 3.3. Elde Edilen İkinci Kat Sonuçları

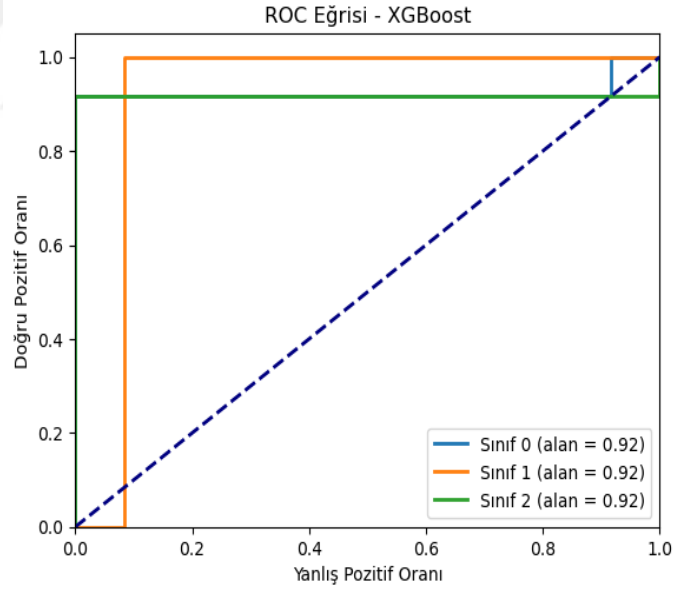
Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
YSA	77.78	79.76	77.78	78.38	78.38	78.38
DVM	83.33	83.08	83.33	82.91	82.91	82.91
EYKA	75	75.05	75	74.46	74.46	74.46
Torbalama	86.11	86.03	86.11	85.98	85.98	85.98
RO	88.89	89.32	88.89	89	89	89
LR	83.33	83.51	83.33	83.32	83.32	83.32
XGBoost	94.44	95.24	94.44	94.54	94.54	94.54

Sonuçlar XGBoost'un %94.44 doğruluk ve %94.54 F1-Skor ile en yüksek performansı sergilediğini ve diğer yöntemlere göre üstün olduğunu ortaya koymaktadır. RO algoritması ise etkileyici sonuçlar sunarak %88.89 doğruluk ve %89 F1-Skor elde edip, XGBoost'u takip etmiştir. Torbalama ise %86.11 doğruluk oranı ile başarılı bir performans göstermiştir. Ancak YSA ve EYKA daha düşük sonuçlarda kalmıştır. XGBoost'un karmaşıklık matrisi Şekil 3.5'te sunulmaktadır.



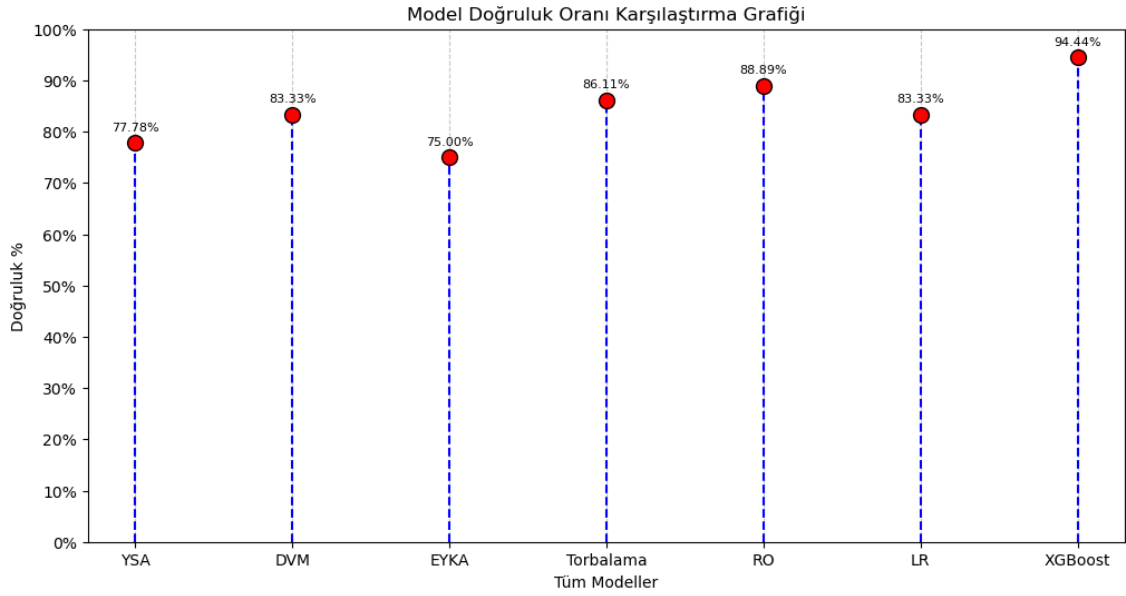
Şekil 3.5. İkinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi

XGBoost modelinin sınıflama performansını daha ayrıntılı bir şekilde değerlendirmek için ROC eğrisi Şekil 3.6' da gösterilmiştir.



Şekil 3.6. İkinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi

Şekil 3.7, tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır.



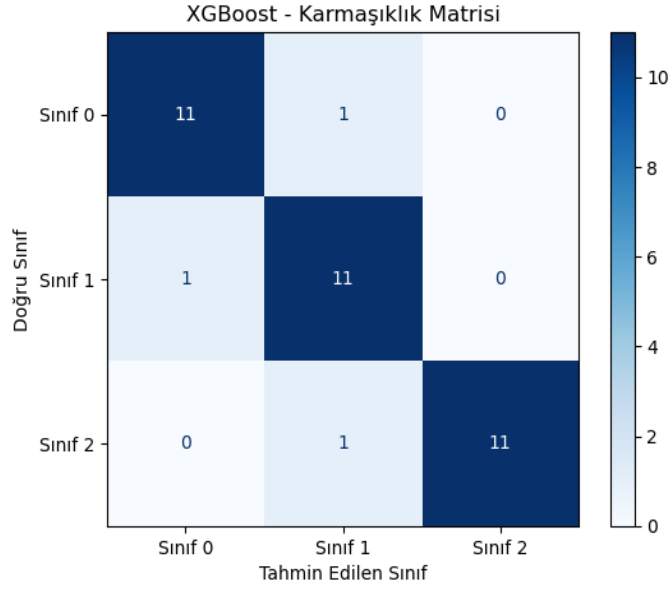
Şekil 3.7. İkinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları

Tablo 3.4' te, 5-kat çapraz doğrulama yöntemi ile elde edilen üçüncü kat sonuçları gösterilmektedir.

Tablo 3.4. Elde Edilen Üçüncü Kat Sonuçları

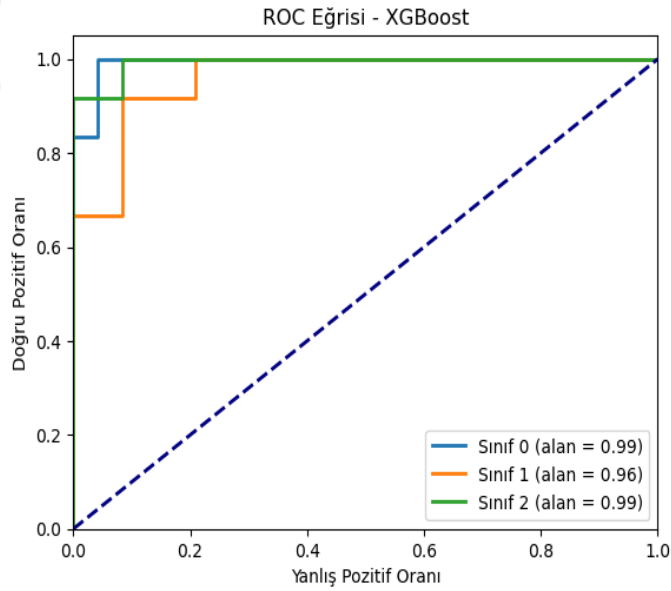
Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
YSA	77.78	78.25	77.78	77.66	77.66	77.66
DVM	88.89	88.97	88.89	88.61	88.61	88.61
EYKA	83.33	84.38	83.33	82.49	82.49	82.49
Torbalama	83.33	83.33	83.33	83.33	83.33	83.33
RO	86.11	86.5	86.11	86.21	86.21	86.21
LR	88.89	90.77	88.89	88.3	88.3	88.3
XGBoost	91.67	92.09	91.67	91.77	91.77	91.77

Sonuçlara göre, XGBoost sınıflayıcısı %91.67 doğruluk ile en yüksek performansı sergilemiştir. Bunun yanında, DVM ve LR algoritmaları güçlü performanslarıyla dikkat çekmektedir; ancak XGBoost'un mükemmel sonuçları bu sınıflayıcıların performansını geride bırakmıştır. XGBoost'un karmaşıklık matrisi Şekil 3.8'de sunulmaktadır.



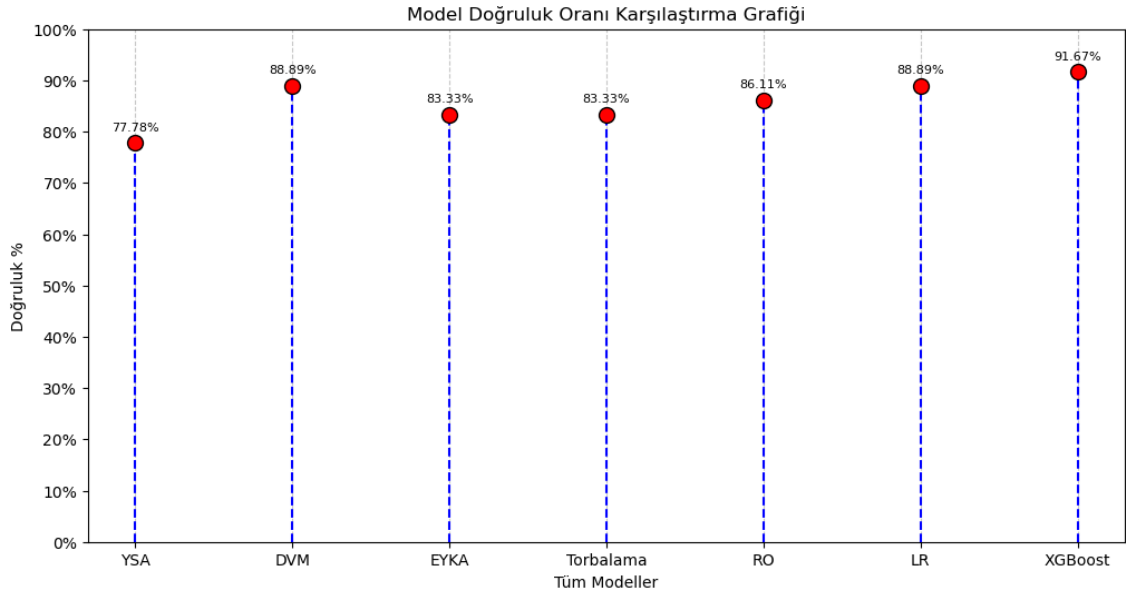
Şekil 3.8. Üçüncü Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi

XGBoost modelinin sınıflama performansını daha ayrıntılı bir şekilde değerlendirmek için ROC eğrisi Şekil 3.9’ da gösterilmiştir.



Şekil 3.9. Üçüncü Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi

Şekil 3.10, tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır.



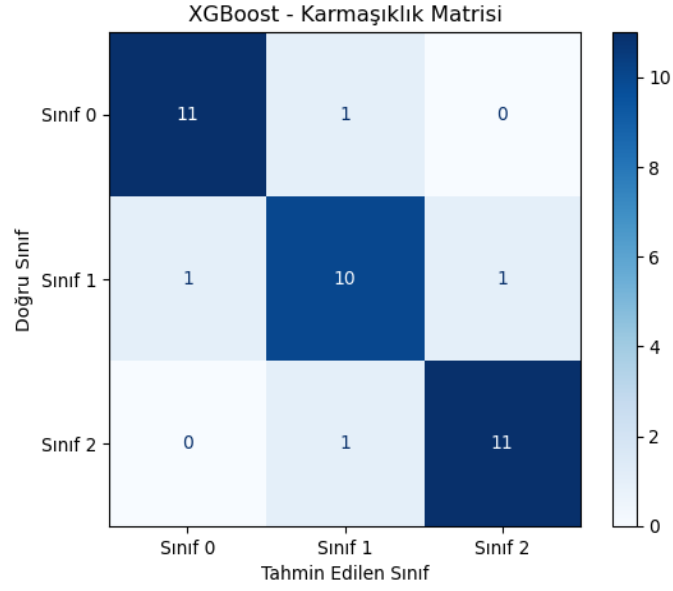
Şekil 3.10. Üçüncü Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları

Tablo 3.5' te, 5-kat çapraz doğrulama yöntemi ile elde edilen dördüncü kat sonuçları gösterilmektedir.

Tablo 3.5. Elde Edilen Dördüncü Kat Sonuçları

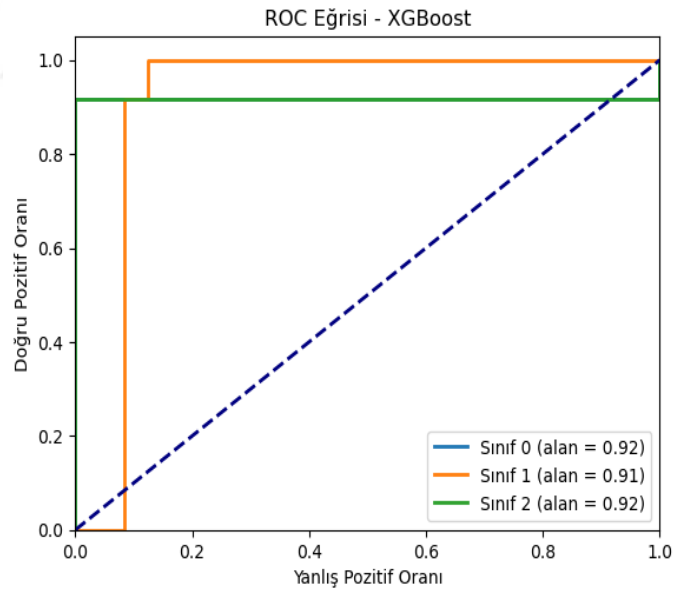
Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
YSA	86.11	86.8	86.11	86.18	86.17	86.17
DVM	86.11	87.18	86.11	86.3	86.3	86.3
EYKA	80.56	82.02	80.56	80.65	80.65	80.65
Torbalama	83.33	83.33	83.33	83.33	83.33	83.33
RO	86.11	87.18	86.11	86.3	86.3	86.3
LR	83.33	83.51	83.33	83.32	83.31	83.31
XGBoost	88.89	88.89	88.89	88.89	88.88	88.88

Dördüncü kat sonuçlarına göre, YSA, DVM ve RO algoritmaları %86.11 doğruluk oranlarıyla dördüncü kat için başarılı sonuçlar ortaya koymuştur. Ancak, XGBoost algoritması bu aşamada da tüm metrikler için %88.89 oranını elde ederek yüksek doğruluk ve F1-Skoru ile başarılı performansını sürdürmüştür. XGBoost'un karmaşıklık matrisi Şekil 3.11'de sunulmaktadır.



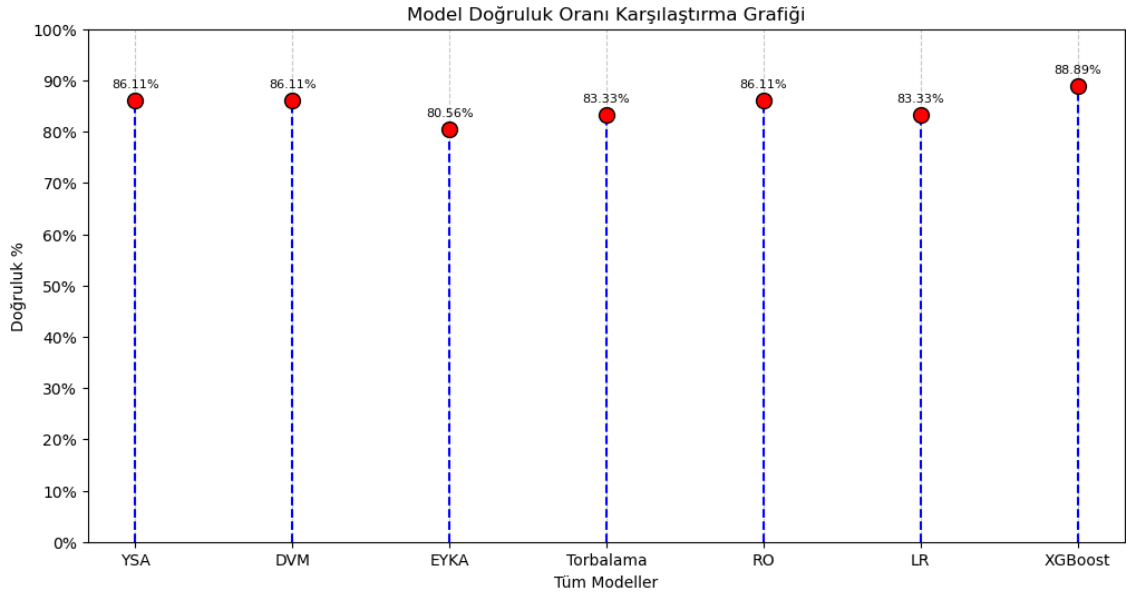
Şekil 3.11. Dördüncü Kat Sonuçlarına Göre LR Sınıflayıcısına Ait Karmaşıklık Matrisi

LR modelinin sınıflama performansını daha ayrıntılı bir şekilde değerlendirmek için ROC eğrisi Şekil 3.12’ de gösterilmiştir.



Şekil 3.12. Dördüncü Kat Sonuçlarına Göre LR Sınıflayıcısına Ait ROC Eğrisi

Şekil 3.13, tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır.



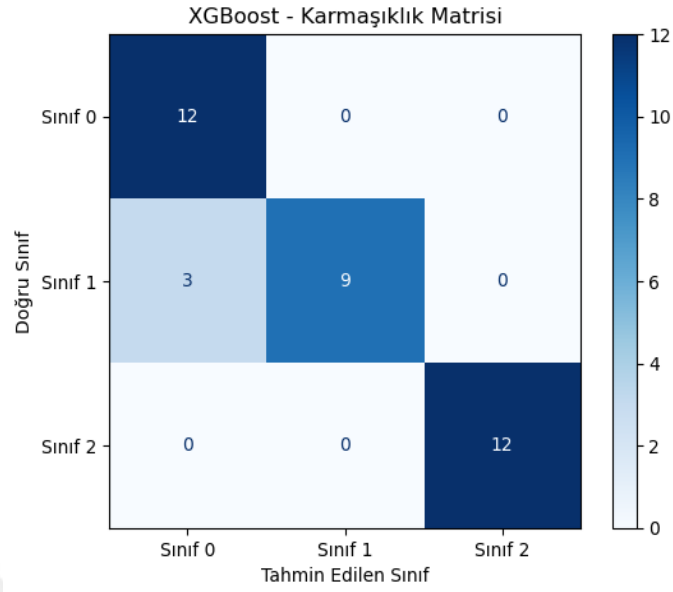
Şekil 3.13. Dördüncü Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları

Tablo 3.6' da 5 kat çapraz doğrulama yöntemi ile elde edilen beşinci kat sonuçları gösterilmektedir.

Tablo 3.6. Elde Edilen Beşinci Kat Sonuçları

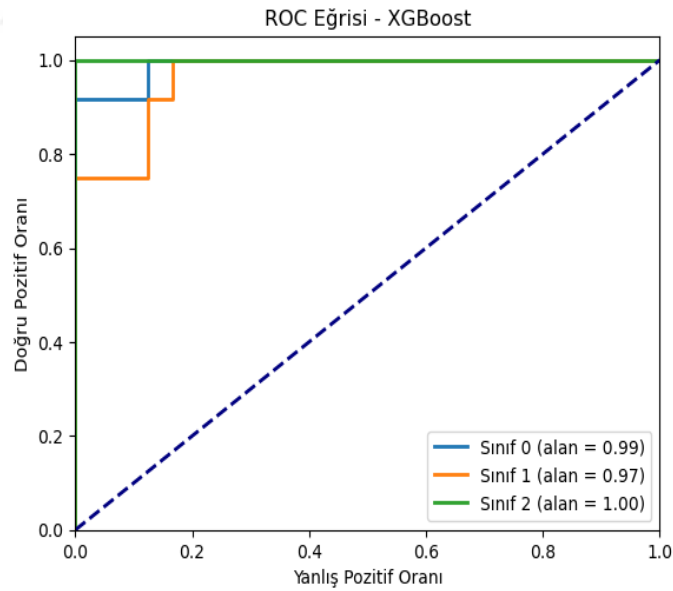
Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
YSA	77.78	77.3	77.78	77.2	77.19	77.19
DVM	83.33	84.04	83.33	82.3	82.29	82.29
EYKA	75	75.12	75	72.91	72.9	72.9
Torbalama	83.33	83.93	83.33	82.31	82.3	82.3
RO	83.33	84.04	83.33	82.3	82.29	82.29
LR	83.33	84.04	83.33	82.3	82.29	82.29
XGBoost	91.67	93.33	91.67	91.53	91.53	91.53

Elde edilen sonuçlara göre, XGBoost'un %91.67 doğruluk ve %91.53 F1-Skoru ile en yüksek performansı göstererek diğer yöntemlerden belirgin bir şekilde ayrıldığını ortaya koymaktadır. Bu algoritmayı, %83.33 doğruluk ve %82.3 F1-Skoru DVM, Torbalama, RO ve LR algoritmaları takip etmiştir. Genel olarak XGBoost algoritmasından elde edilen sonuçlar bu veri seti üzerindeki en etkili sınıflayıcı olduğunu doğrulamaktadır. XGBoost'un beşinci kata ait karmaşıklık matrisi Şekil 3.14'te sunulmaktadır.



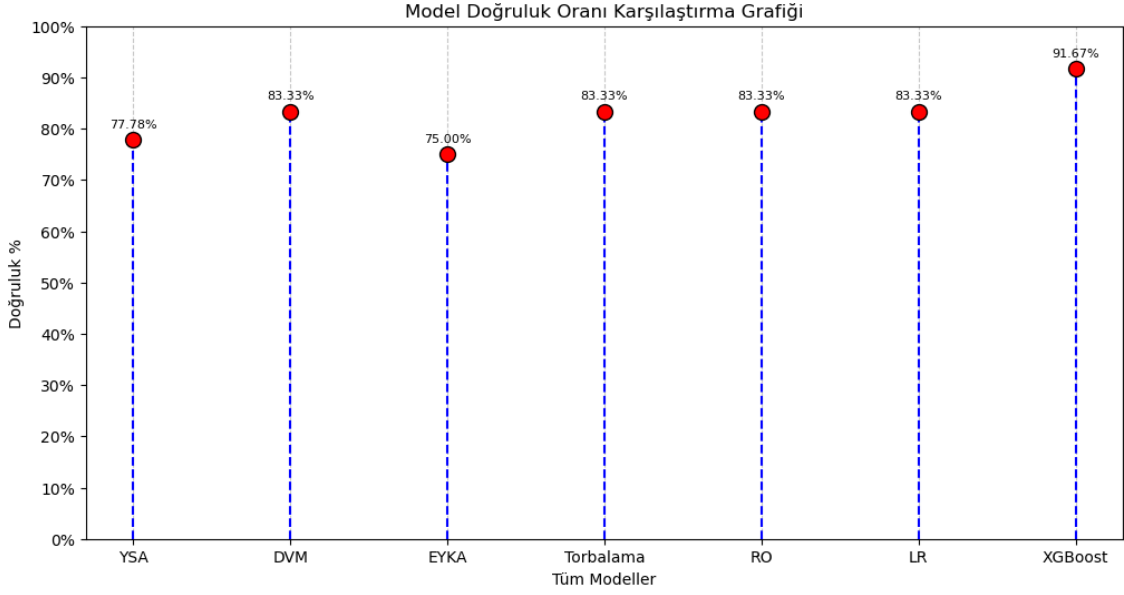
Şekil 3.14. Beşinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait Karmaşıklık Matrisi

XGBoost modelinin sınıflama performansını daha ayrıntılı bir şekilde değerlendirmek için ROC eğrisi Şekil 3.15’ te gösterilmiştir.



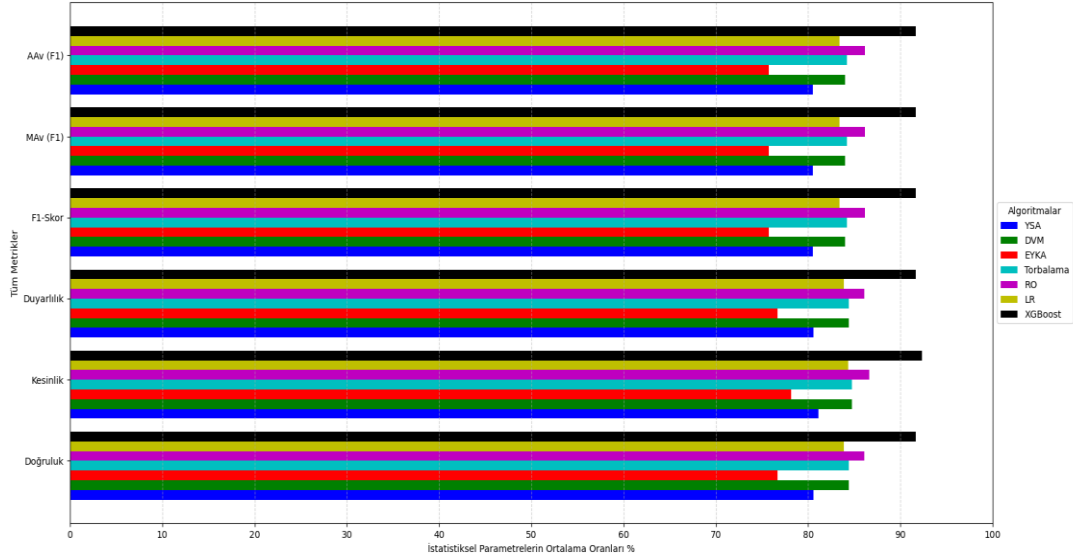
Şekil 3.15. Beşinci Kat Sonuçlarına Göre XGBoost Sınıflayıcısına Ait ROC Eğrisi

Şekil 3.16, tüm algoritmaların doğruluk oranlarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır.



Şekil 3.16. Beşinci Kat Sonuçlarına Göre Tüm Sınıflandırıcıların Doğruluk Oranları

Görüldüğü üzere VAE ile veri çoğaltma işlemi, veri setinin temsil yeteneğini ve modelin genelleme kapasitesini artırarak modelin başarısını olumlu yönde etkilemiştir. Bu işlem modelin öğrenme sürecini zenginleştirerek sınıflama doğruluğunu ve performansını arttırmaya yönelik önemli bir adım olmuştur. Ayrıca, veri standardizasyonu verilerin tüm özelliklerinin benzer bir ölçek üzerinde normalize edilmesini sağlayarak modelin performansını arttırmıştır. Bu sayede farklı ölçeklere sahip verilerin sayısal değerleri arasındaki dengesizlikler ortadan kaldırılmış olup modelin öğrenme süreci daha tutarlı hale getirilmiştir. Bunun yanında, Lasso yöntemi sayesinde ise etkin özellikler belirlenerek sistemin gereğinden fazla efor harcaması engellenmiştir. Ayrıca yapılan bu işlemler, SVM sınıflamada kullanılan 7 farklı makine öğrenmesi algoritmasının başarısına önemli katkıda bulunmuştur. Şekil 3.17, kullanılan tüm sınıflama algoritmalarının 5-kat çapraz doğrulama sonucundaki ortalama performans çıktılarını karşılaştırmalı olarak toplu bir şekilde sunmaktadır. Şekil 3.17 incelendiğinde, VAE ile örnek artırımı ve Lasso yöntemi ile özellik seçimi sonrasında yapılan sınıflandırma işlemlerinde tüm istatistiksel parametrelerde en başarılı sonuçların alındığı makine öğrenmesi algoritması XGBoost olmuştur. XGBoost sınıflandırıcısını ise yine tüm parametrelerdeki başarısıyla RO algoritması izlemiştir.



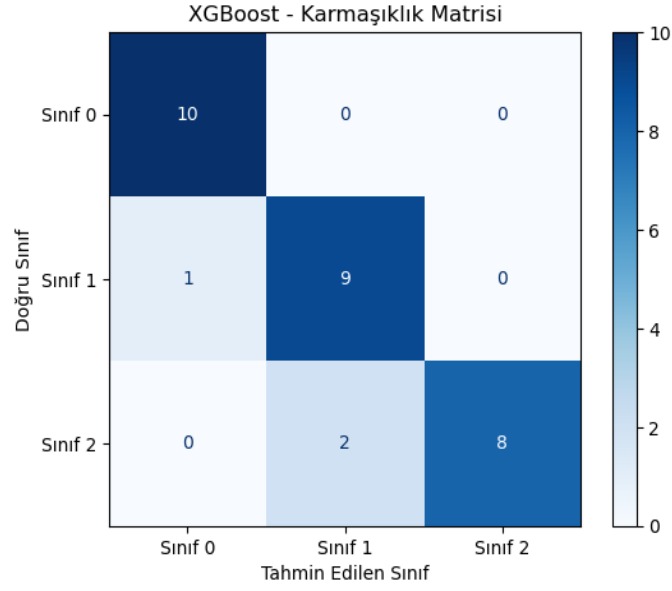
Şekil 3.17. Tüm Sınıflandırıcıların 5-Kat Çapraz Doğrulama Sonucundaki Ortalama Performans Çıktıları

Son aşamada, çalışmanın başında belirlenen ve 30 örnekten oluşan test veri seti, en iyi performans çıktılarının elde edildiği XGBoost algoritmasına sunulmuş ve sonuçlar sınıflar bazında kaydedilmiştir. Tablo 3.7' de bu veri setinin eğitilen XGBoost modeliyle sınıflandırılması sonucunda elde edilen performans çıktıları gösterilmektedir.

Tablo 3.7. Başlangıçta Ayrılan 30 Örneğin Eğitilmiş XGBoost Modeliyle Sınıflandırılma Sonuçları

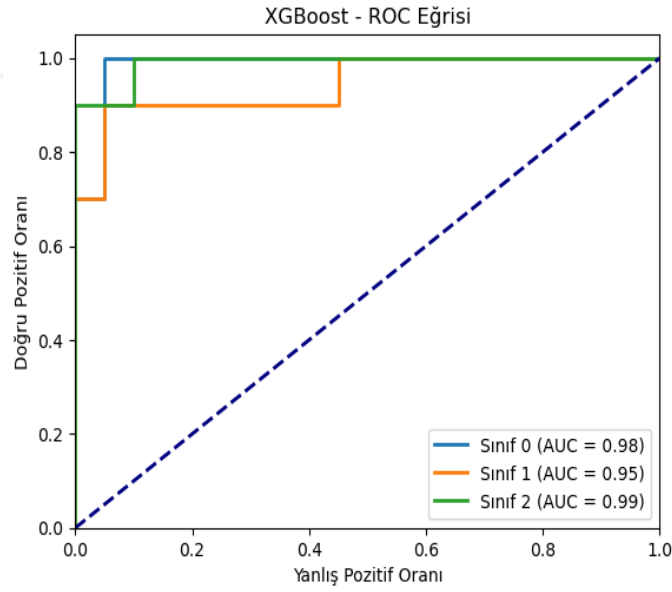
Sınıflandırıcı	Doğruluk %	Kesinlik %	Duyarlılık %	F1-Skor %	MAv (F1) %	AAv (F1) %
Sınıf 0	100	90	100	95.23	95.23	95.23
Sınıf 1	90	81.81	90	85.71	85.71	85.71
Sınıf 2	80	100	80	88.88	88.88	88.88
Ortalama	90	90.9	90	89.94	89.94	89.94

Tablo 3.7' de belirtilen sonuçlara göre, XGBoost %90 doğruluk ve %89.94 F1-Skoru ile hiç karşılaşmadığı 30 örnekli veriyi bile başarılı bir şekilde ayırt edebildiğini ortaya koymaktadır. Bunun yanında diğer tüm parametrelerde de ulaştığı güçlü sonuçlar görülmektedir. Bu durum, XGBoost algoritmasının bu veri seti üzerinde etkili sınıflayıcı olduğunu bir kez daha kanıtlamaktadır. XGBoost'un başlangıçta ayrılan 30 örnekli veri setini sınıflamasına ait karmaşıklık matrisi Şekil 3.18' de sunulmaktadır.



Şekil 3.18. XGBoost'un Başlangıçta Ayrılan 30 örnekli Veri Setini Sınıflamasına Ait Karmaşıklık Matrisi

XGBoost'un başlangıçta ayrılan 30 örnekli veri setini sınıflamasına ait ROC eğrisi Şekil 3.19' da gösterilmiştir.



Şekil 3.19. XGBoost'un Başlangıçta Ayrılan 30 Örnekli Veri Setini Sınıflamasına Ait ROC Eğrisi

Bölüm 3'e ait tüm tablolar ve şekiller ışığında elde edilen bütün sonuçlar ayrıntılı bir şekilde analiz edildiğinde, veri ön işleme adımları öncesinde XGBoost algoritmasının ortalama doğruluk oranı %78.33 iken, veri ön işleme adımları sonrasında

bu oranın %90'a ulaştığı görülmektedir. Aynı şekilde belirtilen ön işlemler öncesinde ortalama F1 skor %78.31 iken, işlemler sonrasında neredeyse %90'a ulaşmaktadır. Veri ön işleme adımları öncesinde XGBoost algoritmasının kesinlik oranı %78.31, duyarlılık oranı ise %78.33 iken veri ön işleme sonrasında bu oranlarda da önemli bir iyileşme gözlemlenmiştir. Bu noktada kesinlik oranı %90.9'a, duyarlılık oranı ise %90'a yükselmiştir. AAv ve MAV metrikleri de dikkatle incelendiğinde her ikisinin de neredeyse %90'a ulaştığı görülmektedir. AAv, sınıf dengesizliği olan veri setlerinde daha sık görülen sınıfların etkisini hesaba katarak genel performansı daha iyi yansıtır. MAV ise her bir sınıfın eşit ağırlıkta değerlendirildiği bir metrik olduğu için, daha dengeli bir değerlendirme yapar. XGBoost algoritmasında her iki metriğin de yüksek değerlerde olması, algoritmanın veri setindeki tüm sınıflara karşı dengeli bir performans sergilediğini göstermektedir. Bu bulgular, XGBoost algoritmasının veri ön işleme adımlarından sonra her açıdan daha güçlü bir performans sergilediğini ve metriklerin her birinde iyileşme sağlandığını ortaya koymaktadır.

İkinci en başarılı model olan RO algoritmasının veri ön işleme adımları öncesindeki ortalama doğruluk, kesinlik ve duyarlılık oranları %80 seviyesindeydi. Ancak veri ön işleme sonrasında bu oranlarda önemli bir iyileşme sağlanarak doğruluk %86.11'e, kesinlik %86.61'e ve duyarlılık %85.96'ya yükselmiştir. F1-Skoru açısından da benzer bir iyileşme gözlemlenmiştir. Veri ön işleme adımları öncesinde %80 olan F1-Skoru, ön işlem sonrasında %85.96'ya çıkmıştır. AAv ve MAV metrikleri incelendiğinde, her iki metriğin de %85.96 ile yüksek bir değere ulaştığı gözlemlenmektedir. Bu bulgular, RO algoritmasının da veri ön işleme adımları sonrasında tüm metriklerde dikkate değer bir performans artışı sağladığını göstermektedir.

Üçüncü en başarılı model olan Torbalama algoritmasının veri ön işleme adımları öncesindeki doğruluk oranı %76.67, kesinlik oranı %77.75, duyarlılık oranı ise %76.67 idi. Ancak veri ön işleme adımları sonrasında yine bu oranlarda önemli bir iyileşme sağlanmıştır. Doğruluk oranı %84.44'e, kesinlik oranı %84.53'e, duyarlılık oranı ise %84.44'e yükselmiştir. F1-Skoru açısından da veri ön işleme adımları öncesinde F1 skoru %76.86 iken, ön işlem sonrasında bu oran %84.19'a çıkmıştır. AAv ve MAV metrikleri de dikkatlice incelendiğinde, AAv %84.19' luk yüksek bir değere ulaşırken, MAV de aynı başarıyı göstererek %84.19 seviyesinde kalmaktadır. Torbalama algoritması tüm bu metriklerindeki artışla birlikte dengeli bir başarı sergileyerek başarılı bir genel model olmuştur.

Tüm metrikler dikkate alındığında, veri ön işleme adımları öncesinde algoritmaların performanslarının genel olarak daha düşük olduğu ve bu adımlardan sonra anlamlı ve belirgin bir artış yaşandığı gözlemlenmektedir.

Model bazında en düşük istatistiksel parametre değerlerinin elde edildiği EYKA algoritması incelendiğinde, gerçekleştirilen ön işlemlerin modelin performansını olumsuz yönde etkilediği gözlemlenmektedir. Başlangıçta EYKA modeli kullanılarak elde edilen %77.5 ortalama doğruluk oranı ve %77.89 ortalama F1-Skor değeri, veri ön işleme adımları sayesinde sırasıyla %76.67'ye ve %75.75'e gerilemiş olsa da, bu değişikliklerin modelin genel performansını iyileştirmeye yönelik bir katkı sağladığını ortaya koymaktadır. Veri ön işleme süreci modelin daha verimli ve anlamlı sonuçlar üretmesine olanak tanıyarak, bu tezdeki hedeflenen başarıya ulaşmada önemli bir rol oynamıştır. Bu durum, yapılan tüm işlem adımlarının bu tez çalışmasında amaçlanan başarıya ulaşmada ne kadar anlamlı ve etkili olduğunu kanıtlamaktadır.

BÖLÜM IV

SONUÇ VE ÖNERİLER

Büyüme ve gelişme evresi tespitinde karşılaşılan zorlukların üstesinden gelmek için disiplinler arası bir yaklaşım benimsemek önemlidir. Farklı uzmanlık alanlarından gelen bilgiler, büyüme ve gelişim süreçlerini daha iyi anlamamıza yardımcı olabilmektedir. Bu tür bir yaklaşım, yalnızca klinik uygulamaların etkinliğini arttırmakla kalmayıp, aynı zamanda araştırmaların kalitesini ve geçerliliğini de yükseltmektedir. Uzmanlar arasındaki sürekli etkileşim ve bilgi paylaşımı, alanın gelişimine önemli katkılarda bulunabilmekte ve ortaya çıkan bulguların uygulamaya geçmesini hızlandırabilmektedir. Bunun yanında, makine öğrenimi teknolojilerinin yapılan çalışmalara entegrasyonu, geleneksel yöntemlerin ötesinde yenilikçi çözümler sunma potansiyeline sahiptir. Oluşturulan uzman sistemlerin, sağlık profesyonellerinin karar verme süreçlerini kolaylaştıracak şekilde tasarlanması, kişiselleştirilmiş tedavi yaklaşımlarının geliştirilmesine olanak tanımaktadır. Bu sayede hastaların bireysel özelliklerine göre özelleştirilmiş tedavi planları oluşturmak, tedavi süreçlerinin daha başarılı ve etkili bir şekilde yürütülmesine katkıda bulunabilmektedir. Böylelikle, sağlık hizmetleri kalitesi artarken, hastaların memnuniyeti de önemli ölçüde yükseltilebilmektedir.

Benzer bir amaçla bu tez çalışmasında, sefalometrik radyografilerde gözlemlenen servikal vertebraların analizinin yapılmasıyla alana özgü olarak seçilmiş makine öğrenme algoritmaları kullanılarak SVM evresi tespiti gerçekleştirilmiştir. Bu çalışmanın amacı, hastaların daha iyi ve kişiye özel tedavi imkânlarına erişimini artırabilirken aynı zamanda sağlık profesyonellerine de daha doğru ve etkili bir çalışma ortamı sunmaktır. Yapılan analizler sonucunda, SVM tespiti kapsamında seçilen makine öğrenmesi modellerinden XGBoost ile %90 oranında yüksek bir ortalama doğruluk oranı elde edilmiştir. Ayrıca bu başarıya 3 sınıflı sınırlı sayıdaki veri ile ulaşılmıştır. Çalışmanın literatüre en önemli katkılarından biri, veri setinin boyutunu ve çeşitliliğini arttıran VAE metodunu kullanılarak gerçekleştirilen veri çoğaltma işlemi olmuştur. Bu sayede küçük veri setiyle çalışmanın getirdiği sınırlamalar aşılmış ve modelin genelleme yeteneği artırılmıştır. Bunun yanında, k-kat çapraz doğrulama sayesinde de kullanılan modellerin performansı daha güvenilir bir şekilde değerlendirilmiş ve aşırı uyum riski azaltılmıştır. Bunun yanında, Lasso yöntemi ile seçilen 7 etkin özellik

sınıflama sürecinde modelin hesaplama maliyetini azaltıp, doğruluğunu arttırmada önemli bir rol oynamıştır. Tüm sayılan bu veri ön işleme adımları, çalışmanın özgünlüğünü desteklemiş ve literatüre bu alanda farklı bir bakış açısı kazandırmıştır. Bu sonuçlara göre çalışma kapsamındaki veri setinin daha geniş ve çeşitli bir veri seti olması ve aynı zamanda daha karmaşık modellerin değerlendirilmesi neticesinde çok daha başarılı sonuçlara ulaşılabileceği öngörülmektedir. Elde edilen bulgular klinik uygulamalarda büyük potansiyele sahiptir ve ortodontik tedavi planlamasında önemli bir rol oynayabilir. Ayrıca makine öğrenme algoritmalarının kullanımıyla ortaya çıkan bu metodoloji, büyüme-gelişme tespiti alanında yeni bir perspektif sunarak gelecek araştırmalar için önemli bir temel oluşturabilir. Bu çalışmanın diş hekimliği alanındaki bilgi birikimini genişleterek bireylerin sağlığına daha etkili ve kişiye özel yaklaşımlar geliştirilmesine önemli katkılarda bulunacağına inanılmaktadır.

Gelecek araştırmalar, makine öğrenimi algoritmalarının farklı disiplinlerde nasıl entegre edilebileceğine dair yenilikçi yaklaşımlar geliştirmeyi hedeflemelidir. Bu disiplinler arası iş birliği, hastaların sağlık durumlarının çok yönlü bir perspektiften incelenmesini sağlayarak, daha doğru tanı ve tedavi yöntemlerinin geliştirilmesine katkıda bulunabilir. Özellikle, veri setlerinin büyütülmesi ve çeşitlendirilmesi, modelin genelleme yeteneğini artıracak ve daha fazla hasta grubu üzerinde test edilmesini sağlayacaktır. Gelecek çalışmalarda, gerçek dünya verilerinin entegrasyonu ile oluşturulacak daha karmaşık ve derin öğrenme temelli modeller, sonuçların doğruluğunu daha da artırabilir. Bu tür bir yenilik, ortodontik tedavi süreçlerinin yanı sıra, genel sağlık uygulamalarında da daha etkili stratejilerin geliştirilmesine olanak tanıyacaktır. Sonuç olarak, büyüme ve gelişme evresi tespitinde kullanılacak yenilikçi yöntemlerin geliştirilmesi, yalnızca bireylerin sağlık durumlarını değil, aynı zamanda toplum sağlığını da olumlu yönde etkileyebilir. Bu çalışmanın sunduğu yöntemler ve bulgular, gelecekteki araştırmalara ilham vererek, ortodonti alanında daha etkili tedavi seçeneklerinin ortaya çıkmasına zemin hazırlayacaktır. Sağlık profesyonellerinin bu teknolojik yenilikleri benimsemesi, bireysel tedavi planlamalarında daha doğru ve verimli kararlar alınmasına katkıda bulunacaktır. Bu çerçevede yapılacak yeni çalışmalar, disiplinler arası bir yaklaşım ve makine öğreniminin entegrasyonu ile, sağlık alanında devrim niteliğinde yenilikler sunabilir.

KAYNAKÇA

- Akkoyun, M., (2013). Adli tıp açısından yaş tayini istenen olguların değerlendirilmesi. *Uzmanlık Tezi, TC Akdeniz Üniversitesi Tıp Fakültesi Adli Tıp Anabilim Dalı, Antalya.*
- Amasya, H., Yildirim, D., Aydoğan, T., Kemaloglu, N., & Orhan, K., (2020). Cervical vertebral maturation assessment on lateral cephalometric radiographs using artificial intelligence: comparison of machine learning classifier models. *Dentomaxillofacial Radiology, 49(5)*, 20190441.
- Baccetti, T., Franchi, L., & McNamara Jr, J. A., (2002). An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. *The Angle Orthodontist, 72(4)*, 316-323.
- Baccetti, T., Franchi, L., & McNamara Jr, J. A., (2005). The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. *Seminars in Orthodontics,*
- Bench, R. W., (1963). Growth of the cervical vertebrae as related to tongue, face, and denture behavior. *American Journal of Orthodontics, 49(3)*, 183-214.
- Bengio, Y., Courville, A., & Vincent, P., (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence, 35(8)*, 1798-1828.
- Bica, C., & Mariana, P., (2010). Relations between cervical vertebral maturation and chronological age. *Romanian Journal of Oral Rehabilitation, 2(1)*, 4-7.
- Carty, H., (2002). Assessment of skeletal maturity and prediction of adult height (TW3 method).: Edited by JM Tanner, MJR Healy, H. Goldstein and N. Cameron. Pp 110. London, etc: WB Saunders, 2001. ISBN: 0-7020-2511-9.£ 69.95. *The Journal of Bone & Joint Surgery British Volume, 84(2)*, 310-311.
- Chatzigianni, A., & Halazonetis, D. J., (2009). Geometric morphometric evaluation of

cervical vertebrae shape and its relationship to skeletal maturation. *American Journal of Orthodontics and Dentofacial Orthopedics*, 136(4), 481. e481-481. e489.

Chen, T., & Guestrin, C., (2016). Xgboost: A scalable tree boosting system.

Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining,

Demirjian, A., & Goldstein, H., (1976). New systems for dental maturity based on seven and four teeth. *Annals of human biology*, 3(5), 411-421.

Ding, M., (2022). The road from MLE to EM to VAE: A brief tutorial. *AI Open*, 3, 29-34.

El-Askary, F. S., El-Banna, A. H., & Van Noort, R., (2012). Immediate vs delayed repair bond strength of a nanohybrid resin composite. *Journal of Adhesive Dentistry*, 14(3), 265.

EMEÇ, M., & ÖZCANHAN, M. H., (2023). Makine Öğrenmesi Algoritmalarında Hiper Parametre Belirleme.

Erverdi, N., (1988). Dis yası, kemik yası ve kronolojik yas arası iliskinin arastırılması.

Investigation of the relationship between dental, skeletal and chronological age.) J Turk Orthod Soc, 1, 30-41.

Fishman, L. S., (1982). Radiographic evaluation of skeletal maturation: a clinically oriented method based on hand-wrist films. *The Angle Orthodontist*, 52(2), 88-112.

Flores-Mir, C., Burgess, C. A., Champney, M., Jensen, R. J., Pitcher, M. R., & Major, P. W., (2006). Correlation of skeletal maturation stages determined by cervical vertebrae and hand-wrist evaluations. *The Angle Orthodontist*, 76(1), 1-5.

Flores-Mir, C., Nebbe, B., & Major, P. W., (2004). Use of skeletal maturation based on hand-wrist radiographic analysis as a predictor of facial growth: a systematic review. *The Angle Orthodontist*, 74(1), 118-124.

- Fonti, V., & Belitser, E., (2017). Feature selection using lasso. *VU Amsterdam research paper in business analytics*, 30, 1-25.
- Franchi, L., Baccetti, T., & McNamara Jr, J. A., (2000). Mandibular growth as related to cervical vertebral maturation and body height. *American Journal of Orthodontics and Dentofacial Orthopedics*, 118(3), 335-340.
- Fudalej, P., & Bollen, A.-M., (2010). Effectiveness of the cervical vertebral maturation method to predict postpeak circumpubertal growth of craniofacial structures. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(1), 59-65.
- Garcia-Fernandez, P., Torre, H., Flores, L., & Rea, J., (1998). The cervical vertebrae as maturational indicators. *Journal of clinical Orthodontics*, 32(4), 221-226.
- GENÇAY, D. K., & KASIMOĞLU, Y., (2016). TÜRK ÇOCUKLARINDA KEMİK YAŞI TAYİNİNDE KULLANILAN ÜÇ YÖNTEMİN KARŞILAŞTIRMALI.
- Grave, K., & Townsend, G., (2003). Cervical vertebral maturation as a predictor of the adolescent growth spurt. *Australasian Orthodontic Journal*, 19(1), 25-32.
- Greulich, W. W., & Pyle, S. I., (1959). Radiographic atlas of skeletal development of the hand and wrist. *The American Journal of the Medical Sciences*, 238(3), 393.
- Gu, Y., & McNamara Jr, J. A., (2007). Mandibular growth changes and cervical vertebral maturation: a cephalometric implant study. *The Angle Orthodontist*, 77(6), 947-953.
- Hägg, U., & Taranger, J., (1982). Maturation indicators and the pubertal growth spurt. *American Journal of Orthodontics*, 82(4), 299-309.
- Hassel, B., & Farman, A. G., (1995). Skeletal maturation evaluation using cervical vertebrae. *American Journal of Orthodontics and Dentofacial Orthopedics*, 107(1), 58-66.
- Hellsing, E., (1991). Cervical vertebral dimensions in 8-, 11-, and 15-year-old children. *Acta Odontologica Scandinavica*, 49(4), 207-213.

- Hinton, G. E., & Salakhutdinov, R. R., (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- İzgi, M. S., & Kök, H., (2020). Kemik yaşı ve maturasyon tespiti. *Selcuk Dental Journal*, 7(1), 124-133.
- Japkowicz, N., & Shah, M., (2011). *Evaluating learning algorithms: a classification perspective*. Cambridge University Press.
- Jayaraman, J., Wong, H. M., King, N. M., & Roberts, G. J., (2013). The French–Canadian data set of Demirjian for dental age estimation: a systematic review and meta-analysis. *Journal of forensic and legal medicine*, 20(5), 373-381.
- Kingma, D. P., (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
- Kök, H., Acilar, A. M., & İzgi, M. S., (2019). Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Progress in Orthodontics*, 20, 1-10.
- Kök, H., İzgi, M. S., & Acilar, A. M., (2021). Determination of growth and development periods in orthodontics with artificial neural network. *Orthodontics & craniofacial research*, 24, 76-83.
- Kök, H., İzgi, M. S., & Acılar, A. M., (2020). Evaluation of the artificial neural network and Naive Bayes models trained with vertebra ratios for growth and development determination. *Turkish Journal of Orthodontics*, 34(1), 2.
- Kucukkeles, N., Acar, A., Biren, S., & Arun, T., (1999). Comparisons between cervical vertebrae and hand-wrist maturation for the assessment of skeletal maturity. *The Journal of clinical pediatric dentistry*, 24(1), 47-52.
- Lamparski, D., (1972). Skeletal age assessment utilizing cervical vertebrae [Master of dental science thesis]. *Pittsburgh: University of Pittsburgh, School of Dental Medicine*.

- Lamparski, D. G., & Nanda, S. K., (2002). Skeletal age assessment utilizing cervical vertebrae. *Craniofacial growth series*, 39, 171-184.
- Lanham, M., (2021). *Generating a new reality: From autoencoders and adversarial networks to deepfakes*. Springer.
- Mağat, G., (2024). Evaluation of The Relationship Between Dental and Cervical Vertebrae Maturation in Terms of Age Determination. *Necmettin Erbakan Üniversitesi Diş Hekimliği Dergisi*, 6(1), 24-33.
- Mahesh, B., (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, [Internet], 9(1), 381-386.
- Mito, T., Sato, K., & Mitani, H., (2002). Cervical vertebral bone age in girls. *American Journal of Orthodontics and Dentofacial Orthopedics*, 122(4), 380-385.
- Nur, B., Kusgoz, A., Bayram, M., Celikoglu, M., Nur, M., Kayipmaz, S., & Yildirim, S., (2012). Validity of Demirjian and Nolla methods for dental age estimation for Northeastern Turkish children aged 5–16 years old. *Medicina oral, patologia oral y cirugia bucal*, 17(5), e871.
- O'Reilly, M. T., & Yanniello, G. J., (1988). Mandibular Growth Changes and Maturation of Cervical Vertebrae: —A Longitudinal Cephalometric Study. *The Angle Orthodontist*, 58(2), 179-184.
- Oflaz, M. U., & Görgülü, S., (2017). Ortodontide Büyüme ve Gelişimin Değerlendirilmesi ve Önemi. *ADO Klinik Bilimler Dergisi*, 8(2), 1571-1580.
- Opitz, J., (2024). A Closer Look at Classification Evaluation Metrics and a Critical Reflection of Common Evaluation Practice. *Transactions of the Association for Computational Linguistics*, 12, 820-836.
- Özer, T., Kama, J. D., & Özer, S. Y., (2006). A practical method for determining pubertal growth spurt. *American Journal of Orthodontics and Dentofacial Orthopedics*, 130(2), 131. e131-131. e136.

- Öztemel, E., (2003). Yapay sinir ağları. *PapatyaYayincilik, Istanbul*.
- Padalino, S., Sfondrini, M. F., Chenuil, L., Scudeller, L., & Gandini, P., (2014). Reliability of skeletal maturity analysis using the cervical vertebrae maturation method on dedicated software. *International Orthodontics, 12(4)*, 483-493.
- Ranstam, J., & Cook, J. A., (2018). LASSO regression. *Journal of British Surgery, 105(10)*, 1348-1348.
- Rezende, D. J., Mohamed, S., & Wierstra, D., (2014). Stochastic backpropagation and approximate inference in deep generative models. International conference on machine learning,
- Robbins, J. W., (1991). Color characterization of porcelain veneers. *Quintessence international, 22(11)*.
- Román, P. S., Palma, J. C., Oteo, M. D., & Nevado, E., (2002). Skeletal maturation determined by cervical vertebrae development. *The European Journal of Orthodontics, 24(3)*, 303-311.
- Sierra, A. M., (1987). Assessment of dental and skeletal maturity: a new approach. *The Angle Orthodontist, 57(3)*, 194-208.
- Singer, J., (1980). Physiologic timing of orthodontic treatment. *The Angle Orthodontist, 50(4)*, 322-333.
- Tanner, J., Oshman, D., Bahhage, F., & Healy, M., (1997). Tanner-Whitehouse bone age reference values for North American children. *The Journal of pediatrics, 131(1 Pt 1)*, 34-40.
- Ülgen, M., (2010). Anomaliler, Sefalometri, Etioloji, Büyüme ve Gelişim, Tanı İçinde: Yüz İskeletinin Büyüme ve Gelişimi. *Ankara Üniversitesi Diş Hekimliği Yayınları. Yayın(30)*.
- Ünver, T., (2013). *El-Bilek Kemikleri, Servikal Vertebra Ve Dişlerin*

Gelişimdönemlerine Göre Kronolojik Yaş Evrelerinin El-Bilek Grafileeri Ve Dental Volumetrik Tomografi Ile Karşılaştırmalı Olarak Değerlendirilmesi Marmara Üniversitesi (Turkey)].

White, S. C., & Pharoah, M. J., (2013). *Oral radiology: principles and interpretation*. Elsevier Health Sciences.

Yeniden Parametrelendirme Hilesi Uygulamasının Yapıyı Nasıl Dönüştürdüğünün

Gösterimi. (2023, Ağustos 21). 17 Aralık 2024 tarihinde <https://wikidocs.net/214485> adresinden alındı.

Zhou, Z.-H., (2021). *Machine learning*. Springer nature.



ÖZGEÇMİŞ

Adı ve Soyadı : Fazilet TÜRKOĞLU

Öğrenim Durumu : Yüksek Lisans

Derece	Bölüm/Program	Üniversite	Yıl
Lisans	Bilgisayar Mühendisliği	Fırat Üniversitesi	2018-2022

ESERLER (Makaleler ve Bildiriler)

1. TÜRKOĞLU, F., YÜCELBAŞ, Ş., “The Role of Data Augmentation Methods in Cervical Vertebra Maturation Stage Detection: A Review”, TOKYO 9TH INTERNATIONAL INNOVATIVE STUDIES & CONTEMPORARY SCIENTIFIC RESEARCH CONGRESS, 10-11 October 2024