

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**YAPAY ÖĞRENME YÖNTEMLERİ İLE AKCİĞERDEKİ TÜMÖRLERİN
BİRİNCİL KÖKENLERİNE GÖRE SINIFLANDIRILMASI**

YÜKSEK LİSANS TEZİ

Abdullah Enes TAHA

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Yüksek Lisans Programı

EKİM 2020

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**YAPAY ÖĞRENME YÖNTEMLERİ İLE AKCİĞERDEKİ TÜMÖRLERİN
BİRİNCİL KÖKENLERİNE GÖRE SINIFLANDIRILMASI**

YÜKSEK LİSANS TEZİ

**Abdullah Enes TAHA
(507171101)**

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Yüksek Lisans Programı

Tez Danışmanı: Dr. Öğr. Üyesi Ömer Faruk BEYCA

EKİM 2020

İTÜ, Fen Bilimleri Enstitüsü'nün 507171101 numaralı Yüksek Lisans Öğrencisi Abdullah Enes TAHA, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “YAPAY ÖĞRENME YÖNTEMLERİ İLE AKCİĞERDEKİ TÜMÖRLERİN BİRİNCİL KÖKENLERİNE GÖRE SINIFLANDIRILMASI” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Dr. Öğr. Üyesi Ömer Faruk BEYCA**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Nizamettin Bayyurt**
İstanbul Teknik Üniversitesi

Prof. Dr. Selim Zaim
İstanbul Sabahattin Zaim Üniversitesi

Teslim Tarihi : 19 Ekim 2020
Savunma Tarihi : 19 Ekim 2020





Kanserden vefat eden dayıma,



ÖNSÖZ

Takip edilemeyecek bir hızla gelişen teknolojinin olduğu şu zamanlarda bile kanser hastalığına karşı hala kesin bir çözüm bulunabilmiş değil. Geç tanı konulmuş vakaların çoğu ne yazık ki acı bir son ile bitiyor. Erken teşhisin son derece önemli olduğu kanser tedavilerinde, yapmış olduğum bu çalışmanın az da olsa tanı koyarken yaşanan zorluklara çözüm getirmesini temenni ediyorum.

Tezim sırasında bana yardımlarını esirgemeyen başta ailem olmak üzere desteklerini her zaman hissettiğim aileme ve arkadaşlarıma, tecrübelerinden devamlı faydalanmamı sağlayan büyüklerime ve bana birlikte çalışma imkanı tanıyan başta Dr. Abdusselim Adil Peker olmak üzere tüm Bezmialem Hastanesi Radyoloji Bölümü ekibine teşekkür ediyorum.

Kanser sebebiyle Ağustos 2019'da vefat eden dayımı rahmetle yad ediyorum.

Ekim 2020

Abdullah Enes Taha
(Endüstri Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xiii
ŞEKİL LİSTESİ.....	xiiiv
ÖZET.....	xvii
SUMMARY	xix
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	5
3. AKCİĞER KANSERİ	9
3.1 Akciğerler.....	9
3.2 Kanser.....	10
3.3 Akciğer Kanseri.....	11
3.4 Metastaz	11
4. MATERYAL METHOD	13
4.1 Yapay Öğrenme.....	13
4.1.1 Denetimli öğrenme	13
4.1.2 Denetimsiz öğrenme.....	15
4.2 Regresyon Analizi	15
4.3 Çok Değişkenli Regresyon Analizi	16
4.4 Lojistik Regresyon	16
4.5 k-En yakın komşular	18
4.6 Destek vektör makineleri	20
4.7 Karar ağaçları ve Rastgele Orman.....	22
4.8 Gradyan Artırma	25
4.9 Yapay Sinir Ağları	27
4.10 Özellik seçimi.....	30
4.11 Doku analizi ve Haar Benzeri Özellikler	30
4.11.1 Gri seviyeli eş-oluşum matrisleri	31
4.11.2 Haar benzeri özellikler	32
4.12 Değerlendirme Metrikleri.....	33
4.12.1 Doğruluk oranı	34
4.12.2 Kesinlik değeri	34
4.12.3 Duyarlılık değeri	34
4.12.4 F1 skoru.....	34
5. VAKA ANALİZİ.....	37
5.1 Veri Seti.....	37
5.2 Veri Ön İşleme	38
5.3 Bulgular	39
5.3.1 Lojistik regresyon.....	39

5.3.2 KNN	39
5.3.3 Destek vektör makineleri.....	40
5.3.4 Rastgele orman	40
5.3.5 Gradyan artırma.....	41
5.3.6 Yapay sinir ağıları	41
6. SONUÇ VE ÖNERİLER.....	43
KAYNAKLAR.....	45
EKLER.....	51
ÖZGEÇMİŞ.....	84



KISALTMALAR

KNN	: k-En Yakın Komşu
DVM	: Destek Vektör Makineleri
KA	: Karar Ağacı
RO	: Rastgele Orman
BT	: Bilgisayarlı Tomografi
LR	: Lojistik Regresyon
YSA	: Yapay Sinir Ağları
GA	: Gradyan Arttırma
GSEM	: Gri Seviyeli Eş-oluşum Matrisi fMRG
fMRG	: Fonksiyonel Manyetik Rezonans Görüntüleme



ÇİZELGE LİSTESİ

Sayfa

Çizelge 4.1 : Çekirdek tipleri ve fonksiyonları.....	22
Çizelge 5.1 : Resimden çıkarılan özellikler.....	38
Çizelge 5.2 : Lojistik regresyon metrikleri.....	39
Çizelge 5.3 : KNN yöntemi parametreleri ve metrikleri.....	39
Çizelge 5.4 : DVM yöntemi parametreleri ve metrikleri.....	40
Çizelge 5.5 : RO parametreleri ve hesaplanan metrikleri.....	41
Çizelge 5.6 : Gradyan artırma parametreleri ve metrikleri.....	41
Çizelge 5.7 : YSA parametreler ve metrikleri.....	42
Çizelge 6.1 : Yöntemlerin değerlendirme metrikleri.....	43



ŞEKİL LİSTESİ

Sayfa

Şekil 3.1 : Akciğerin anatomik yapısı.....	9
Şekil 3.2 : Metastazın evreleri	12
Şekil 4.1 : Yapay öğrenmenin genel süreçleri.....	13
Şekil 4.2 : Denetimli öğrenme süreçleri	14
Şekil 4.3 : Lineer regresyon ve lojistik regresyon farkının grafikte gösterimi.....	17
Şekil 4.4 : En yakın komşu çalışma prensibi.....	19
Şekil 4.5 : Destek vektör makineleri gösterimi.	21
Şekil 4.6 : Karar ağacı diyagramı ve örnek düzlemde gösterimi.....	24
Şekil 4.7 : Rastgele orman yöntemi.....	25
Şekil 4.8 : Gradyan arttırma modeli	26
Şekil 4.9 : Yapay sinir ağı yapı.....	27
Şekil 4.10 : Sinir hücresinin kısımları... ..	29
Şekil 4.11 : Yapay sinir ağı kısımları... ..	29
Şekil 4.12 : Örnek 4x4'lük resim piksel değerleri.....	31
Şekil 4.13 : Referans-komşu ilişki tekrarı matrisi... ..	31
Şekil 4.14 : Transpozisi ile toplanan yeni matris.....	32
Şekil 4.15 : Hata matrisi ve elemanları	33
Şekil 5.1 : Segmente edilmemiş ve segmente edilirken akciğer BT görüntüsü.....	37
Şekil 5.2 : Segmente edilmiş tümörler	37



YAPAY ÖĞRENME YÖNTEMLERİ İLE AKCİĞERDEKİ TÜMÖRLERİN BİRİNCİL KÖKENLERİNE GÖRE SINIFLANDIRILMASI

ÖZET

Dünyadaki kansere bağlı ölümlerde akciğer kanseri birinci sırada yer alıyor. Ayrıca erkekler arasında en yaygın görülen kanser türüdür. Hesaplamalara göre 2018 yılında dünya genelinde 154,050 kişi akciğer kanseri sebebiyle yaşamını yitirmiştir. Akciğerler, toraks dışı kötü huylu hücreler için en sık görülen ikinci metastatik büyüme bölgesidir ve vakaların %20-54'ünde ikincil akciğer lezyonları saptanır. Bilgisayarlı tomografisinde tespit edilen çoklu nodüllerin %73'ü akciğer metastazıdır. Tümör hücresinin belirli bir kitlesel büyüklüğe ulaşması ve dönüşümünü tamamlaması metastazın gerçekleşmesi için gereklidir. Büyümekte olan tümör hücreleri gerekli oksijen ve besini sağlamak için besin kaynaklarına doğru ilerlemesi gerekmektedir. Bu sebeple tümör kitlesi 1-2 mm çapa sahip olduktan sonra kılcal damar oluşumu gözlemlenir. Bu kılcal damar ağı zaman geçtikçe yoğunlaşır ve tümör kitlesinin etrafını sarar. Daha sonra oluşan bu yeni damarlar sayesinde tümör hücreleri dolaşıma katılarak kan yolu ile uzak organlara doğru ilerleyerek farklı organlara tutulur. Dolayısıyla bu süreç haftaları hatta ayları bulabilmektedir. Günümüzde yapay öğrenme, özellikle teşhis sorunlarında tıbbi problemleri çözmek için çok sayıda araç sağlamaktadır. Yapay öğrenme teknikleri kanser tespitinde yaklaşık otuz yıldır kullanılmaktadır.

Bu tez çalışması Bezmialem Vakıf Üniversitesi Hastanesi'nden alınan veriler ile yapılmıştır. Henüz konu seçilmeden önce Bezmialem Vakıf Üniversitesi Radyoloji Bölümü ile görüşmeler yapıldı. Daha sonra radyoloji ve onkoloji bölümleriyle birlikte bir toplantı yapılarak uygulamada fayda sağlayacak hangi çalışmaların yapılabileceği tartışıldı. Sonuç olarak akciğerinde tümör bulunan hastaların bilgisayarlı tomografi görüntülerinden o tümörün metastaz olup olmadığının tespit edilmesinin uygun olacağı belirlendi. Mevcut uygulamada kötü huylu tümör tanısı konulmuş hastalarda kemoterapi ve radyoterapi tedavilerine başlayabilmek için öncelikle tümörün birincil kökenini belirlenmesi gerekmektedir. Biyopsi gerektiren bu işlem her zaman kolay olmayabiliyor ve hastadan farklı yerlerinden defalarca parça alınabiliyor. Bu süreç haftaları hatta ayları bulabiliyor. Kanser tedavilerinde erken tanının önemi düşünüldüğünde birkaç hafta önem arz etmektedir. Bu tezdeki amacımız, akciğerinde tümör tespit edilen hastaların tomografi görüntülerinden kanserin kökeninin akciğer mi yoksa farklı bir organ mı olduğunu tespit edebilmek.

Çalışmada toplam 100 hastanın bilgisayarlı tomografi görüntüleri kullanıldı. Birincil kökeni akciğer olan 50 hasta, akciğere metastaz yapmış 50 hastanın akciğerlerindeki tümörlerin her birisi ayrı bir gözlem olarak kabul edilmiştir. Sonuç olarak 80 akciğer, 147 metastaz tümörü uzman radyologlar tarafından segmente edildi. Segmente edilen tümör görüntüleri önce normalize edildi, daha sonra gri seviyeli eş-oluşum matrisi çıkarılarak Haar benzeri özellikler elde edildi. Elde edilen özelliklerin değerleri incelendi ve aykırı değer olan bir gözlem çıkarıldı. Birincil kökeni akciğer olan

tümörler 1, metastaz olan tümörler 0 olarak etiketlendi. Toplam 226 gözlem %70 eğitim, %30 test verisi olmak üzere rastgele olarak ikiye ayrıldı. Hesaplanan toplam 26 özellik lojistik regresyon, k-en yakın komşu, destek vektör makinesi, rastgele orman, gradyan arttırma ve yapay sinir ağları olmak üzere 6 farklı yapay öğrenme yönteminde eğitildi ve test edildi. Başlıca parametreler sürekli değiştirilerek tekrar eğitildi ve test edildi. Modellerin her parametre grubu için yaptığı tahminlerin performansı doğruluk oranı, kesinlik değeri, duyarlılık değeri ve F1 skoru olmak üzere 4 farklı değerlendirme metriğiyle hesaplandı. Her yöntem önce kendi içinde karşılaştırıldı. Eğitim verileriyle yapılan tahminler arasında da doğruluk oranı hesaplandı ve aşırı öğrenme tespit edilen modeller değerlendirilmedi. Denenen farklı parametre gruplarından en yüksek F1 skoru alan parametre grubu en iyi olarak seçildi. Lojistik regresyonun en iyi modelinin F1 skoru 0,67, k-en yakın komşunun 0,76, destek vektör makinesinin 0,77, rastgele ormanın 0,68, gradyan arttırmanın 0,74 ve yapay sinir ağlarının 0,83 olarak hesaplandı. Yapay sinir ağları 6 yöntem arasında en iyi sonucu veren yöntem olarak seçildi. Yapılan çalışmada gözlem sayısındaki kısıt modelin daha iyi kurulmasında engele sebep olmaktadır. Az veri olmasının en büyük sebebi segmentasyon işleminin uzun sürmesi ve uzmanlar tarafından yapılabilmesi. Ayrıca çekilen tomografi görüntülerinde bir standart olursa segmente edilebilecek örnek sayısı artacaktır.

CLASSIFICATION OF TUMORS IN THE LUNG ACCORDING TO PRIMARY ORIGINS USING MACHINE LEARNING METHODS

SUMMARY

Lung cancer ranks first in cancer-related deaths in the world. It is also the most common type of cancer among men. According to calculations, 154,050 people worldwide died from lung cancer in 2018. Lungs are the second most common metastatic growth site for non-thoracic malignant cells and secondary lung lesions are detected in 20-54% of cases.

When a single nodule is detected in the clinical setting in a patient with no history of malignant nodules, the probability of metastasis of this nodule is between 0.4% and 9%. If it is known that the patient has a malignant tumor, the probability of metastasis of a nodule in the lung is 25%. 73% of multiple nodules that are detected in computed tomography are lung metastases.

It is necessary for the tumor cell to reach a certain mass size and complete its transformation for metastasis to occur. Growing tumor cells must move towards food sources to provide the necessary oxygen and nutrients. For this reason, capillary formation is observed after the tumor mass has a diameter of 1-2 mm. This capillary network gets denser as time passes and surrounds the tumor mass. Via these new vessels that are formed later, the tumor cells join the circulation and move towards distant organs by blood and cling to different organs.

Machine learning is a scientific discipline that focuses on how computers learn from data. It occurs with the intersections of statistics aiming to learn relationships from data and computational algorithms are created with computer science. It is very difficult to create a statistical structure from big data that contains billions or even trillions of data. For this reason, two different disciplines as mathematics and computer science have come together. The types of learning used by computers are divided into categories such as supervised learning and unsupervised learning.

Pattern classification has been detecting abnormalities such as mammography masses and nodules on chest radiology images based on features that define the regional image appearance for decades. With advances in computer hardware, it has become possible to train more complex models on more data. In the last few years, the processes of segmenting, recognizing and saving images using supervised learning methods have been accelerated. These methods are generally more accurate than traditional quantitative analysis based on simple cross and density measurements. The main obstacle to the fact that artificial learning methods cannot be used more widely in medical imaging today is the insufficient amount of educational data.

Today, machine learning provides a large number of tools to solve medical problems, especially in diagnostic problems. There is a relationship between radiomics and machine learning based on mutual benefit. Due to its ever-growing high-dimensional nature, radiomics requires stronger analysis tools, and machine learning is a potential candidate for this purpose. On the other hand, machine learning studies in medical image analysis need radiomics for feature extraction. Machine learning techniques have been used in cancer detection for nearly thirty years. According to PubMed's results on machine learning and cancer, more than 7510 articles have been published so far. The vast majority of these publications make use of one or more machine learning algorithms and combine data from heterogeneous sources to predict tumors as well as a cancer type.

This thesis study was conducted with the data obtained from Bezmialem Vakıf University Hospital. Before the topic was chosen, meetings were held with the Radiology Department of Bezmialem Vakıf University. Then, a meeting was held with the radiology and oncology departments to discuss which practices could be useful in this study. As a result, it was determined with the consensus that it would be appropriate to determine whether the tumor had or had not metastasis from the computed tomography images of patients with lung tumors.

In current practice, the primary origin of the tumor should be determined first in order to start chemotherapy and radiotherapy treatments in patients diagnosed with malignant tumors. This procedure, which requires biopsy, is not always easy. It may be necessary to take samples from different parts of the patient's body many times. This process might take weeks or even months. Considering the importance of early diagnosis in cancer treatments, each week is important. Our aim in this thesis is to determine whether the origin of the cancer is the lung or a different organ from the CT images of patients with lung tumors.

Computed tomography images of 100 patients were used in the study. Each of the tumors in the lungs of 50 patients whose primary origin was lung and 50 patients who had metastasized to the lung were considered as a separate observation. As a result, 80 lungs and 147 metastasis tumors were segmented by specialist radiologists. First, segmented tumor images were normalized, then the gray level co-formation matrix was formed to obtain Haar-like features. The values of the obtained properties were examined and an outlier observation was removed. Tumors whose primary origin was lung were labeled as 1, tumors with metastasis as 0. Total of 226 observations were randomly splitted into two as 70% training and 30% test data. The calculated 26 properties were trained and tested in six different machine learning methods: logistic regression, k-nearest neighbor, support vector machine, random forest, gradient boosting and artificial neural networks.

Major parameters have been constantly changed and retrained and tested. The performance of the estimations made by the models for each parameter group was calculated with 4 different evaluation metrics: accuracy rate, precision value, recall value and F1 score. Each method was first compared within itself. The accuracy rate between the predictions and the training data was also calculated and the models with excessive learning were not evaluated. The parameter group with the highest F1 score in the tested parameter groups was chosen as the best. The F1 score of the best model of logistic regression was 0.67, k-nearest neighbor 0.76, support vector machine 0.77, random forest 0.68, gradient enhancement 0.74, and artificial neural networks 0.83. Artificial neural networks were chosen as the best method out of six methods.

In the study, the scantiness in the number of observations causes an obstacle to establishing a better model. The foremost reason for the lack of data is that the segmentation process takes a long time and can only be done by experts. In addition, if there would be a standard in the tomography images taken, the number of samples that can be segmented will increase meaningfully.





1. GİRİŞ VE AMAÇ

Dünyadaki kansere bağlı ölümlerde akciğer kanseri birinci sırada yer alıyor. Ayrıca erkekler arasında en yaygın görülen kanser türüdür. Hesaplamalara göre 2018 yılında dünya genelinde 154,050 kişi akciğer kanseri sebebiyle yaşamını yitirmiştir. (Mukherjee vd., 2019). Akciğerler, toraks dışı kötü huylu hücreler için en sık görülen ikinci metastatik büyüme bölgesidir ve vakaların %20-54'ünde ikincil akciğer lezyonları saptanır. Akciğer metastazı çoğunlukla akciğer atar damarı dolaşımında yuva yapan tümör hücrelerinin damar içi yayılımı ile ilişkilidir. Bu yayılma modeli nedeniyle, akciğer metastazları bazal ve periferik bölgelerde daha bol olma eğilimindedir (Stella ve diğ, 2019).

Herhangi bir kötü huylu nodül geçmişi olmayan hastada klinik ortamda tekli bir nodül tespit edildiğinde bu nodülün metastaz olma ihtimali %0,4 ile %9 arasındadır. Eğer hastada kötü huylu tümör olduğu biliniyorsa akciğerde bulunan bir nodülün metastaz olma olasılığı %25'tir. Bilgisayarlı tomografisinde tespit edilen çoklu nodüllerin %73'ü akciğer metastazıdır (Hirikata ve diğ, 1995).

Örüntü sınıflandırması on yıllardır, mamografideki kitleler ve göğüs radyoloji görüntülerindeki nodüller gibi anormallikleri bölgesel görüntü görünümünü tanımlayan özellikleri temel alarak tespit etmektedir. Bilgisayar donanımındaki gelişmelerle birlikte daha fazla veri üzerinde daha karmaşık modeller eğitmek mümkün hale geldi. Son birkaç yıldır denetimli öğrenme yöntemleri kullanılarak görüntüleri segmente etme, tanıma ve kaydetme işlemleri hızlanmıştır. Bu yöntemler segmentasyon sistemlerinde basit yoğunluk ve gradyan yöntemlerin yerini almaktadır. Birkaç yeni yöntem verilerde çok değişkenli sınıflandırma veya regresyon yöntemlerini kullanarak resimden tanı koymayı öğrenmektedir. Bu yöntemler, genellikle basit hacim ve yoğunluk ölçümlerine dayalı geleneksel nicel analizlerden daha yüksek doğruluğa sahiptir. Günümüzde tıbbi görüntülemede yapay öğrenme yöntemlerinin daha yaygın olarak kullanılamamasındaki başlıca engel eğitim verilerinin yeterli miktarda olmamasıdır. Ayrıca denetimli öğrenme tekniklerinin eğitimleri sırasında standartlaştırılmış görüntüleme protokolleri ile elde edilen

resimler kullanıldığında umut vaat eden sonuçlar verirken standardın dışında çekilmiş bir resmi test ettiğinde performansları hızlı şekilde bozulabilmektedir. Bunun sebebi, bu yöntemler hem eğitim hem de test veri kümelerinin aynı veri kümesinden alınan rastgele gözlemler olduğu varsayımı altında çalışmasıdır. Ancak uygulamada eğitim verileri, farklı cihazların farklı görüntüleme protokollerinden, farklı yöntemlerden ve farklı bir hasta popülasyonundan elde edilmiş görüntüleri kullanır ve bu varsayımı ihlal eder. Giderek artan bu sorunları önlemenin yolu transfer öğrenmesi veya alan uyarlaması teknikleridir (Brujine, 2016).

Radyomik analizi, tıbbi görüntülerden çok sayıda özellik çıkarılmasıdır. Radyomik analizi, piksel veya voksel dağılımı ve görüntüleri hakkında yararlı uzamsal bilgilere sahip şekil, boyut ve doku özelliklerini içerir. Bu radyomik özellikler çeşitli organlarda tanı ve tedaviye destek amacıyla istatistiksel modeller oluşturmada kullanılır. Günümüzde yapay öğrenme, özellikle teşhis sorunlarında tıbbi problemleri çözmek için çok sayıda araç sağlamaktadır. Radyomik ve yapay öğrenme arasında karşılıklı fayda ilişkisi vardır. Sürekli büyüyen yüksek boyutlu doğası nedeniyle radyomik daha güçlü analiz yapabilen araçlara ihtiyaç duyar ve yapay öğrenme bu amaç için potansiyel bir adaydır. Diğer bir yandan tıbbi görüntü analizinde yapay öğrenme çalışmaları özellik çıkarımı konusunda radyomiğe ihtiyaç duyar (Koçak ve diğ, 2019).

Bu tez çalışması Bezmialem Vakıf Üniversitesi Hastanesi'nden alınan veriler ile yapılmıştır. Henüz konu seçilmeden önce Bezmialem Vakıf Üniversitesi Radyoloji Bölümü ile görüşmeler yapıldı. Daha sonra radyoloji ve onkoloji bölümleriyle birlikte bir toplantı yapılarak uygulamada fayda sağlayacak hangi çalışmaların yapılabileceği tartışıldı. Sonuç olarak akciğerinde tümör bulunan hastaların bilgisayarlı tomografi görüntülerinden o tümörün metastaz olup olmadığının tespit edilmesinin uygun olacağı ortak kanaat ile belirlendi. Mevcut uygulamada kötü huylu tümör tanısı konulmuş hastalarda kemoterapi ve radyoterapi tedavilerine başlayabilmek için öncelikle tümörün birincil kökenini belirlenmesi gerekmektedir. Biyopsi gerektiren bu işlem her zaman kolay olmayabiliyor ve hastadan farklı yerlerinden defalarca parça alınabiliyor. Bu süreç haftaları hatta ayları bulabiliyor. Kanser tedavilerinde erken tanının önemi düşünüldüğünde birkaç hafta önem arz etmektedir.

Bu tezdeki amacımız, akciğerinde tümör tespit edilen hastaların tomografi görüntülerinden kanserin kökeninin akciğer mi yoksa farklı bir organ mı olduğunu tespit edebilmek. Biyopsi gibi uzun, hasta için yorucu bir sürece gerek olmaksızın

böyle bir teşhisin yapılmasının kanser tedavisinde önemli bir gelişme olacağına inanıyoruz. Araştırmalarımız sonucunda daha önce tam olarak böyle bir çalışmanın yapılmadığı sonucuna vardık. Bu tezde yapılan çalışmaların ilerleyen süreçte gerçek uygulamada tanı ve tedavide yardımcı modeller oluşturulmasında katkısının olmasını ümit ediyoruz.

Çalışmamızda segmente edilmiş tümörden doku analizi yöntemleriyle çeşitli özellikler elde edilmiştir. Bu özellikler kullanılarak 6 farklı yapay öğrenme yöntemi eğitilmiş ve test edilmiştir. Çıkan sonuçlar karşılaştırılmış ve en iyi model seçilmiştir.





2. LİTERATÜR ARAŞTIRMASI

Bir hastalığın seyrinin tahminindeki başarısı şüphesiz tıbbi bir teşhisin kalitesine bağlıdır. Fakat prognostik bir tahmin basit bir teşhis kararından daha fazla dikkate alınmalıdır. Kansere duyarlılık tahmini, kanserin nüksetmesi ve kanserli bir şekilde hastanın hayatta kalmasının öngörülmesi. İlk iki vakada, bir tür kanser geliştirme olasılığını ve tamamen veya kısmi bir remisyondan sonra bir kanser tipini yeniden geliştirme olasılığını bulmaya çalışıyor. Son durumda, kanser teşhisi veya tedavisi sonrası hastalıklara özgü veya genel hayatta kalma gibi bir hayatta kalma sonucunun öngörülmesi ana amaçtır. Kanser sonucunun öngörülmesi genellikle yaşam beklentisi, hayatta kalma, ilerleme ve tedavi duyarlılığı durumlarını ifade eder.

Yapay öğrenme teknikleri kanser tespitinde yaklaşık otuz yıldır kullanılmaktadır. PubMed'in yapay öğrenme ve kanser konusundaki sonuçlarına göre şu ana kadar 7510'dan fazla makale yayınlandı. Bu yayınların büyük çoğunluğu, bir veya daha fazla yapay öğrenme algoritmasından yararlanır ve tümörlerin yanı sıra bir kanser tipinin öngörülmesi için heterojen kaynaklardan gelen verileri birleştirir (Kourou ve diğ, 2014).

Yapay öğreniminin radyolojide bir başka uygulaması da radyoloji metin raporlarının işlenmesidir. Günlük radyoloji pratiğinden toplanan raporlar büyük metin veri tabanlarını doldurmaktadır. Bu radyoloji rapor veri tabanlarını modern bilgi işlem teknolojilerini kullanarak kullanmak, rapor aramayı ve almayı iyileştirebilir ve radyologların tanı koymasına yardımcı olabilir. Anahtar kelimeleri kullanarak yapılan arama raporlarıyla karşılaştırıldığında doğal dil işleme ve doğal dil anlaması, radyoloji raporlarında gizli olan ilgili bilgileri organize etmek ve almak için daha etkili bir yol sağlar. Bu yöntemlerin avantajı, büyük ölçekli verileri işleyebilmeleri ve anlamlı bir bilgiyi insan okuyucular tarafından mümkün olmayacak şekilde çıkartabilmeleridir. Doğal dil işleme ve doğal dil anlaması, yapay öğreniminin önemli altkümeleridir ve radyoloji rapor veri tabanlarının metin analizi için uygun bir yol sunar. İnsan dilinden faydalı bilgiler çıkarabilir ve bunları ayrıştırma ağaçları veya birinci dereceden mantık

gibi daha resmi temsiller halinde düzenleyebilir. Bu resmi gösterimler, insan dilini doğrudan kullanma ile karşılaştırıldığında daha verimli bilgisayar işleme olanağı sağlar. Kısaca, yapay öğrenme radyoloji alanında kullanıldığında birçok zaman ve emekten tasarruf edilecek ve daha tutarlı sonuçlar elde edilecektir. (Wang ve Summers, 2012).

Beyin fonksiyonu ve aktivite analizi biliş, psikoloji ve beyin hastalığı teşhisinde araştırmalarda önemli roller oynar. Fonksiyonel manyetik rezonans görüntüleme (fMRG), beyin aktivitesini değerlendirmek için müdahalesiz ve etkili bir yol sağlar. İnsan beyninin karmaşıklığı ve beyin aktivitesindeki farklılıklar nedeniyle, fMR görüntüleri genellikle karmaşık gösterir, örüntüler ve yorumlanması genellikle bilgisayar analizi gerektirir. Son yıllarda yapay öğrenme algoritmaları, fMRG'den uyarıcıların, zihinsel durumların, davranışların ve diğer ilgili değişkenlerin kodunu çözmek için gittikçe daha fazla kullanılmıştır. Özellik çıkarımından sonra, beyin aktivite modellerini ayırt etmek için destek vektör makineleri, doğrusal diskriminant analizi ve sinir ağı gibi bir sınıflandırıcı yetiştirebiliriz. FMR görüntülerinden kalıpları öğrenmek için otomatik yöntemler geliştirmek zordur. Çünkü veriler küçük alan kaplayan, oldukça yüksek boyutlu ve gürültülüdür. T. Mitchell ve arkadaşları, bilişsel durumları beyin görüntülerinden çözmeye çalışmışlardır. Bir insana bir özne resmi veya cümle gösterdiler. Ya da farklı kavramları (yemek, insanlar, bina vb.) tanımlayan bir kelime okumalarını istediler. Amaç, fMRG tarafından tespit edilecek farklı beyin aktivitesi döngülerinin aktive edilmesi idi. Gaussian naif Bayes, SVM'ler ve KNN dahil olmak üzere fMRG verilerini analiz etmek için birkaç sınıflandırıcı araştırdılar. Elde ettiği sonuçlar, beynin çeşitli bilişsel durumlarını yapay öğrenme algoritmaları kullanarak ayırt etmenin mümkün olduğunu gösterdi. Daha sonra farklı dürtüler altında beyin davranışını tahmin etmek için insan bilişi üzerine daha fazla araştırma yaptılar. Önceki araştırmalar, uzaysal sinirsel aktivasyon kalıplarının, farklı nesneler ve kavramlar hakkında düşünme ile ilişkili olduğunu göstermiştir. fMRG verileri eğitim süreci tarafından kullanılmamış sözcüklerle ilişkili sinir yapılarını tahmin edebilen bir hesaplama modeli şeklinde gösterilmiştir. Çarpıcı bir bulgu, modellerinin bir metin korpusunda binlerce isim için sadece 60 ismin fMRG verilerine dayanarak doğru tahminler yapabilmesiydi (Wang ve Summers, 2012).

Toplam 115 hastanın beyin MR görüntüsüyle sınıflandırma çalışması yapmıştır. Beyindeki kitlenin metastaz mı yoksa nekroz mu olduğu konusunda yapılan

sınıflandırma çalışmasında 83 metastaz hastası, 32 nekroz hastasının verileri kullanılmıştır. Metastaz hatalarının 23'ü radyoterapi ile tedavi olmuş hastalardan oluşmaktadır. Segmente edilmiş tümör görüntülerinden gri seviyeli eş-oluşumlu matrisi, histogram, dalgalıklar ile özellikler 155 farklı istatistiki özellik çıkarılmıştır. Daha sonra veri seti ikiye bölünmüştür. Birinci veri setinde tedavi edilmemiş metastazlar çıkarılmıştır. İkinci veri setinde eğitim setine tedavi edilmemiş metastazlar konulmuş, test setine ise tedavi edilmiş metastazlar ve nekroz gözlemleri rastgele şekilde konulmuştur. Daha sonra her iki veri seti de destek vektör makineleri yinmeli özellik eleme yöntemini kullanılarak temsil etmeyen özellikler çıkarılıyor ve sınıflandırma yapılıyor. İlk veri setinde 10, ikinci veri setinde 7 anlamlı özellik belirleniyor. İlk veri setinden 0,94, ikinci veri setinden 0,93 kesinlik değeri elde ediliyor (Larroza ve diğ, 2015).

Karaciğerdeki tümörün iyi huylu mu kötü huylu olduğunu tespit etmek için 20 kötü huylu tümöre sahip hastanın ve 10 iyi huylu tümöre sahip hastanın bilgisayarlı tomografi görüntüleri kullanılmıştır. Gri seviyeli eş-oluşumlu matristen elde edilen özellikleri kullanarak olasılıklı sinir ağı modeli oluşturulmuştur. Oluşturulan model, 20 kötü huylu hastanın 17'sini, 10 iyi huylu hastanın 8'ini doğru tahmin etmiştir (Chen ve diğ, 1998).

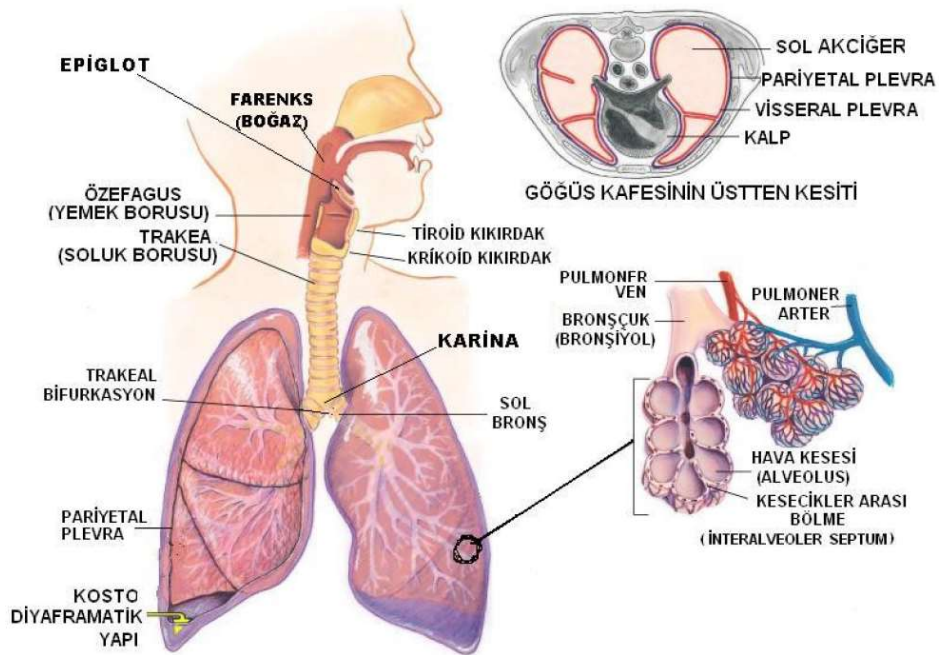
Akıllı tekerlekli sandalyedeki bir kameradan alınan görüntüler kullanarak aşağı giden merdiven, KNN kullanarak tespit edilmeye çalışıldı. Kameradan alınan 640x480 piksellik görüntülerin önce en altındaki 200x480 boyutundaki kısım kullanılmak üzere kesildi. Daha sonra bu kısım 3 kanallı RGB renk skalasından gri skalaya dönüştürüldü. Dönüştürülen resmin 0,45,90 ve 135 derecelik GSEM'leri oluşturdu ve oluşturulan GSEM'lerden haar benzeri özellikler elde edildi. Daha sonra $k=3,5,7$ değerleriyle her farklı açıdaki haar benzeri özellikler kullanılarak KNN modeli oluşturuldu. Oluşturulan modeller arasında %92,5 doğruluk oranına sahip $k=7$ ve 0 derece parametreleri en iyi parametre olarak seçildi (Johan ve diğ, 2020).



3. AKCİĞER KANSERİ

3.1 Akciğerler

Göğüs boşluğunun içerisinde bulunan akciğerler, solunum sisteminin birincil organıdır ve sağ ve sol olmak iki tanedir. Alt solunum soluk borusu ve akciğerlerden oluşmaktadır. Göğüs boşluğunun ana hava yolu soluk borusudur. Soluk borusu, gırtlığın alt ucundan başlayarak akciğerlere kadar devam eder. Akciğerlerin bulunduğu karina seviyesine geldiğinde sağ ve sol ana bronşlara bölünür. Ana bronşlar ise bölünerek sekonder bronşları onlar da bölünerek segmenter bronşları meydana getirirler. Segmenter bronşlar daha sonra mikroskobik seviyedeki bronşiolle onlar da alveollere dönüşürler. Alveoller akciğerin fonksiyonel birimidir. Bu seviyede bulunan zengin kılcal damarlar ile alveoller arasında karbondioksit ve oksijen alışverişi olur (Çevikbaş, 2016). Akciğerler trakeadan sonra 24 kez ikiye bölünerek 300 milyon alveol meydana gelir. Bu alveollerin yaklaşık toplam yaklaşık 140 metrekare yüzey alanı vardır (Coşar, 2014).



Şekil 3.1 : Akciğerin anatomik yapısı (Çevikbaş, 2016).

3.2 Kanser

Canlı dokulara gömülü olan hücreler, sürekli olarak ortamı biyokimyasal ve mekanik olarak algılar. Bu hücre dışı sinyalleri hücre içi sinyallere dönüştürür. Daha sonra bu sinyalleri birleştirerek cevap verir. Bu cevaplar gen ifadesi, farklılaşma, büyüme, hücre bölünmesi, hücre hareketi ve apoptozu içerir. Hücre kendini dışarıda gelen olumsuzluklara karşı korur. Bu koruma, doku ve çevresindeki sinyallerin doku genelinde birleştirilmesini ve verdiği cevabı içerir. Tespit ve yanıt adımları sırasında herhangi bir bozulma hücrenin anormal ve kontrolsüz büyümesine sebep olur. Bu da tümör oluşumuna neden olur. Tümörler iyi huylu veya kötü huylu karakterde olabilirler. (Kim ve diğ, 2011).

Bir hücre veya hücre grubunun organizmanın kontrol mekanizmasından çıkarak kontrolsüz çoğalmasına tümör denir. Organizmada bu kontrol mekanizması gerektiği gibi çalışmadığı zamanlar sağlıklı hücreler yapısal ve işlevsel değişim geçirerek tümör karakteri kazanabilirler (Hanahan ve Weinberg, 2011). Hücrede oluşan genetik ve/veya epigenetik değişimlerinin toplamı sonucu meydana gelen hastalığa kanser denir (Ponder, 2001).

Normal hücreler diğer hücrelere veya dokudaki hücre dışındaki alana adezyon molekülleri ile bağlanır. Bu bağlantı noktalarında elektron yoğun plakalar meydana gelir. Böylece hücrelere durma potansiyeli kazandırır. Fakat bu durum her hücrede bulunmaz. Tümör hücrelerinin durma potansiyelinin olmamasının sebebidir. Dolayısıyla tümör hücreleri bulunduğu organda kitlesel bir şekilde gelişirler. Eğer bu hücreler bulunduğu organdan çıkmazsa bu tümöre iyi huylu tümör denir. Fakat bazı tümör hücreleri bulunduğu birincil organdan ayrılarak farklı organlara doğru gelişme eğilimindedir. Bu birincil odaktan ayrılarak farklı yerlere gitme eğilimine metastaz denir ve metastaz yapabilen tümörler kötü huylu tümör olarak tanımlanmaktadır (Chiang ve Massague, 2008).

Kötü huylu tümörlerin davranışlarına bakarak altı özellik tanımlanmıştır. Bu özellikler sırasıyla, dışarıdan gelen büyüme sinyallerine ihtiyaç duymama, büyümeyi durdurmayı sağlayan sinyallere karşı direnç, programlı hücre ölümüne yakalanmama, yaşlanmaya karşı dayanıklılık ve sınırsız çoğalma potansiyeline sahip olma, yeni kan damarları oluşturabilme ve bir organdan farklı bir organa metastaz yapabilme. Bu altı

özellik kötü huylu tümörlerin ana özellikleri olup bu özelliklere göre tedavisinde bu özellikler hedef alınmaktadır (Hanahan ve Weinberg, 2000).

3.3 Akciğer Kanseri

Akciğer kanseri, akciğer dokusunun malign transformasyonu ile meydana gelen bir kanser türüdür (Kaya, 2012).

Dünyadaki kansere bağlı ölümlerde akciğer kanseri birinci sırada yer alıyor. Ayrıca erkekler arasında en yaygın görülen kanser türü. Hesaplamalara göre 2018 yılında dünya genelinde 154,050 kişi akciğer kanseri sebebiyle yaşamını yitirdi. Bu korkunç hastalığın sebeplerine bakıldığında %85'lik oran ile uzun süreli sigara içme en başta geliyor. Hastalar incelendiğinde akciğer kanserinin vücutta birçok yerde metastaz yaptığı ortaya çıkmıştır (Mukherjee ve diğ, 2019).

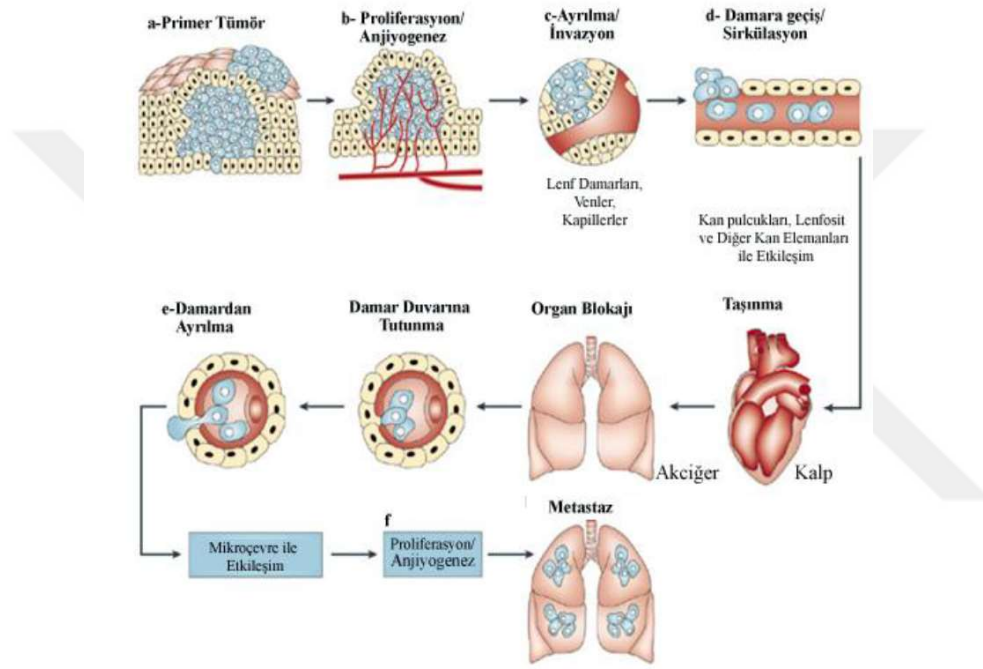
Akciğer kanserinin görülme sıklığı yaş ile doğru orantılıdır. Hastaların çoğu 50-70 yaş arasında olup hastalığın yaş ortalaması yaklaşık 60'tır. Hastalık erkeklerde daha sık görülür. Ayrıca siyah ırkta beyaz ırka göre daha sık görülür (Müsellim, 2007). Akciğer kanserinin görülme ihtimali toplam içilen sigara miktarıyla orantılıdır ve pasif içiciler de risk grubundadır (Wald ve diğ, 1986).

Akciğer kanserinde tanı ve kanserin evresini belirlemek için çeşitli yöntemler vardır. Tedavi, kanserin alt türüne ve evresine göre belirlenir. Birkaç yıl öncesine kadar olmayıp şu an olan birçok kişiye özgü tedavi yöntemi vardır (Nasim ve diğ, 2019). Akciğer kanseri tedavisinde tümörün biyolojisini anlamada, ileri düzeyde erken teşhiste, çok yöntemli tedavilerde ve ilaçlara karşı gösterilen direncin moleküler mekaniğinin çözülmesinde muazzam gelişmeler sağlanmasına rağmen özellikle metastatik hastalıklarda olumlu sonuçlar pek fazla olmuyor (Cryer ve Thorley, 2019).

3.4 Metastaz

Tümör metastazı bir takım sıralı olayların sonucunda gerçekleşmektedir. Bu olayların herhangi birindeki başarısızlık metastaz oluşumunu engelleyecektir (Poste ve Fidler, 1980). Hücre belirli bir kitlesel büyüklüğe ulaşması ve dönüşümünü tamamlaması metastazın gerçekleşmesi için gereklidir. Büyümekte olan tümör hücreleri gerekli oksijen ve besini sağlamak için besin kaynaklarına doğru ilerlemesi gerekmektedir. Bu sebeple tümör kitlesi 1-2 mm çapa sahip olduktan sonra

anjiyogenetik faktörlerin sentezlenmesi ve salgılanmasıyla kılcal damarların oluşumu gözlenir. Bu kılcal damar ağı zaman geçtikçe yoğunlaşır ve tümör kitlesinin etrafını sarar. Daha sonra oluşan bu yeni damarlar sayesinde tümör hücreleri dolaşıma katılarak kan yolu ile uzak organlara doğru ilerlerler. Kanda bulunan immün hücrelerinin tümör hücrelerine saldırma ihtimali vardır. Bu sebeple tümör hücreleri fazla kanda dolaşmadan bir an önce dokuya gitme eğilimindedir. Dolaşımdan ayrılıp yeni organlara ulaştığında tekrar büyüme eğilimindedir ve anjiyogenez olarak adlandırılan yeni damar oluşumunu tetiklerler. (Fidler,2003).



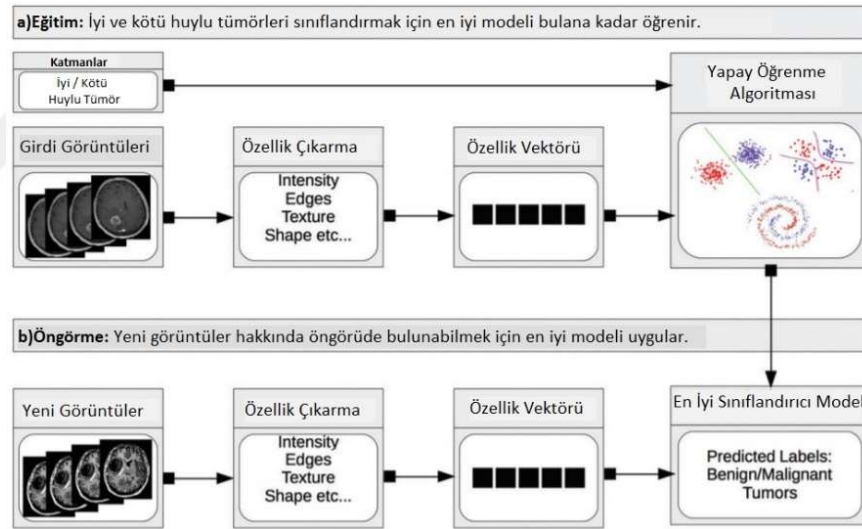
Şekil 3.2 : Metastazın evreleri (Dilmaç, 2014).

Anjiyogenez tümör dokusuna damar desteği sağlarken aynı zamanda ona taşıdığı oksijen ve besinle büyümesine yardımcı oluyor ve metastaz yapıp farklı organlara yayılmasını desteklemektedir. Tümör hücreleri farklı bir dokuya metastaz yaptığında hemen agresif bir şekilde hemen çoğalmazlar. Tek bir hücreden veya mikro metastazlardan oluşan tümör dormansi denilen uyku halinde yıllarca kalabilirler. Tümörden salınan faktörler veya mikro çevrenin değişmesiyle dormansi halinden kurtulup agresif hale geçebilirler. Agresif hale geçtikten sonra tümör hücreleri hızlıca büyüyerek yeni organlara metastaz arayışına girerler (McGee ve diğ, 2006).

4. MATERYAL METHOD

4.1 Yapay Öğrenme

Yapay öğrenme, bilgisayarların verilerden nasıl öğrendiğine odaklanan bilimsel bir disiplindir. Verilerden ilişkiler öğrenmeyi amaçlayan istatistiklerin kesişimleriyle ortaya çıkar ve bilgisayar bilimleriyle hesaplama algoritmaları oluşturulur. Milyarlarca ve hatta trilyonlarca veri içeren büyük veriden istatistiksel bir yapı oluşturulması oldukça zordur. Bu sebeple matematik ve bilgisayar bilimleri olarak iki farklı disiplin bir araya gelmiştir. Bilgisayarlar tarafından kullanılan öğrenme türleri, denetimli öğrenme ve denetimsiz öğrenme gibi kategorilere ayrılır (Deo, 2015).



Şekil 4.1 : Yapay öğrenmenin genel süreçleri (Erickson ve diğ, 2017).

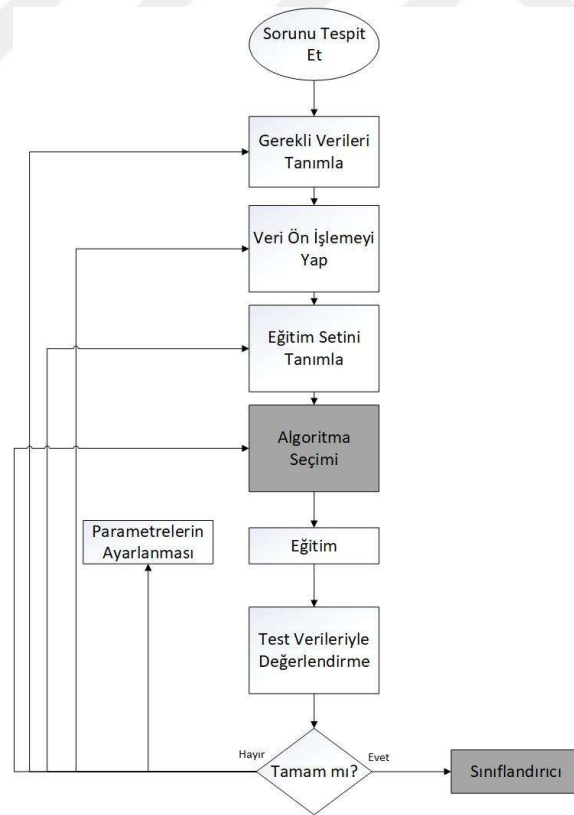
4.1.1 Denetimli öğrenme

Denetimli öğrenme, etiketli bir girdi ile istenilen çıktının tahmin edilmesinde kullanılır. El yazısı tanıma, nesnelerin görüntülerini sınıflandırma ve doküman sınıflandırması gibi denetimli öğrenme problemleri örnek olarak verilebilir. Bu problemlerin hepsi eğitilmiş bir kişinin iyi yapabileceği görevlerdir ve bu nedenle bilgisayar genellikle insan performansını yaklaştırmaya çalışılıyor. Denetimli öğrenme, yeni bir veri örneğini en iyi şekilde tanımlamak için alt gruplar arasında

seçim yapmayı ve bilinmeyen bir parametrenin tahmin edilmesini içeren öngörmeyi içerir.

Tıpta denetimli öğrenmeye örnek olarak, kardiyolojide sınırlı bir tanı dizisinden seçim yapmak için örüntü tanımanın yapıldığı EKG'nin otomatik yorumu verilebilir. Radyolojide, bir akciğer nodülünün otomatik olarak göğüs röntgeninden tespiti de denetimli öğrenmeyi temsil eder. Bu iki durumda da geliştirilen model, iyi bir doktorun zaten yüksek hassasiyetle yapabileceklerini yapıyor. Denetimli öğrenme genellikle riski tahmin etmek için kullanılır. Koroner kalp hastalığı için Framingham Risk Skoru aslında tıpta denetimli öğrenmenin en yaygın kullanılan örneği olabilir (Deo, 2015)

Denetimli öğrenmenin bir türü de sınıflandırmadır. Algoritmanın bir vektörü birkaç sınıftan birine ataması için fonksiyonun çeşitli girdi ve çıktı örneklerine bakması gerekir. Yapay öğrenme, bir eğitim setindeki örneklerden bir takım kuralları öğrenme sürecidir. Yani yeni örneklerden genelleme yapmak için kullanılacak bir sınıflandırıcı oluşturur. Gerçek dünyadaki bir soruna denetimli yapay öğrenme süreci Şekil 4.2’de verilmiştir (Osisanwo ve diğ., 2017).



Şekil 4.2 : Denetimli öğrenme süreçleri (Osisanwo ve diğ., 2017).

4.1.2 Denetimsiz öğrenme

Denetimsiz öğrenmede, sadece bir gözlem setimiz vardır ve her örnek için etiket bilgisi yoktur. Genellikle bu gözlemler veya özellikler, gözlemlenmemiş veya gizli değişkenlerden oluşur. Denetimsiz öğrenmenin temel amacı, örnekler arasındaki ilişkileri keşfetmek veya gözlemlerin arkasındaki gizli değişkenleri ortaya çıkarmaktır. Kümeleme, veri öğelerini tanımlamak için kategorileri veya kümeleri bulmaya çalışır. Denetimsiz öğrenmeye örnek olarak K-means, ortalama kaydırma, hiyerarşik kümeleme, Gauss karışım modeli ve bulanık C-means verilebilir (Deo, 2015).

4.2 Regresyon Analizi

Regresyon analizi, en yaygın kullanılan istatistiksel yöntemlerden biridir (Ruppert ve Matteson, 2015) Regresyon analizi, aralarında sebep sonuç ilişkisi olan birden fazla değişkenin arasındaki bağıntıyı araştırmak ve ilgili tahmini yapmak için matematiksel model oluşturan istatistiksel bir analizdir (Şahinler, 2000). Regresyon analizinde, bağımlı değişken olarak adlandırılan çıktıyı bağımsız değişken olarak adlandırılan girdiler ile yazmak amaçlanır (Alpaydın, 2018).

Regresyon analizi, iki veya daha fazla değişken arasındaki ilişkiyi kullanan bu istatistiksel araçlardan biridir. Regresyon modelleri doğrusal ilişki modelleri ile doğrusal olmayan ilişki modelleri olarak iki gruba ayrılabilir. Doğrusal modeller çoğu regresyon uygulaması için yeterlidir. Doğrusal modelin hiçbir şekilde uygulanamadığı yerde doğrusal olmayan model kullanılır (Habil, 2012).

Tek değişkenli regresyon analizi bağımlı değişkenin sadece bir tane bağımsız değişken ile açıklandığı analizdir.

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \dots + \beta_p X_{i,p} + \varepsilon_i \quad (4.1)$$

Çoklu doğrusal regresyon analizi denklem 4.1'de görüldüğü gibi bağımlı değişkenin birden fazla bağımsız değişkenlerle açıklandığı analizdir. Burada ε_i genellikle hata değişkeni olarak adlandırılır çünkü kurulan Y değeri tahmin edilmeye çalışırken yapılan tahmin hatasıdır. Regresyon analizlerinde hata değişkenine ait bir takım varsayımlar vardır:

- 1- Hata değişkenleri birbiri arasında bağımsızdır
- 2- Sabit varyansa sahiptir

3- Ortalaması 0 olan normal dağılıma göre dağılım gösterirler.

Bağımlı değişkenin birden fazla olması durumuna çok değişkenli regresyon analizi denilmektedir (Ruppert ve Matteson, 2015).

4.3 Çok Değişkenli Regresyon Analizi

Çok değişkenli regresyon analizinde oluşturulan modeldeki bazı bağımsız değişkenler bağımlı değişkeni açıklamada yetersiz kalabilir. Bu sebeple modeli oluştururken bağımlı değişkeni en uygun şekilde açıklayan bağımsız değişkenler seçilmeli ve önemsiz değişkenler seçilmemelidir. Bu süreç “değişken seçimi” olarak adlandırılmaktadır (Alpar, 2003).

Bir regresyon Y bağımlı değişkenini açıklayan $X_1, X_2, X_3, \dots, X_k$ bağımsız değişkenleri ile bu değişkenleri birden fazla değişken seçimi mevcuttur ve bu değişkenlerden çok sayıda model oluşturulabilir. Bağımsız değişkenlerin fazla olması modelin sapma hatasının düşük olmasını sağlarken aynı zamanda bilgi edinme maliyetlerinin artmasına sebep olmaktadır. Bu sebeple kabul edilebilir düzeyde az değişkenin seçilmesi tercih edilmelidir ve dolayısıyla en iyi regresyon modelinin seçilmesi önem arz etmektedir. En iyi modelin seçilmesinde birden fazla yöntem ve süreç vardır. Bu yöntem ve süreçler denenerek en iyi modeller belirlenir ve uygulamaya göre kabul görecektir şekilde oluşturulur (Çakır Zeytinoğlu, 2007).

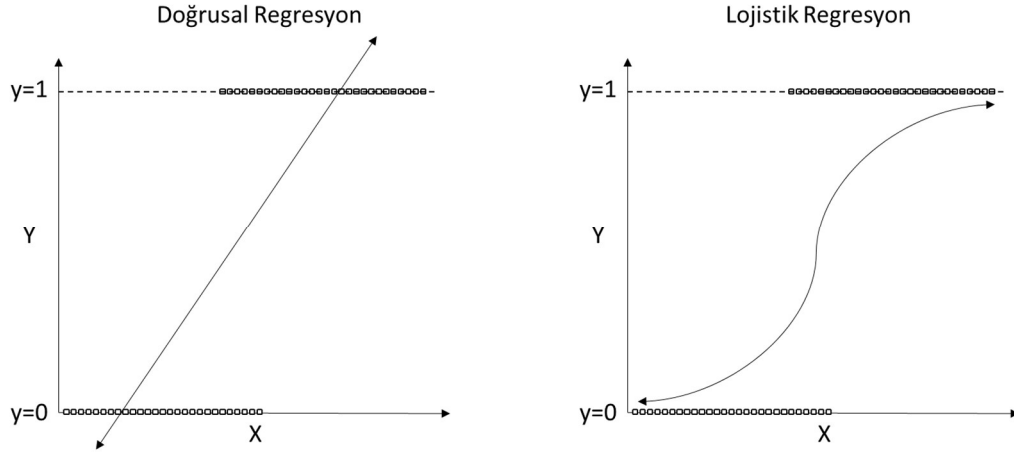
4.4 Lojistik Regresyon

Lojistik regresyon modeli ile doğrusal regresyon modeli arasındaki fark, lojistik regresyon modelinin bağımlı değişkeninin ikili olmasıdır. Lojistik ve doğrusal regresyon arasındaki bu fark hem parametrik model seçiminde hem de varsayımlarda önemli rol oynamaktadır. Bu fark hesaba katıldığında, lojistik regresyon analizinde doğrusal regresyonda kullanılan yöntemlerin genel prensipleri izlenir.

Herhangi bir regresyon probleminde kilit değer, bağımsız değişkenin bilindiğinde, bağımlı değişkenin ortalamasıdır. Bu değere koşullu ortalama denir ve $E(Y|x)$, "x değeri değişkenini belirtirken x, bağımsız değişkenin değerini belirtir. $E(Y|x)$, "x değeri verildiğinde Y'nin beklenen değeri" olarak okunur. Doğrusal regresyonda, bu ortalamanın 4.2'deki gibi doğrusal bir denklem olarak ifade edilebileceği varsayılır.

$$E(Y|x) = \beta_0 + \beta_1 x \quad (4.2)$$

Koşullu ortalama sıfıra veya 1'e yaklaştıkça, x'deki birim değişimin $E(Y|x)$ üzerindeki değişimi giderek azalır. Şekil 4.3'te de görüleceği üzere eğrinin S şeklinde olduğu söylenebilir. Rastgele değişkenlerin kümülatif dağılım grafiğini andırır. Y'nin ikili olduğu durumlarda $E(Y|x)$ 'e bir model sağlamak için iyi bilinen bazı kümülatif dağılımlar kullanılır. Kullanacağımız model, lojistik dağılım modelidir.



Şekil 4.3 : Lineer regresyon ve lojistik regresyon farkının grafikte gösterimi (Le, 2018).

İkili bağımlı değişkeninin analizinde kullanılması için birçok dağılım fonksiyonu önerilmiştir. Lojistik dağılımın seçilmesinin iki temel nedeni vardır. Birincisi, matematiksel olarak son derece esnek ve kolay kullanılan bir fonksiyondur. İkincisi, kendisini klinik olarak anlamlı bir yoruma yönlendirir. Lojistik regresyonlarda x bilinirken Y 'nin koşullu ortalamasını hesaplarken $\pi(x) = E(Y|x)$ eşitliği kullanılır. $x - \infty$ ile $+\infty$ arasında değişirken $\pi(x)$ 0 ile 1 arasında değer alır ve denklem 4.3'te gösterilmiştir.

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (4.3)$$

$\pi(x)$ 'in dönüşümüne logit dönüşümü denir ve denklem 4.4'te gösterilmiştir.

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x \quad (4.4)$$

Şu ana kadar tek değişkenli lojistik regresyon işlendi. Şimdi birden fazla bağımsız değişkene sahip olan “çok değişkenli lojistik regresyon” incelenecek. $x' = (x_1, x_2, \dots, x_p)$ vektörü ile gösterilen p bağımsız değişkenden oluşan bir koleksiyon düşünün. Çok değişkenli lojistik regresyonun logit eşitliği denklem 4.5'teki gibidir

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (4.5)$$

Lojistik regresyon modeli ise denklem 4.6'da gösterilmiştir.

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \quad (4.6)$$

Bağımsız değişkenlerden bazıları kesikli ise (ırk, cinsiyet, tedavi grubu vb.) nominal değişkenler denir. Bu değişkenlerin aralık ölçeği ile ifade edilmesi uygun değildir. Modelin, bağımlı değişkeninin ikiden fazla kategorili ve nominal nominal olması durumunda multinomial lojistik regresyon modeli olarak adlandırılır (Hosmer ve Lemeshow, 2000).

4.5 k-En yakın komşular

KNN, yeni örnek sorgulamasının sonucunun k en yakın komşu kategorisinin çoğunluğuna göre sınıflandırıldığı denetimli öğrenme algoritmasıdır. Bu algoritmanın amacı, yeni bir nesneyi niteliklere ve eğitim örneklerine göre sınıflandırmaktır (Dharani ve Aroquiaraj, 2013).

KNN algoritması, özellik uzayındaki en yakın eğitim örneklerine göre nesneleri sınıflandırmak için bir yöntemdir ve tüm yapay öğrenme algoritmalarının en basitleri arasındadır (Kim ve diğ., 2012).

KNN, parametrik olmayan tembel bir öğrenme algoritmasıdır. Bu sebeple veri dağılımı ile ilgili herhangi bir varsayımda bulunmaz. Pratik verilerin çoğu teorik varsayımlara uymadığından KNN algoritmasının parametrik olmaması ciddi bir avantaj sağlar. KNN aynı zamanda tembel bir algoritmadır, bu da herhangi bir genelleme yapmak için eğitim veri noktalarını kullanmadığı anlamına gelir. Yani, eğitim aşaması oldukça hızlıdır. KNN, tüm eğitim verilerine göre karar verdiği için tüm eğitim verilerini tutmaktadır. Bu da genelleme olmaması anlamına gelir.

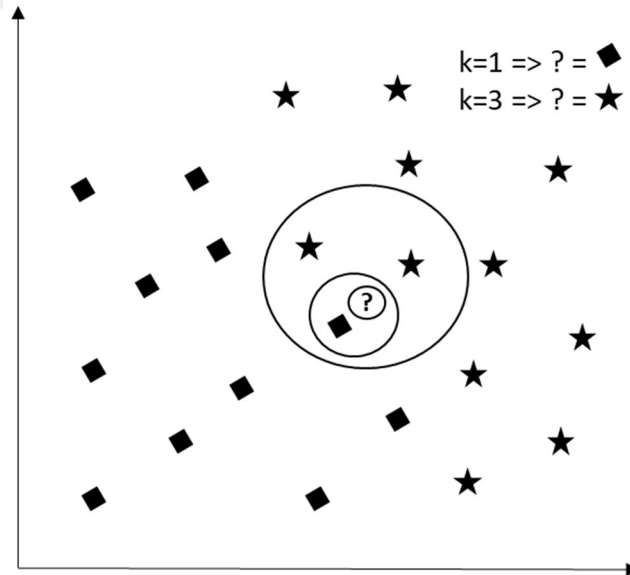
KNN aşağıdaki gereksinimlere sahiptir:

- KNN, verilerin bir öznitelik uzayında olduğunu varsayar. Daha doğrusu, veri noktaları metrik uzayındadır.
- Veriler skaler veya çok boyutlu vektörler olabilir. Noktalar öznitelik uzayında olduğundan, bir uzaklık kavramına sahiptirler.
- KNN, toplu olarak test veri kümesi adı verilen yeni veri noktalarını sınıflandırmak için eğitim verilerini referans olarak kullanır

- Eğitim verileri, dizi vektörden ve sınıf etiketinden oluşur. En basit durumda, 1 veya 0 olabilir. Ancak KNN daha fazla sınıf sayısı ile eşit derecede iyi çalışabilir.
- Ayrıca tek bir 'k' numarası verilir. Bu sayı, kaç komşunun (mesafe ölçüsüne göre tanımlandığı yerler) sınıflandırmayı etkileyeceğini belirler (Anchalia ve Roy, 2014).

KNN sınıflandırması, hem girdi hem de hedef değişkenleri içeren bir eğitim veri seti kullanılarak ve daha sonra sadece giriş değişkenlerini içeren test verilerinde bilinmeyen girdinin K'ye olan mesafesini karşılaştırılarak gerçekleştirir. KNN, sınıf atamasını ya en yakın K referans noktalarının sınıf numaralarının ortalamasını alarak veya onlar için çoğunluk oyu alarak belirler (Ramteke ve Monali, 2012).

KNN, bir nesneyi en benzer sınıfa veya sınıflara atayarak giriş vektörünü sınıflandırır. Örnek nesneye en yakın olan komşuların veya bilinen nesnelerin sayısı k'dır. Eğer k = 1 ise bilinmeyen nesne basitçe en yakın komşunun sınıfına atanır. Bir örnek nesnenin diğerine ne kadar yakın olduğunu belirleyen benzerlik işlevi, giriş vektörünün değerleri ile diğer örnekler için vektörün değerleri arasındaki Öklid uzaklığı olabilir. Bununla birlikte, özellik vektörlerindeki değerlerin normalleşmesinin doğru şekilde yapılması kritik önem taşır. (Erickson ve diğ, 2017)



Şekil 4.4 : En yakın komşu (Erickson ve diğ, 2017).

Uzaklık metriği, KNN sınıflandırmasında önemli bir rol oynar, çünkü doğruluğun performansını etkiler. Minkowski uzaklığı denklem 4.7'de gösterildiği gibidir.

$$(\sum |x_i - y_i|^p)^{1/p} \quad (4.7)$$

$p=1$ olduğunda Manhattan uzaklığı elde edilir. Aynı zamanda L_1 uzaklığı, taksi uzaklığı veya şehir blokları uzaklığı olarak da bilinir. Denklem 4.8’de gösterilmiştir.

$$\sum |x_i - y_j| \quad (4.8)$$

$p=2$ olduğunda öklid uzaklığı elde edilir. Öklid uzaklığı literatürde kullanılan en popüler uzaklık yöntemidir. Denklem 4.9’da gösterilmiştir (Ooni ve diğ, 2013).

$$\sqrt{\sum (x_i - y_i)^2} \quad (4.9)$$

KNN algoritması, ağırlık fonksiyonu tekdüze ve mesafe olarak göre iki gruba ayrılır. Tekdüze ağırlıklandırmada bütün noktalar eşit ağırlık alarak hesaplanır. Mesafeye göre ağırlıklandırmalarda ise yakın olan noktaların ağırlıklıkları daha fazla, uzak olan noktaların ağırlıkları daha azdır (Pedregosa ve diğ, 2011).

4.6 Destek vektör makineleri

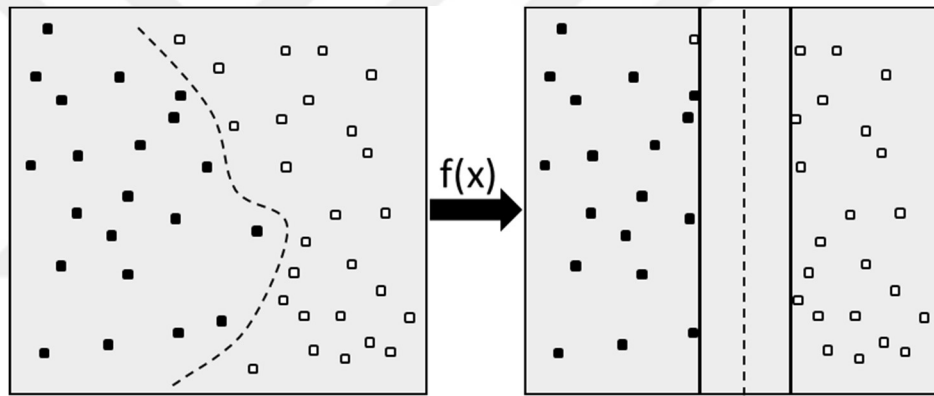
DVM, Vapnik tarafından diğer araştırmacılarla birlikte önerildiğinden beri, birçok alanda geniş çapta çalışılmış ve uygulanmıştır. DVM'nin temel fikri, iki sınıf arasındaki mesafeyi maksimize etmeye çalışmaktır (Huang ve diğ, 2013).

Denetimli sınıflandırmanın değeri o sınıflandırmanın doğruluk fonksiyonudur. Sınıflandırmadaki temel hedeflerden biri, az sayıda eğitim verisi ile yüksek doğruluk elde ederek sınıflandırma sürecini olabildiğince faydalı ve ekonomik hale getirmektir. Bunun için kullanılabilecek cazip sınıflandırıcılardan birisi de destek vektör makineleridir. DVM sınıflandırmaları en büyük olabilirlik yönetiminden, karar ağaçlarından ve yapay sinir ağı tabanlı yaklaşımlardan daha tutarlı sınıflandırma yapabilir. DVM, sınıf dağılımlarının kenarında yer alan ve destek vektör olarak adlandırılan eğitim örneklerine odaklanarak sınıflar arasında optimum ayırıcı aşırıdüzlemi yerleştirmeyi amaçlar. Optimum ayırıcı aşırıdüzlem, destek vektörleri kümeleri arasında maksimum mesafeye yerleştirilecek şekilde yönlendirilir. Bu yönlendirme sayesinde DVM’nin görülmeyen durumlarda, sinir ağları gibi eğitim hatasını en aza indirmeyi amaçlayan sınıflandırıcılara göre daha doğru genellemesi beklenir. DVM’ler ikili sınıflandırma için tasarlanmıştır. Bununla birlikte DVM, çok sınıflı senaryolar için genişletilebilir. (Mathur ve Foody, 2008).

DVM, yapay öğrenme ve örüntü sınıflandırması alanında son zamanlarda önem kazanmıştır. Sınıflandırma, girdi uzayında doğrusal veya doğrusal olmayan bir ayırma

yüzeyi gerçekleştirilerek elde edilir. DVM sınıflandırmasında ayırma işlevi, destek vektörleri ile ilişkili çekirdeklerin doğrusal bir kombinasyonu olarak ifade edilebilir (Vishwanathan ve Murty, 2002).

DVM, giriş verilerini iki sınıf arasındaki en geniş düzlem veya destek vektörünü üretecek şekilde dönüştürür. DVM, geniş düzlemde dolaylı olarak yanlı olan nokta sayısına karşı geniş bir ayırma düzlemine sahip olmak istediği derecenin esnek bir şekilde seçilmesine izin verir. Bu öğrenme makineleri yakın zamanda oluşturuldu ve son zamanlardaki popülerliklerinin nedeni, doğrusal olmayan ilişkileri kullanarak noktaları başka boyutlarla eşleştirebilen ve böylece doğrusal olarak ayrılamayan örnekleri sınıflandırabilen temel işlevlerin eklenmesidir. Bu yetenek, DVM algoritmalarını diğer birçok yapay öğrenme yöntemine göre büyük bir avantaj sağlar. (Erickson ve diğ, 2017)



Şekil 4.5 : Destek vektör makineleri (Erickson ve diğ, 2017).

En basit DVM sınıflandırma senaryosu, sınıfların doğrusal olarak ayrılmasıdır. Bu senaryonun veri seti denklem 4.10’da gösterilmiştir.

$$\{x_i, y_i\}, i = 1, \dots, k \quad (4.10)$$

N boyutlu bir uzayda $x \in \mathbb{R}^N$ ve $y \in \{-1, +1\}$ sınıf etiketleridir. Bu eğitim örüntüleri, denklem 4.11’deki gibi ayırıcı düzlemin yönünü belirleyen w vektörü ve ayırıcı düzlemin orijine mesafesini belirleyen b skaler değeri varsa doğrusal olarak ayrılabilir.

$$y_i(w x_i + b) - 1 \geq 0 \quad (4.11)$$

DVM, aşırıdüzlemde dik bir çizgi boyunca ölçülen sınıflar arasındaki mesafenin maksimize edildiği ayırıcı aşırıdüzlemleri bulur. Bu işlemi denklem 4.12’de belirtilen ve denklemin 4.11’in kısıtı olduğu amaç fonksiyonu ile yapar.

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (4.12)$$

Doğrusal olarak ayrılamayan sınıflar için, belirli bir sınıfın tüm eğitim durumlarının optimum aşırıdüzlemin aynı tarafında yer alması kısıtlaması, zayıf değişken $E_i \geq 0$ eklenmesiyle gevşetilebilir. Bu durumda, DVM mesafeyi en büyükleyen ve hatalı sınıflandırmayı en küçükleyen aşırıdüzlemi bulmayı amaçlar. Mesafe ve hatalı sınıflandırma arasındaki bu değiş tokuşu pozitif C değişkeniyle kontrol eder. Doğrusal olarak ayrılamayan veriler için denklem 4.13'teki amaç fonksiyonu kullanılır.

$$\min_{w,b,E_i} \left[\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^k E_i \right] \quad (4.13)$$

Doğrusal olmayan durumlarda $x \in \mathbb{R}^N$ özellik vektörü doğrusal olmayan $\Phi: \mathbb{R}^N \rightarrow F$ vektör fonksiyonu aracılığı ile daha yüksek boyutlu bir F öklid uzayına taşınır. F boyutundaki optimum mesafe problemi x_i, x_j yerine $\Phi(x_i), \Phi(x_j)$ şeklinde yazılır. Bu sorunu çözmek için doğrusal olmayan DVM'lerde çekirdek işlevi K kullanılır. Denklem 4.14'te gösterilmiştir (Pal ve Foody, 2010).

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (4.14)$$

Çekirdekler birer fonksiyondur ve bu çekirdeklerde kullanılan değişkenler çizelge 4.1'de gösterilmiştir (Doğan, 2019).

Çizelge 4.1 : Çekirdek tipleri ve fonksiyonları

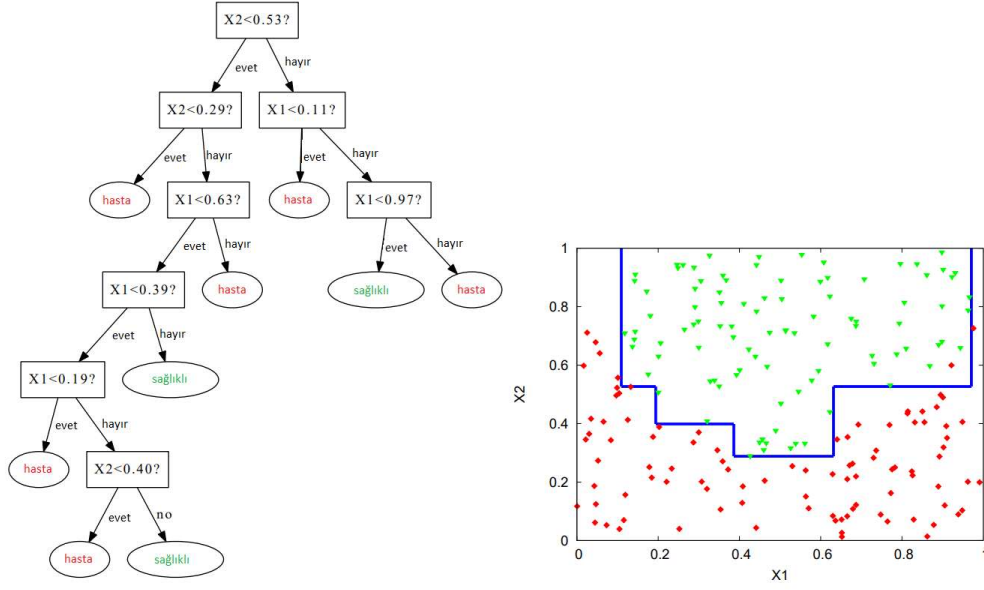
Çekirdek Tipi	Fonksiyon
Doğrusal	$K(x_i, x_j) = x_i^T x_j$
Polinom	$K(x_i, x_j) = (g x_i^T x_j + b)^d, g > 0$
Radyal Tabanlı Fonksiyon (RTF)	$K(x_i, x_j) = \exp(-g \ x_i - x_j\ ^2), g > 0$
Sigmoid	$K(x_i, x_j) = \tanh(g x_i^T x_j + b)$

4.7 Karar ağaçları ve Rastgele Orman

KA, günümüzde en ilgi çekici denetimli öğrenme algoritmalarından biridir. Ağaç temelli öğrenme yöntemlerinin kökeni, kavram öğrenme sistemi ile Hunt'a atfedilir. Daha sonra bu yöntem genişletildi ve popüler hale geldi. Başlangıçta sınıflandırma problemlerini çözmek için tasarlanmış olan KA algoritmaları, daha karmaşık problemleri çözebilecek şekilde genişletildi (Geurts ve diğ, 2009).

KA fikri sıradan bir ağaçtaki dal ve yapraklardan gelir. KA, kök ve dalların bölündüğü konum olan düğümlerden oluşur. Dairelerle gösterilen düğümlerle bir karar ağacı oluşturulur ve bu düğümleri birbirine bağlayan bölümler dallar ile temsil edilir. KA, köken başlayıp aşağı doğru ilerler ve genellikle soldan sağa doğru çizilir. Ağacın başladığı düğüme kök düğüm, bittiği düğümlere ise yaprak düğüm denir. Yaprak düğüm olmayan iç düğümlerden iki veya daha fazla dal uzatılabilir. Bir düğüm belirli bir özelliği temsil ederken dallar bir değer aralığını temsil eder (Ali ve diğ, 2012).

Karar ağaçları, girdilerin sınıflanmasında insan tarafından okunabilirlik sağlar. Bu da diğer yapay öğrenme algoritmalarına karşı büyük bir avantajdır. Karar ağaçları çoğu insan için tanındır ve genellikle evet veya hayır cevabı verilecek şekilde bir sorular sorulur. Yapay öğrenme, en doğru sonucu veren en basit ağacı oluşturacak karar noktaları için olası birçok kombinasyonu hızlı bir şekilde arayarak karar ağacının yönünü oluşturur. Algoritma çalıştırıldığında, maksimum derinliği ve aranacak maksimum genişliği belirler ve daha fazla karar noktasına göre doğru sonuçlara sahip olmanın ne kadar önemli olduğunu belirler. Bazı durumlarda, birden fazla karar ağacının yapıldığı bir topluluk yöntemi kullanarak doğruluk artırılabilir. Rastgele bir orman sınıflandırıcısı, sınıflandırma oranını arttırmak için birkaç karar ağacı kullanmasına rağmen, genellikle yüksek performans gösterse de verileri yeniden örneklemiyor. Ek olarak, rastgele ormanlarda, toplam özellik sayısının yalnızca bir alt grubu rastgele seçilir ve alt kümeden en iyi ayrılma özelliği, her düğümü bir ağaçta bölmek için kullanılır. Torbalamadan farklı olarak, tüm özellikler bir düğümü bölmek için düşünülür (Erickson ve diğ, 2017).



Şekil 4.6: Karar ağacı diyagramı ve örnek düzlemi (Geurts ve diğ., 2009).

KA oluşturma esnasında iki ana işlem vardır. Her özellik için ayrımların değerlendirilerek en iyi düğümün bulunması ve bulunan düğümü kullanarak yeni ayrımlar yapılması. Genellikle en iyi düğümü bulduktan sonra ayırma kriterini basitçe uygulayarak yeni düğümler oluşturulur. Bu işlemlerdeki karmaşıklık ise en iyi düğümü belirlemektir. Bir özelliğin ayrımlarının iyiliğini değerlendirmek için bir ayırma endeksi kullanılır. Bu ayırma endekslerinden olan Entropi ve Gini endeksleri m sınıf içeren S veri kümesi ile denklem 4.15 ve 4.16’da sırasıyla gösterilmiştir.

$$Entropi(S) = - \sum_{j=1}^m P_j \log P_j \quad (4.15)$$

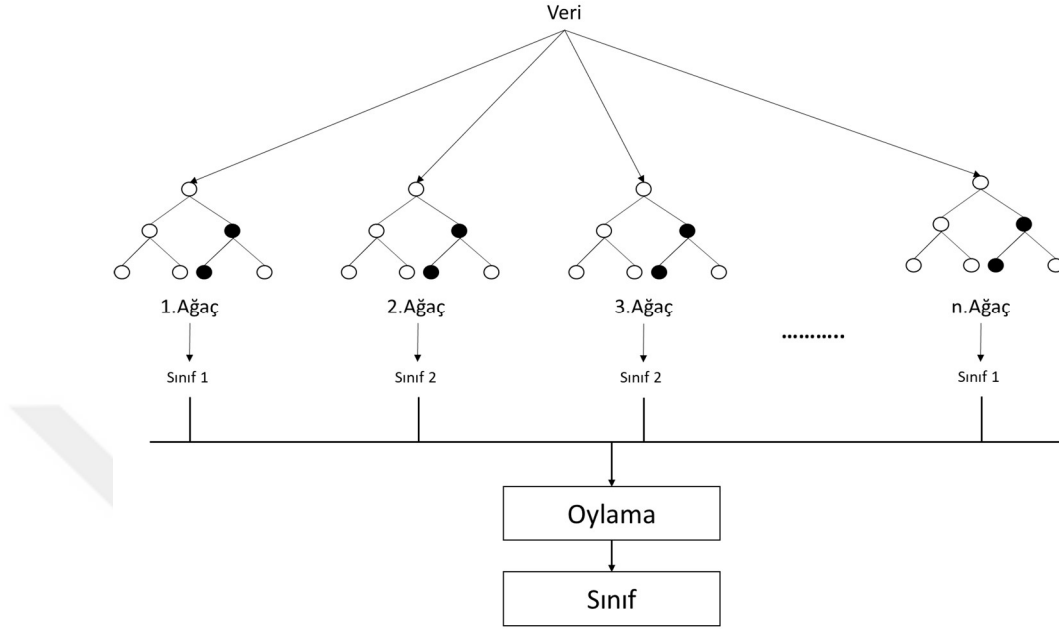
$$Gini(S) = 1 - \sum_{j=1}^m P_j^2 \quad (4.16)$$

Burada P_j , S kümesindeki j sınıfının göreceli frekansıdır. Entropi veya gini endeksine bağlı olarak veri kümesini ayırmak için A özneliği kullanılırsa bilgi kazancı hesaplanabilir. İlgili formül denklem 4.17 ve 4.18’de gösterilmiştir.

$$Kazanç(S, A) = Entropi(S) - \sum_{v \in A} \left(\frac{|S_v|}{|S|} * Entropi(S_v) \right) \quad (4.17)$$

$$Kazanç(S, A) = Gini(S) - \sum_{v \in A} \left(\frac{|S_v|}{|S|} * Gini(S_v) \right) \quad (4.18)$$

v , A özniteliğinin olası değerlerini temsil eder. S_v ise A özniteliğinin v değerine sahip olduğu S 'nin alt kümesini temsil eder. $|S_v|$, S_v kümesinin eleman sayısı, $|S|$ ise S kümesinin eleman sayısıdır (Du ve Zhan, 2002).



Şekil 4.7 : Rastgele orman yöntemi (Soylu, 2018).

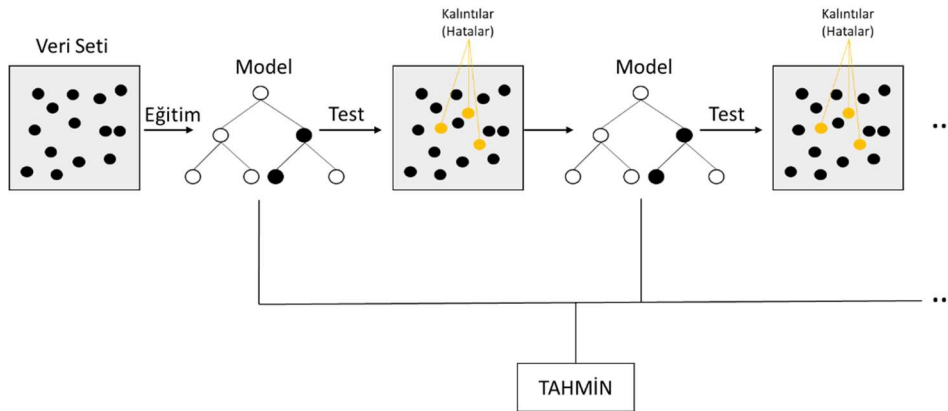
Rastgele orman aynı zamanda torbalı karar ağaçları olarak bilinir. RO, yöntemi birden fazla KA içeren toplu yöntemdir. Yani RO'ya her girdi verilişinde bu girdi kurucu KA'ların her birinden aşağı aktarılır. Her KA bağımsız olarak bir sınıflandırmayı öngörür ve ilgili sınıf için oy verir. RO ise bu verilen oyların çoğunluğuna göre tahmine karar verir. KA'nın bu oyu, tek bir KA çıktısına göre doğal olarak daha az gürültülü ve aykırı değerlere karşı daha az hassastır. Böylece küçük verilerden kaynaklanan oynaklığı azaltarak tahminlerin doğruluğunu artırır. RO özellik seçme sistemine sahiptir. Böylece boyut indirgemeye gerek olmadan çok sayıda girdi parametresi ile işlem yapabilir (Shaikhinave diğ, 2019).

4.8 Gradyan Artırma

GA, parametreleştirilmiş basit bir fonksiyonu her yinelemede kalıntıları en küçük kareler yöntemine göre yerleştirerek toplam regresyon modelini oluşturur. Kalıntılar, mevcut adımda değerlendirilen her eğitim verisinin model değerine göre en aza indirilen kayıp fonksiyonunun gradyanıdır. GA hem doğruluğu hem de hızı büyük ölçüde iyileştirdiği kanıtlanmıştır (Friedman, 2002).

GA, regresyon ve sınıflandırmada kullanılan bir yapay öğrenme yöntemidir. GA, topluluk öğrenme metodunun yükseltme yöntemlerinden biridir. Topluluk öğrenme ikiye ayrılır. Rastgele orman yönteminin örneklerinden biri olduğu bulunduğu torbalama yöntemi ve yükseltme yöntemi. Torbalama yönteminde birden fazla algoritma birbirinden bağımsız çalışır ve yaptıkları tahminlere göre nihai tahmin seçilir. Yükseltme yönteminde ise algoritmalar birbiriyle ilişki içinde sıralı olarak çalışır.

Denetimli öğrenme yöntemlerinin amacı kayıp fonksiyonunu en küçükmektir. GA'nın arkasında yatan mantık ise basit sınıflandırma algoritmalarını kullanarak bir model oluşturmak. Daha sonra oluşan kalıntıları tekrar tekrar kullanarak modeli güçlendirmek ve daha iyi hale getirmektir. Bu işlem kalıntıları herhangi bir örüntü oluşturamayana kadar devam eder. Daha fazla devam edilmesi aşırı öğrenmeye sebep olur. Bu ardışık işlem Şekil 4.8'de gösterilmiştir (Grover, 2017).



Şekil 4.8 : Gradyan artırma modeli (Ledell, 2016).

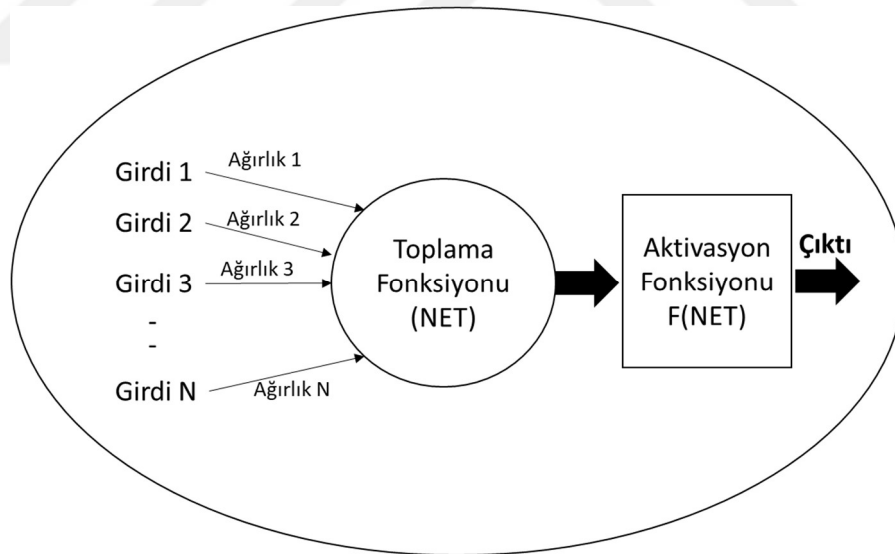
GA'nın 3 ana bileşeni vardır. Kayıp fonksiyonu, zayıf öğreniciler ve toplanır modeldir. Kayıp fonksiyonu, modelin tahminlerde ne kadar iyi olduğunu tahmin eder. Zayıf öğrenici, verileri sınıflandıran ancak çok zayıf tahmin yapabilen modellerdir. Diğer bir deyişle hata oranı yüksektir. En yaygın kullanılan zayıf öğrencilerden birisi karar ağaçlarıdır. Katkı modeli ise bu zayıf öğrencilerin her seferinde bir adım eklemenin yinelenmeli ve sıralı yaklaşımıdır. Her yineleme sonrası kayıp fonksiyonunun değeri azalmalı ve nihai modele yaklaşmalıdır (Kurama, 2020).

4.9 Yapay Sinir Ağları

Bir bilgi işlem teknolojisi olan yapay sinir ağları, insan beyninin çalışma ve düşünebilme yeteneğinden yola çıkarak oluşturulmuştur (Aktaş ve diğ, 2003). Lojistik regresyon veri sınıflandırmada yapay sinir ağlarının basit halidir (Adepoju, 2019). Yapay sinir ağları biyolojik sinir ağından ilham alınarak geliştirilmiş ve örüntü tanıma, tahmin, optimizasyon ve kontrol konularında birçok porblemi çözebilen zaki sistemlerdir (Jain ve diğ, 1996).

Yapay sinir ağaları, gürültüyü ve eksik verileri tolere edebilen, bağlantılarını değiştirerek kendi kendine öğrenebilen, girdi verileriyle karmaşık ilişkiler kurabile esnek ve parametrik olmayan bir model yaklaşımıdır (Tang ve Chi, 2005).

İnsan vücudundaki sinir ağlarının sinir hücreleri olduğu gibi yapay sinir ağlarının da yapay sinir hücreleri vardır. Yapay sinir ağlarına mühendislik biliminde süreç elemanları da denilmektedir. Şekil 4.9'da gösterildiği gibi her süreç elemanının girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktıları olmak üzere toplam beş temel elemanı vardır (Öztemel, 2006).



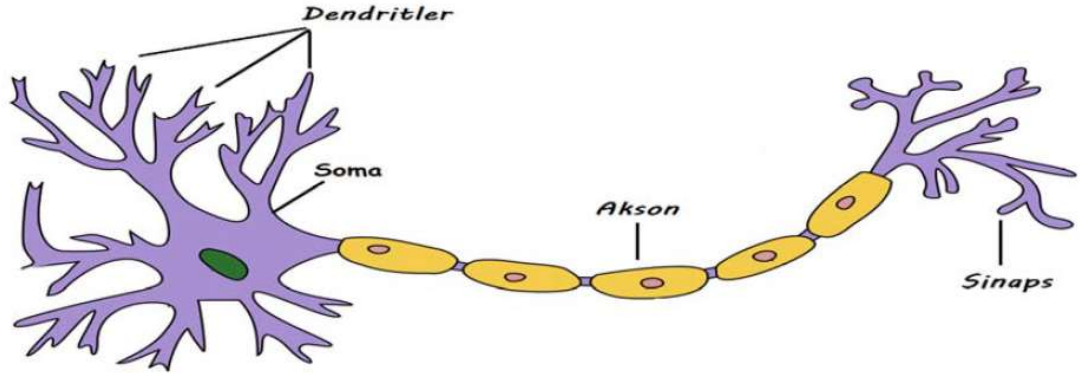
Şekil 4.9 : Yapay sinir ağı yapısı (Öztemel, 2006).

Öğrenme olayı biyolojik beynin en önemli özelliklerinden birisidir. Hayvanlar ve insanlar devamlı bulundukları ortam ile etkileşimi sonucunda öğrenme işlemi gerçekleştirir. Öğrenilen bilgiler beynin işlevlerini etkileyecek davranışlarla kendini gösterir. Yapay sinir ağlarının temelini bu özellik oluşturur (Öztemel, 1996).

YSA, yapay öğrenme yöntemlerinde gözetimli öğrenme yaklaşımlı modellerinden biridir ve etiketli girdiler kullanır (Mohri ve diğ, 2018).

İnsan beyinde çok sayıda nöron veya sinir hücresi vardır ve her sinir hücresi sinapslar yoluyla diğer birçok sinir hücresine bağlanır ve çok karmaşık sinir ağları oluşturur. Bu yüzden sinaptik fonksiyon çok önemlidir. Bir nöron başka bir nörona sinapslar aracılığı ile mesaj gönderir. İnsanlar kendi kendini öğrenme yeteneğine sahiptir, çünkü sinir ağları çok sayıda paralel hesaplama ve dağıtılmış bilgiyi işleyebilir. Bu nedenle, insan ve hayvanların beynini taklit etme becerilerini kullanmak için yapay sinir ağları modeli geliştirildi. Yapay sinir ağları, beyin ve sinir sistemi ile ilgili bilgileri işlemek için taklit edilen bir teknolojidir ve düğüm olarak da adlandırılan birçok yapay nöron sınıfından oluşur. Yapay sinir ağları, bilgisayarların hızlı hesaplama özelliğini kullanarak çok sayıda paralel hesaplama yapabilir. Olgun insan beyni altı katmandan oluşan katmanlı bir yapıya sahiptir. Tüm kavrama işlemleri beyinde gerçekleşir. Böylesine karmaşık bir öğrenmenin gerçekleşebilmesi beyindeki nöron sayısına, katmanların yapısına ve bağlantı türüne bağlıdır. Yapay sinir ağları, matematiksel fonksiyonu ve parameterleri ile insan beyindeki bu yapıyı taklit etmeye çalışır (Lin, 2017).

Sinir hücresi diye adlandırılan nöronlar bilgi işleyen özel biyolojik hücrelerdir. Soma olarak da tanımlanan gövdesi ve akson ve dendritlerden oluşan ağaç benzeri dallardan oluşmaktadır. Hücre gövdesi, kalıtsal özellikler hakkında bilgilerin bulunduğu çekirdeğe ve nöronun ihtiyaç duyduğu malzemeleri üreten moleküler ekipmanları tutan bir plazmaya sahiptir. Nöronlar, alıcı görevi gören dendritler vasıtasıyla diğer nöronlardan impuls olarak bilinen sinyalleri alır ve hücre gövdesi tarafından üretilen sinyalleri akson yani verici boyunca iletir ve son olarak bu sinyaller uçlarında sinapsların olduğu ipliklere ve alt ipliklere iletilir. Sinapslar bir nöronun dendritiyle diğer nöronun akson ipliği arasındaki temel yapı ve fonksiyonel birimdir. Sinyaller sinapsların ucuna ulaştığında nörotransmitter adı verilen bazı kimyasallar salgılanır. Nörotransmitterler sinapsların tipine bağlı olarak reseptör nöronun elektriksel uyarı yayma eğilimini arttırmak veya engellemek için sinaptik boşlukta yayılır. İnsan beyninin selebral korteks olarak adlandırılan beyin zarında toplam 1011 nöron bulunur. Her nöron yaklaşık 104 farklı nörona bağlıdır. Bu da bir insan beyinde toplam 1015 ara bağlantı vardır (Jain ve diğ, 1996).



Şekil 4.10 : Sinir hücresinin kısımları (Ghrai, 2019).

Şekil 4.10’da da gösterildiği gibi yapay sinir ağları beş kısımdan oluşur. Bu matematiksel modeli denklem 4.19’da verilmiştir. x_1 ve x_2 son nörondan gelen girdiler, w_1 ve w_2 nöronun ağırlıkları, b bias değeri, ve n ise eşik değerdir.

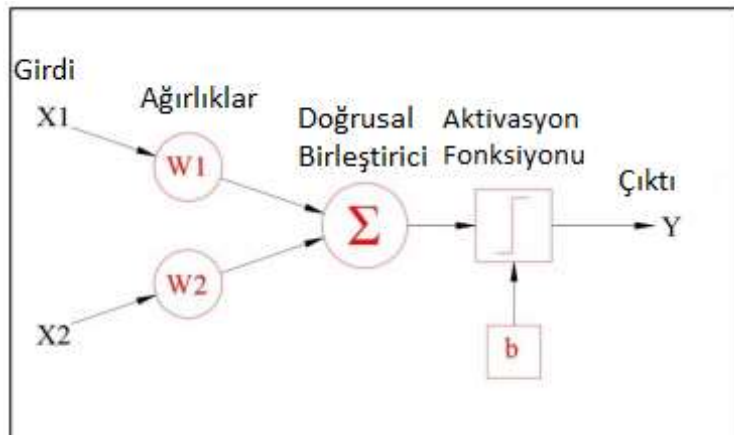
$$n = [w_1 w_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b \quad (4.19)$$

Denklem 4.20’de ise bu eşitliğin matris açılımı verilmiştir.

$$n = w_1 x_1 + w_2 x_2 + b \quad (4.20)$$

F fonksiyonu doğrusal olmayan ve türevlenebilir bir fonksiyondur. Sonucunda yapay sinir ağının çıktısı olan y değerini verir. Denklem 4.21’de gösterilmiştir (Lin, 2017).

$$y = f(n) \quad (4.21)$$



Şekil 4.11 : Yapay sinir ağı kısımlar (Lin, 2017).

4.10 Özellik seçimi

Yapay öğrenmede ilk adım, karar vermek için kullanılacak olan bilgileri içeren özellikleri çıkarmaktır. İnsanlar radyoloji birimlerinde görsel olarak önemli özellikleri öğrenirler, fakat bir özelliği hesaplamak veya göstermek zor olabilir. Görüntü özellikleri, tıbbi görüntüleme verileriyle çalışırken görülen en yaygın değişikliklerden bazıları olduğundan, gürültü, yoğunluk ve dönüş açılarındaki değişikliklere karşı esnek olmalıdır.

Bir görüntüden birçok özelliği hesaplamak mümkün olsa da çok fazla özelliğe sahip olmak, kararın gerçek temelini öğrenmek yerine aşırı uyuma neden olabilir. Özellik seçimi en iyi tahmini yapabilmek için kullanılması gereken özelliklerin alt kümesini seçme işlemidir. Özellikler arasındaki korelasyonları araştırmak, bu özellikleri seçmede bir yöntemdir. Çok sayıda ilişkili özelliğe sahip olmak muhtemelen özelliklerin sayısının bilgi kaybı olmadan azaltılabileceği anlamına gelir. Bununla birlikte, bazı durumlarda daha karmaşık bir ilişki vardır ve izolasyondaki bir özelliğin değerlendirilmesi tehlikelidir. Örneğin, size her kilonun obeziteyi gösterip göstermediğine ya da göstermediğine ilişkin ikili sınıflandırmalara sahip bir ağırlık listesi verildiğini düşünün. Bazı tahminler yapılabilir, ancak yükseklikleri eklemek doğruluğu artırabilir: Düşük bir yükseklik değeriyle birlikte oldukça yüksek bir ağırlık değerinin obeziteyi yansıtmaması daha yüksek bir yüksek değer ile birlikte yüksek bir ağırlık değerinden daha yüksektir (Erickson ve diğ, 2017).

4.11 Doku analizi ve Haar Benzeri Özellikler

Günlük hayatta doku denilince akla cisimlerin yüzeylerindeki kaba, ipeksi, engebeli ve parmakla kolaylıkla anlaşılabilen yüzeydeki geçişler gelir. Görüntü dokusu da çok farklı değildir. Sadece cisimlerin yüzeylerindeki pürüzlerdeki yükseklik ve alçaklıklar yerine resimlerde piksel değerlerinde geçiş vardır. Gri düzeydeki resimlerde bu geçişler parlaklık geçişleri olarak isimlendirilebilir. Bu geçişleri anlamak için gri seviyeli eş-oluşum matrisleri kullanılır. Bu matrislerden türetilen doku ölçüleri oldukça karmaşık olmalarına rağmen tanımlayıcı ölçülerdir.

Gri seviyeli eş-oluşum matrisleri ile doku ölçüm hesaplamalarının çoğu 1970’li yıllarda Robert Haralick ve arkadaşları tarafından bir dizi makalede sistematikleştirildi. Bu çalışmalarında matematiksel desen analizi ile çalışan birkaç

araştırmacının çalışmalarını temel aldılar. Harlick ve arkadaşları 1973'te 4 farklı yönde 14 farklı doku ölçümü hesaplanmasını ve bunların ortalamasının alınmasını önerdi. Ancak bu çok benimsenmedi.

4.11.1 Gri seviyeli eş-oluşum matrisleri

GSEM, bir seferde iki piksel arasındaki referans ve komşu piksel adı verilen ilişkiyi dikkate alır. Şekil 4.12'da 0-4 arasında piksel değerlerine sahip olan 4x4'lük bir görüntünün piksel matrisi verilmiştir.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Şekil 4.12 : Örnek 4x4'lük resim piksel değerleri.

Her bir matris değerinin sağındaki değer o değer komşu pikseli olarak adlandırılmaktadır. Referans değerinden sonra gelen komşu değerlerinin kaç kez tekrarlandığı bilgisini gösteren yeni bir matris oluşturuluyor. Örneğin 0 referans pikselinden sonra yanına 0 komşu pikseli gelmesi Şekil 4.9'daki matriste 2 defa tekrarlanmıştır. Bu sebeple yeni oluşturulacak matrisin 0,0 bölümüne 2 yazılıyor. Oluşturulan yeni matris Şekil 4.13'te gösterilmiştir.

		Komşu			
		0	1	2	3
Referans	0	2	2	1	0
	1	0	2	0	0
	2	0	0	3	1
	3	0	0	0	1

Şekil 4.13 : Referans-komşu ilişki tekrarı matrisi.

Oluşturulan yeni matris simetrik değil ve analizlerde daha iyi sonuç almak için simetrik matris kullanılması amaçlanmaktadır. Bu sebeple oluşturulan matris transpozesi ile toplanarak simetrik hale getiriliyor. Yapılan işlem Şekil 4.11'de gösterilmiştir.

$$\begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Şekil 4.14 : Transpozesi ile toplanan yeni matris.

Elde edilen yeni matris değerleri denklem 4.22’te gösterildiği normalize edilerek doku ölçümleri için gerekli olan $P_{i,j}$ değerlerine dönüştürülüyor. V_{ij} matrisin i . Satırındaki j . Sütündaki değeri temsil etmektedir (Myrka, 2017).

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} v_{i,j}} \quad (4.22)$$

4.11.2 Haar benzeri özellikler

Myrka (2017)’ye göre 6 adet Haar benzeri doku ölçüm özelliği belirlendi. Bunlar kontrast, benzemezlik, türdeşlik, açısıl ikinci moment, enerji ve korelasyon.

Kontrastın hesaplanması denklem 4.23’te gösterilmiştir.

$$Kontrast = \sum_{i,j}^{N-1} P_{i,j} (i - j)^2 \quad (4.23)$$

Benzemezlik ölçüsünün hesaplanması denklem 4.24’te gösterilmiştir.

$$Benzemezlik = \sum_{i,j}^{N-1} P_{i,j} |i - j| \quad (4.24)$$

Türdeşlik ölçüsünün hesaplanması denklem 4.25’te gösterilmiştir.

$$Türdeşlik = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (4.25)$$

Açısıl ikinci moment (AİM) doku ölçüsünün hesaplanması denklem 4.26’da gösterilmiştir.

$$Açısıl İkinci Moment = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (4.26)$$

Enerji doku ölçüsünün hesaplanması denklem 4.27’de gösterilmiştir.

$$Enerji = \sqrt{AİM} \quad (4.27)$$

Korelasyon doku çözümlemesinin hesaplanması denklem 4.28’de gösterilmiştir.

$$Korelasyon = \sum_{i,j}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (4.28)$$

4.12 Değerlendirme Metrikleri

Yapay öğrenmede bir sınıflandırıcının değerlendirilmesi için doğruluk yaygın bit etkinlik ölçüsüdür. Doğru sınıflandırılan gözlem sayısının toplam gözlem sayısına oranı ile hesaplanır (Junker ve Hoch, 1999).

Kesinlik gibi diğer performans ölçüleri de her bir sınıftaki performansa duyarlı kaldıklarında daha uygun olabilir (Landgrebe ve diğ, 2006).

Yapay öğrenmede en iyi modelin hangisi olduğuna karar verebilmek için değerlendirme metrikleri kullanılır. Şekil 4.15'te hata matrisini ve elemanları gösterilmiştir.

	Tahminlenen	
Gerçekleşen	Doğru Pozitifler (DP)	Yanlış Negatifler (YN)
	Yanlış Pozitifler (YP)	Doğru Negatifler (DN)

Şekil 4.15 : Hata matrisi ve elemanları.

Doğru pozitif ve doğru negatif modelin doğru tahmin ettiği gözlemleri, yanlış pozitif ve yanlış negatifler ise modelin yanlış tahmin ettiği gözlemleri gösterir. Örneğin, birincil kökeni akciğer olan tümörler 1, metastaz olan tümörler 0 olduğunu varsayalım. O zaman doğru pozitif, tümörün birincil kökeninin akciğer olduğu ve modelin doğru tahmin ettiği gözlemler; doğru negatif, tümörün metastaz olduğu ve modelin doğru tahmin ettiği gözlemler; yanlış pozitif, tümörün metastaz olduğu fakat modelin yanlış tahmin ettiği gözlemler; yanlış negatif ise birincil kökeni akciğer olan fakat modelin yanlış tahmin ettiği gözlemlerdir. Ayrıca, yanlış pozitifte tip-1 hata, yanlış negatife tip-2 hata denilmektedir.

4.12.1 Doğruluk oranı

Doğruluk oranı, bir modelin başarısını ölçmek için kullanılır, ancak tek başına yeterli değildir. Doğru tahmin edilen gözlemlerin tüm gözlemlere oranı olarak hesaplanır. Denklem 4.29’da gösterilmiştir.

$$\text{Doğruluk Oranı} = \frac{DP+DN}{DP+DN+YP+YN} \quad (4.29)$$

4.12.2 Kesinlik değeri

Kesinlik skoru, tahmin edilen pozitif gözlemler arasında doğru olarak tahmin edilen pozitif gözlemlerin oranını göstermektedir. Denklem 4.30’da gösterilmiştir. Kesinlik değeri özellikle yanlış pozitiflerin olduğu durumlarda önem arz etmektedir. Bu sebeple kesinlik değerinin yüksek olması model seçiminde önemli bir etkindir.

$$\text{Kesinlik Değeri} = \frac{DP}{DP+Y} \quad (4.30)$$

4.12.3 Duyarlılık değeri

Duyarlılık değeri, pozitif olarak tahmin edilmesi gereken gözlemlerin ne kadarının pozitif olarak tahmin edildiğini gösteren bir metrik. Gerçekte pozitif olmasına rağmen negatif tahminler yüksekse duyarlılık değeri düşük çıkar. Model seçiminde yüksek olması beklenmektedir. Denklem 4.31’de hesaplaması gösterilmiştir.

$$\text{Duyarlılık Değeri} = \frac{DP}{DP} \quad (4.31)$$

4.12.4 F1 skoru

F1 skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır. Aritmetik ortalama yerine harmonik ortalama kullanılmasının başlıca sebebi uç örneklerde harmonik ortalamanın daha doğru sonuçlar vermesi. Örneği duyarlılık değeri 1 kesinlik değeri 0 olan durumlarda aritmetik ortalama 0,5 çıkarken harmonik ortalama 0 çıkacaktır. Denklem 4.32’de nasıl hesaplandığı gösterilmiştir.

$$F1 \text{ Skoru} = 2 * \frac{\text{kesinlik} * \text{duyarlılık}}{\text{kesinlik} + \text{duyarlılık}} \quad (4.32)$$

Model seçiliminde doğruluk oranı yerine F1 skorunun seçilmesinin en temel sebebi sınıflardaki gözlem sayıları eşit dağılmayan veri kümelerinde hatalı model seçimi

yapmamaktır. Ayrıca F1 skoru sadece yanlış pozitif ve yanlış negatif değil tüm hata maliyetlerini içeren bir değerlendirme metriğidir (Öğündür, 2019).

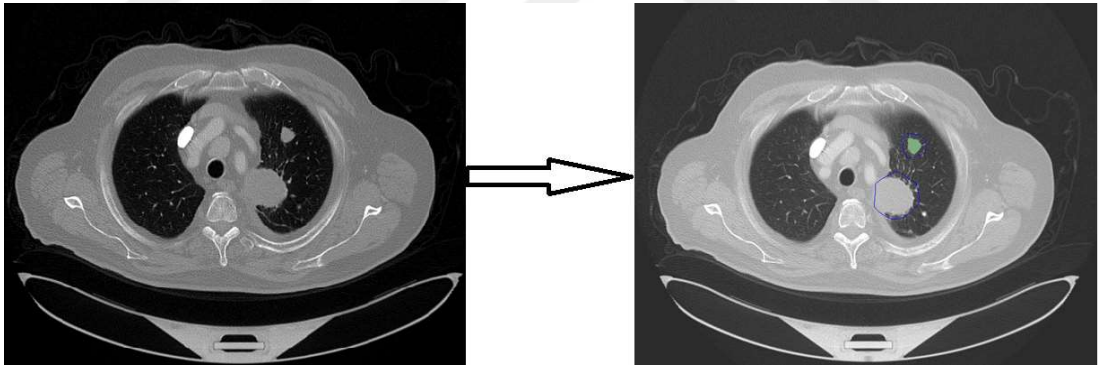




5. VAKA ANALİZİ

5.1 Veri Seti

Çalışmada kullanılan veriler, Bezmialem Vakıf Üniversitesi Hastanesi'nden alınmıştır. Veriler, BT çekilen ve akciğerinde kötü huylu tümör tespit edilen hastalar arasından uzman radyologlar tarafından incelenerek oluşturuldu. 10 mm'den daha büyük çapa sahip olan nodüller, 3D Slicer programındaki “segmentation wizard” modülünü kullanarak yarı otomatik segmentasyon ile segmente edildi. Segmente edilen kanser nodülleri “nrrd” formatında kaydedildi. Şekil 5.1’de solda 2 nodüle sahip olan bir hastanın orjinal BT kesit görüntüsü, sağda ise segmente edilirken alınan bir görüntü vardır. Şekil 5.2’de ise Şekil 5.1’de gösterilen tümörlerin segmente edildikten sonraki hali gösterilmiştir.



Şekil 5.1 : Segmente edilmiş ve edilmemiş akciğer BT görüntüsü.



Şekil 5.2 : Segmente edilmiş tümörler.

Kanserinin birincil kökeni akciğer olan 50 hasta, birincil kökeni akciğer olmayan ve akciğere metastaz yapmış 50 hastanın akciğer BT görüntülerinden segmente edilmiş tümör görüntülerine ön hazırlık işlemleri yapıldı. Hastalardaki her tümör farklı bir girdi olarak değerlendirildi. Böylece birincil kökeni akciğer olan 80 gözlem, akciğere metastaz yapmış 147 gözlem elde edilmiş oldu.

5.2 Veri Ön İşleme

Bu tezde verilere Python dili kullanarak yapay öğrenme yöntemleri uygulanmıştır. Hastaneden alınan nrrd uzantılı veriler nrrd kütüphanesi kullanarak gri sakala 512x512 boyutlarında ve kesit sayısı derinliğinde matrislere dönüştürüldü. Elde edilen matrislerdeki piksel değerleri 0-255 aralığına normalize edildi. Tümörün olduğu kesitler belirlendi. Temsili en yüksek olduğu düşünülen orta kesitteki tümör görüntüsünden, kesitlerin ortalamasından oluşan görüntüden ve kesitlerin standart sapmasından oluşan görüntüde ayrı ayrı özellik çıkarımı yapıldı. Birincil kökeni akciğer olan girdilerin etiketi 1, metastaz olan girdilerin etiketi 0 olarak belirlendi.

Öncelikle tümörün yatay ve dikey düzlemdeki çap uzunlukları hesaplandı. Daha sonra oluşturulan 3 farklı görüntülerin her birinden 6 adet haar-like özellikleri çıkarıldı. Ayrıca görüntülerin piksel değerlerinin ortalaması ve standart sapması alındı. Böylece toplam 26 özellik oluşturuldu. Bu özellikler Çizelge 5.1’de gösterilmiştir.

Çizelge 5.1 : Elde edilen özellikler.

Haar-like Özellikleri	Fiziki ve İstatistikî Özellikler
Kontrast	Yatay Çap
Benzemezlik	Dikey Çap
Türdeşlik	Ortalama
Enerji	Standar Sapma
Açısal İkinci Moment	
Korelasyon	

Oluşturulan bu 26 özellik ile 227x26 boyutlarında girdi matrisi elde edildi. Bir adet aykırı değer belirlendi ve girdilerin arasından çıkarıldı. Verilerin %30’u rastgele olarak test grubu olarak ayrıldı. Eğitim verileri ölçeklendirildi ve bu ölçeklendirmeye göre test verileri ölçeklendirildi. Böylece veriler analiz etmeye hazır getirildi.

5.3 Bulgular

Analiz etmeye hazır hale getirilen verilere lojistik regresyon, k en yakın komşu, destek vektör makineleri, rastgele orman, gradyan arttırma ve yapay sinir ağları yöntemleri uygulandı. Daha sonra bu yöntemlerdeki belirlenen parametreler değiştirilerek en iyi parametreler bulunmaya çalışıldı. Elde edilen tahminler doğruluk oranı, kesinlik skoru, duyarlılık skoru ve F1 skoru ile değerlendirildi. F1 skoru en yüksek olan parametreler en iyi parametre kabul edildi. Ayrıca aşırı öğrenmeyi kontrol etmek için eğitim verileriyle tahminler arasında doğruluk oranı da hesaplandı. Aşağıda her yöntemin en iyi F1 skoru olan başlıca parametreleri verilmiştir.

5.3.1 Lojistik regresyon

Çizelge 5.2 : Lojistik regresyon metrikleri.

doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
0,8088	0,8291	1	0,5	0,6667

Lojistik regresyon yönteminde herhangi bir parametreyle oynanmamıştır. Kesinlik ve duyarlılık skorlarında da anlaşılabileceği üzere lojistik regresyon yanlış bir şekilde öğrenmiştir. Birincil kökeni akciğer olmayan tümörlerin hepsini doğru tahmin etmiştir. Fakat birincil akciğer gözlemlerinin yarısını yanlış tahmin etmiştir. Lojistik regresyonu yönteminin parametreleriyle bulunan metrikler Çizelge 5.2'deki gibidir.

5.3.2 KNN

KNN yönteminde k değeri 1-29 arasındaki değerler ile tek tek denendi. Ağırlık olarak tekdüze veya uzaklık ağırlıkları denendi. Uzaklık yönteminde minkowski kullanıldı ve p değeri 1-9 arasında farklı değerle denendi. Elde edilen veriler F1 skoruna göre karşılaştırıldı ve aşırı öğrenme olan parametreler dikkate alınmadı. Ayrıca Ek-B'de hesaplanan diğer parametrelerin olduğu tablo verilmiştir.

Çizelge 5.3 : KNN yöntemi parametreleri ve metrikleri.

k değeri	ağırlık	p değeri	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
5	tekdüze	4	0,8529	0,7722	0,9	0,6923	0,7826
5	tekdüze	5	0,8529	0,7658	0,9	0,6923	0,7826
5	tekdüze	6	0,8529	0,7595	0,9	0,6923	0,7826
5	tekdüze	2	0,8382	0,7848	0,8571	0,6923	0,766
3	tekdüze	2	0,8235	0,8481	0,7917	0,7308	0,76

KNN yöntemi ile bulunan başlıca parametreler Çizelge 5.2'deki gibidir. Her ne kadar F1 değeri diğer parametrelere göre daha düşük olsa da $k=3$, $p=2$ parametreleri ile daha tutarlı metrikler elde edilmiştir. Diğer parametrelerde eğitim verilerindeki doğruluk oranı test verilerine göre daha düşük çıkması eksik öğrenme ihtimali olabileceğini göstermektedir. Bu sebeple %82,35 doğruluk oranına sahip $k=3$, $p=2$, ağırlık = tekdüze parametreleri KNN için en iyi parametrelerdir.

5.3.3 Destek vektör makineleri

DVM yönteminde çekirdek parametresiyle oynandı. Doğrusal, polinom, radyal tabanlı fonksiyon ve sigmoid çekirdekleri denendi. Ayrıca C katsayısı 10^{-3} ile 10^3 arasında 10'un katları olacak şekilde değerlerle denendi. Polinom çekirdeği için polinom derecesi 1-4 arasında değerler ile denendi. Çıkan sonuçlar Çizelge 5.4'te gösterilmiştir. Ayrıca Ek-C'de hesaplanan diğer parametrelerin olduğu tablo verilmiştir.

Çizelge 5.4 : DVM yöntemi parametreleri ve metrikleri.

çekirdek	derece	C	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
polinom	4	100	0,8382	0,9114	0,8571	0,6923	0,7660
rtf	-	10	0,8382	0,9177	0,8947	0,6538	0,7556
polinom	1	100	0,8235	0,8418	0,9375	0,5769	0,7143
rtf	-	1	0,8235	0,7658	1	0,5385	0,70
polinom	4	1000	0,75	0,9620	0,6552	0,7308	0,6909

Çekirdek fonksiyonu 4. Dereceden polinom olan ve $C=100$ olan parametresi F1 skoru 0,77 ile en yüksek F1 skoruna sahip parametre olduğu Çizelge 5.4'te görülüyor. Doğruluk oranlarına bakılacak olursa %83,82 ile en yüksek doğruluk oranına sahip. Fakat eğitim verileri ile hesaplanan doğruluk oranı %91,14. Bu parametrelerle aşırı öğrenmiş olabilir. Birinci dereceden polinom ve $C=100$ olan parametre de değerlendirilebilir.

5.3.4 Rastgele orman

RO yönteminde kullanılacak karar ağacı sayısı (n_{KA}) 1-19 arasında denendi. Her bir karar ağacı için gini ve entropi ayırma endeksleri denenerek metrikler hesaplandı.

Çizelge 5.5 : RO parametreleri ve hesaplanan metrikleri.

n KA	ayırma end	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
13	entropi	0,8235	0,9937	0,85	0,6538	0,7391
18	gini	0,8088	0,9937	0,8095	0,6538	0,7234
18	entropi	0,7941	0,9937	0,80	0,6154	0,6957
10	gini	0,7794	0,9810	0,7391	0,6538	0,6939
5	gini	0,75	0,9684	0,6552	0,7308	0,6909
1	entropi	0,7794	0,8734	0,7619	0,6154	0,6809

RO’da kullanılan parametrelerden en iyi sonuç veren 6 parametre Çizelge 5.5’te gösterilmiştir. Her ne kadar diğer ilk beş parametrelerin F1 skorlarından daha düşük F1 skoruna sahip olsa da tek karar ağacı ve entropi ayırma endeksi en iyi parametre seçilmiştir. Diğer parametrelerde eğitim verileriyle hesaplanan doğruluk oranı çok yüksek olup test verileriyle hesaplanan doğruluk oranları ile arasında ciddi fark vardır. Bu ise aşırı öğrenme olduğunu gösterir. Ayrıca Ek-D’de hesaplanan diğer parametrelerin olduğu tablo verilmiştir.

5.3.5 Gradyan arttırma

GA’da arttırma yapılacak aşama sayısı 5-200 arasında 5’in katları olacak şekilde denendi. Kriter parametresi 0.1-1 arasında 0.05 arttırarak denendi. Sapma ve üstel kayıp fonksiyonları denendi.

Çizelge 5.6 : Gradyan arttırma parametreleri ve metrikleri.

aşama sayısı	kriter	kayıp fon	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
10	0,2	üstel	0,8235	0,9177	0,85	0,6538	0,7391
5	0,5	üstel	0,8088	0,9304	0,8421	0,6154	0,7111
15	0,2	üstel	0,7941	0,9304	0,7727	0,6538	0,7083

Gradyan arttırma metrikleri Çizelge 5.6’da gösterilmiştir. 10 aşamalı kriter değeri 0,2, kayıp fonksiyonu üstel parametreleri en iyi parametre olarak seçilmiştir. F1 skoru %73.91, doğruluk oranı %82,35’tir. Ayrıca Ek-E’de hesaplanan diğer parametrelerin olduğu tablo verilmiştir.

5.3.6 Yapay sinir ağları

YSA yönteminde 3 gizli katman kullanılmıştır. Katmanlarda kullanılacak nöron sayısı 8-48 arasında 8’in katları olacak şekilde denenmiştir. Aşırı öğrenmeyi engellemek için kullanılan dropout değerini 0.1-0.5 arasında 0.1 artacak şekilde denenmiştir. En

iyileştirici olarak RMSprop ve Adam eniyileştiricileri kullanılmıştır. Kaç tekrar yapılacağını belirleyen epoch büyüklüğü 100-400 arasında 100'ün katları olacak şekilde denenmiştir.

Çizelge 5.7 : YSA parametreler ve metrikleri.

nöron sayısı	dropu t	en iyileştirici	epoch	doğruluk_test	doğruluk_eğ_t	kesinlik	duyarlılık	F1
32	0,2	RMSprop	100	0,8824	0,8797	0,9091	0,7692	0,8333
16	0,3	RMSprop	200	0,8824	0,8797	0,9091	0,7692	0,8333
24	0,3	adam	100	0,8824	0,8544	0,9091	0,7692	0,8333
16	0,2	RMSprop	200	0,8676	0,8924	0,8696	0,7692	0,8163
40	0,2	RMSprop	200	0,8676	0,9557	0,8696	0,7692	0,8163
40	0,3	RMSprop	200	0,8676	0,9430	0,8696	0,7692	0,8163
40	0,1	adam	100	0,8529	0,9304	0,8333	0,7692	0,8
48	0,4	adam	300	0,8676	0,9747	0,9474	0,6923	0,8
40	0,4	adam	200	0,8529	0,9494	0,8636	0,7308	0,7917

ANN yönteminde kullanılan parametrelerin metrikleri Çizelge 5.7’de verilmiştir. Nöron sayısı 16, dropout değeri 0,2, en iyileştirici RMSprop ve epoch büyüklüğü 100 olan model ile nöron sayısı 16, dropout değeri 0,3, en iyileştirici RMSprop, epoch büyüklüğü 200 olan model 0,83 F1 skoru değeri ile en iyi parametrelere sahip olan model olarak seçilmiştir. Ayrıca Ek-F’de hesaplanan diğer parametrelerin olduğu tablo verilmiştir.

6. SONUÇ VE ÖNERİLER

Bu tezde akciğerlerinde 10 mm'den büyük çapta tümör olan 100 hastanın BT görüntüleri 6 farklı yapay öğrenme yöntemiyle analiz edilmiştir. Ve her yöntemin parametreleriyle oynayarak en iyi parametreye sahip model bulunmaya çalışılmıştır. Sırasıyla lojistik regresyon, k en yakın komşu, destek vektör makineleri, rastgele orman, gradyan artırma ve yapay sinir ağları yöntemlerinin en iyi parametreleri çizelge 6.1'de gösterilmiştir. Bu değerlere göre yapay sinir ağları yöntemi en iyi sonuç veren yöntem olarak seçilmiştir.

Çizelge 6.1 : Yöntemlerin değerlendirme metrikleri.

metot	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
LR	0,8088	0,8291	1	0,5	0,6667
KNN	0,8235	0,8481	0,7917	0,7308	0,76
DVM	0,8382	0,9114	0,8571	0,6923	0,766
RO	0,7794	0,8734	0,7619	0,6154	0,6809
GA	0,8235	0,9177	0,85	0,6538	0,7391
YSA	0,8824	0,8797	0,9091	0,7692	0,8333

%88 doğruluk oranına sahip, %83 F1 skoruna sahip bu yöntem ile hastaların akciğerlerindeki tümörlerin metastaz mı yoksa kökeni akciğer olan tümör mü yüksek doğruluk oranıyla tahmin edebilecek.

Çalışmada kullanılan verilerin sınırlı olması yapılan analizi olumsuz yönde etkilemiştir. Verilerin az olması sebebiyle her bir tümörün farklı bir girdi olarak alınması, aynı hastadaki birden fazla tümörün birbirleri arasındaki konum bilgisi yok sayılmıştır. Yine veri setinin yeterince büyük olmaması sebebiyle çapraz doğrulama yapılmamış sadece parametreler değiştirilerek kontrol edilmiştir. Ayrıca çekilen akciğer BT'lerde kesit kalınlığı ve kontrastlı/kontrastsız olması konusunda bir standart oluşursa analizler için kullanılacak verilerin sayısı artacaktır.

Hasta BT görüntüleri analizindeki en büyük engel tümörlerin segmente edilmesi. İşlemin uzun sürmesi, uzmanlık gerektirmesi ve bilgisayarlarda gerekli programın yüklü olması gözlemlenen nedenler arasındadır. Segmentasyon sürecini verim/zaman

olarak en yüksek olduđu tecr be edinilen 3 Slicer Python temelli a ık kaynak kodlu tıbbi g r nt leme yardımcı programı uzman doktorların bilgisayarlarına y klenmeli ve gerekli eđitimler verilmeli. B ylece uzmanlar segmentasyonu daha hızlı yapabilir ve analiz i in gerekli veri sayısı arttırılabilir.

Bundan sonraki  alıřmalarda, konvansiyonel yapay sinir ađları kullanarak segmente edilmiř t m r g r nt leri ile  zellik  ıkarımı yapılmadan sınıflandırma yapılabilir. Ayrıca segmentasyon yapılmadan t m rl  akciđer BT g r nt leriyle de konvansiyonel sinir ađları kullanarak sınıflandırma yapılabilir. Daha sonra bu    farklı sınıflandırma y nteminden elde edilen sonu lar karřılařtırılarak en iyi sonucu veren y ntem belirlenebilir.

Akciđerde bulunan t m rl  b lgenin metastaz mı deđil mi sınıflandırmasından sonra metastaz olan t m rlerin birincil k keninin hangi organ olduđu konusunda bir  alıřma yapılabilir. Bu  alıřma ikili sınıflandırmadan  ıkıp  ok sınıflı sınıflandırma olacađından daha fazla g zlem verisine ihtiya  duyulacaktır.

 n hazırlık s re lerinin ve analizlerin yapıldıđı Python kodları Ek-A'da verilmiřtir.

KAYNAKLAR

- Adepoju, O.** (2019). A step-by-step tutorial on coding Neural Network Logistic Regression model from scratch. 12 Haziran 2020 tarihinde <https://link.medium.com/wFwPY8olf7> adresinden erişildi.
- Akçay, M. Ç.** (2014). *Radyoterapi Uygulanan Akciğer Kanseri Tanılı Vakalarda Radyasyon Pnömonisi Gelişimini Etkileyen Faktörler*. (Doktora tezi). Eskişehir Osman Gazi Üniversitesi Tıp Fakültesi, Eskişehir.
- Aktaş, R., Doğanay M. M. ve Yıldız B.** (2003). Mali Başarısızlığın Öngörülmesi: İstatistiksel Yöntemler Ve Yapay Sinir Ağı Karşılaştırması, *IJCSI International Journal of Computer Science Issues*, 58 (4), 1-24.
- Ali, J., Khan, R., Ahmad, N. ve Maqsood, I.** (2012). Random Forests and Decision Trees, *Ankara Üniversitesi SBF Dergisi*, 9 (3), 272-278.
- Alpar, R.** (2003). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş*. Nobel Akademik Yayıncılık.
- Alpaydın, E.** (2018). *Foundations of Machine Learning* (2. bs.). London, The MIT Press.
- Alpaydın, E.** (2018). *Yapay Öğrenme* (4. bs.). İstanbul, Boğaziçi Üniversitesi Yayınevi.
- Ancalia, P. P. & Roy K.** (2014). The k-Nearest Neighbor Algorithm Using MapReduce Paradigm, *2014 Fifth International Conference on Intelligent Systems, Modelling and Simulation*, 15 (3), 513-518.
- Bradley, A. P., Duin, R. P. W., Paclik, P. ve Landgrebe T. C. W.** (2006). Precision-recall operating characteristic (P-ROC) curves in imprecise environments, *18th International Conference on Pattern Recognition*, 4, 123-127.
- Brujine, M.** (2016). Machine learning approaches in medical image analysis: from detection to diagnosis, *Medical Image Analysis journal*, 33, 94-97.
- Chen, E., Chung, P., Chen, C. ve Chang, C.** (1998). An automatic diagnostic system for CT liver image classification, *IEEE Transactions on Biomedical Engineering*, 45 (6), 783-794.
- Chiang, A. C. & Massague, J. M.** (2008). Molecular Basis of Metastasis, *The New England Journal of Medicine*, 359 (26), 2814-2823.
- Creyer, A. M. & Thorley, A. J.** (2019). Nanotechnology in the diagnosis and treatment of lung cancer, *Pharmacology & Therapeutics*, 198, 189-205.
- Çakır Zeytinoğlu, F.** (2007). İşletmelerin Dönen Varlıklarının Satışlar Üzerindeki Etkileri: En İyi Regresyon Denkleminin Seçimi ve Sektörel Karşılaştırma, *Marmara Üniversitesi İ.İ.B.F. Dergisi*, 22 (2), 331-349.

- Çevikbaş, G.** (2016). *Akciğer Tümörlerinde Yoğunluk Ayarlı Ark Radyoterapi Tekniği Kullanılarak İki Farklı Optimizasyon Algoritmasının Doz Dağılımına Etkisinin İncelenmesi*. (Yüksek lisans tezi). İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, İstanbul.
- Deo, R. C.** (2015). Machine Learning in Medicine, *Basic Science for Clinicians*, 132, 1920-1930.
- Dharani, T. & Aroquiaraj I. L.** (2013). Content Based Image Retrieval System using Feature Classification with Modified KNN Algorithm. 12 Ekim 2020 tarihinde <https://arxiv.org/ftp/arxiv/papers/1307/1307.4717.pdf> adresinden erişildi.
- Dilmaç, S.** (2014). *Metastatik Ve Non Metastatik Tümörlerde Nefronektin Ekspresyonu*. (Yüksek lisans tezi). Akdeniz Üniversitesi Sağlık Bilimleri Enstitüsü Histoloji ve Embriyoloji Anabilim Dalı, Antalya.
- Doğan, O.** (2019). *Kentsel Dokunun, Destek Vektör Makineleri Yöntemi ile Sınıflandırılması: İstanbul Örneği*. (Yüksek lisans tezi). Sivas Cumhuriyet Üniversitesi Fen Bilimleri Enstitüsü, Sivas.
- Du, W. & Zhan Z.** (2002). Building Decision Tree Classifier on Private Data, *Electrical Engineering and Computer Science*, 8, s.y.
- El-Habil, A.** (2012). An Application on Multinomial Logistic Regression Model, *Pakistan Journal of Statistics and Operation Research*, 8 (2), 271-291.
- Erickson, B. J., Korfiatis, P., Akkus, Z. ve Kline, T. L.** (2017). Machine Learning for Medical Imaging, *Radiographics*, 37 (2), 505-515.
- Fidler, I. J.** (2003). Mammary gland biology and breast cancer, *Conference on Common Molecular Mechanisms of Mammary Gland Development and Breast Cancer Progression*, 7 (11), 1084-1088.
- Fidler, I. J.** (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited, *Nature Journals*, 3, 453-458.
- Friedman, J. H.** (2002). Stochastic Gradient Boosting, *Computational Statistics & Data Analysis*, 38 (4), 367-378.
- Geurts, P., Irtuthum, A. ve Wehenkel, L.** (2009). Supervised Learning With Decision Tree-Based Methods in Computational And Systems Biology, *Mol Biosyst*, 5 (12), 1593-1605.
- Grover, P.** (2017). Gradient Boosting from scratch. 16 Ekim 2020 tarihinde <https://medium.com/mlreview/gradient-boosting-from-scratch-1e317ae4587d> adresinden erişildi.
- Hanahan, D. & Weinberg, R. A.** (2011). Hallmarks of Cancer: The Next Generation, *Cell Cancer Journal*, 144 (5), 646-674.
- Hirakata, K., Nakata H. ve Nakagawa, T.** (1995). CT of Pulmonary Metastases With Pathological Correlation, *Seminars in Ultrasound, CT, and MRI*, 16 (5), 379-394.
- Hosmer, D. & Lameshow, S.** (2000). *Applied Logistic Regression*. New York, A Wiley-Interscience Yayınevi.

- Huang, X., Shi, L. ve Suykens, J. A. K.** (2014). Support Vector Machine Classifier with Pinball Loss, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36 (5), 984-997
- Jain, A. K., Mao, J. ve Mohiuddin, K. M.** (1996). Artificial neural networks: A tutorial, *IEEE Journals & Magazine*, 29 (3), 134-140.
- Johan, A. W. S. B., Utaminingrum, F. ve Budi, A. S.** (2020). K-Value Effect Based on Combination GLCM Angle and KNN for Detecting Smart Wheelchair, *Journal of Information Technology and Computer Science*, 5 (1), 23-31.
- Junker, M., Hoch, R. ve Dengel, A.** (1999). On the evaluation of document analysis components by recall, precision, and accuracy, *Proc. of 5th International Conference on Document Analysis and Recognition*, 29 (3), 713-716.
- Kaya, G.** (2012). *Akciğer Kanseri Tedavisinde Uygulanan Konformal Ve Yoğunluk Ayarlı Radyoterapi Tekniklerinin Karşılaştırılmalı Değerlendirilmesi*. (Yüksek lisans tezi). Ankara Üniversitesi Nükleer Bilimleri Enstitüsü, Ankara.
- Kaya, G.** (2019). *Konvolüsyonel Sinir Ağları Kullanılarak Çiçek Türlerinin Sınıflandırılması*. (Yüksek lisans tezi). Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Kim, J., Kim, B. S. ve Savarese, S.** (2011). The role of the microenvironment in tumor growth and invasion, *Progress in Biophysics & Molecular Biology*, 106 (2), 133-138.
- Kim, Y., Stolarska, M. A. ve Othmer, H. G.** (2012). Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines, *Applied Mathematics in Electrical and Computer Engineering*, 106 (2), 353-379.
- Kocak, B., Durmaz E. Ş., Ateş E. ve Kılıçkesmez, Ö.** (2019). Radiomics with artificial intelligence: a practical guide for beginners, *Diagn Interv Radiol* 2019, 25, 485–495.
- Kourou, K., Exarchos T. P., Exarchos K.P., Karamousiz M. V. ve Fotiadis, D. I.** (2014). Machine Learning applications in cancer prohnosis and prediction, *Computational and Structural Biotechnology Journal*, 13, 8-17.
- Kurama, V.** (2020). Gradient Boosting In Classification: Not a Black Box Anymore!. 16 Ekim 2020 tarihinde <https://blog.paperspace.com/gradient-boosting-for-classification/> adresinden erişildi.
- Larroza, A., Moratal, D., Paredes-Sanches A., Soria-Olivas E., Chust, M. L., Arribas, L. A. ve Arana E.** (2015). Support Vector Machine Classification of Brain Metastasis and Radiation Necrosis Based on Texture Analysis in MRI, *J Magn Reson Imaging*, 42 (5), 1362-1368.
- Le, J.** (2018). Logistic Regression in R Tutorial. 10 Haziran 2020 tarihinde <https://www.datacamp.com/community/tutorials/logistic-regression-R> adresinden erişildi.

- Lin, J. W.** (2017). Artificial Neural Network Related to Biological Neuron Network: A Review, *Advanced Studies in Medical Sciences*, 5 (1), 55-62.
- Mathur, A. & Foody, G. M.** (2008). Multiclass and Binary SVM Classification: Implications for Training and Classification Users, *IEEE Geoscience And Remote Sensing Letter*, 5 (2), 241-245.
- Mukherjee, A., Paul, M. ve Mukherjee, S.** (2019). Recent Progress in the Theranostics Application of Nanomedicine in Lung Cancer, *Publisher of Open Access Journal*, 11 (5), 597-614.
- Müsellim, B.** (2007). Akciğer Kanserinin Epidemiyolojisi ve Etyolojisi, *İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Sürekli Tıp Eğitimi Sempozyum Dizisi*, 58, 113-118.
- Myrak, H.** (2017). GLCM Texture: A Tutorial v. 3.0 March 2017. 12 Ekim 2020 tarihinde <https://prism.ucalgary.ca/handle/1880/51900> adresinden erişildi.
- Nasim, F., Sabath, B. F. ve Eapen, G. A.** (2019). Lung Cancer, *The Medical Clinics of North America*, 103 (3), 463-473.
- Ooi, H. L., Ng, S. C. ve Lim, E.** (2013). ANO Detection with K-Nearest Neighbor Using Minkowski Distance, *International Journal of Signal Processing Systems*, 1 (2), 208-211.
- Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O. ve Akinjobi, J.** (2017). Supervised Machine Learning Algorithms: Classification and Comparison, *International Journal of Computer Trends and Technology (IJCTT)*, 48 (3), 128-138.
- Öğündür, G.** (2019). Doğruluk (Accuracy) , Kesinlik(Precision) , Duyarlılık(Recall) ya da F1 Score ?. 16 Ekim 2020 tarihinde <https://medium.com/@gulcanogundur/do%C4%9Fruluk-accuracy-kesinlik-precision-duyarl%C4%B1l%C4%B1k-recall-ya-da-f1-score-300c925feb38> adresinden erişildi.
- Öztemel, E.** (1996). Bilgisayarda Öğrenme ve Yapay Sinir Ağları, *Otomasyon Dergisi*, 134-140.
- Pal, M. & Foody, G. M., 2010.** Feature selection for classification of hyperspectral data by SVM, *IEEE Transactions on Geoscience and Remote Sensing*, 48, 2297-2307.
- Pedrogosa, F., Varoquax, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E.** (2011). Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, 12, 2825-2830.
- Ponder, B. A. J.** (2001). Cancer Genetics, *Macmillan*, 411, 336-341.
- Poste, G. & Fidler, I. J.** (1980). The pathogenesis of cancer metastasis, *Nature Journals*, 283, 139-146.
- Ramteke, R. J. & Khacane, M. Y.** (2012). Automatic Medical Image Classification and Abnormality Detection Using KNearest Neighbour, *International Journal of Advanced Computer Research*, 2 (4), 190-196.
- Ruppert, D. & Matteson D. S.** (2018). *Statistics and Data Analysis for Financial Engineering with R examples* (2. bs.). New York, Springer.

- Shaikhina, T., Lowe, D., Daga, S., Higgins, R. ve Khovanova, N.** (2019). Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation, *Biomedical Signal Processing and Control*, 52, 456-462.
- Soylu, K.** (2018). *Kredi Kartı Sahte İşlem Tespiti*. (Yüksek lisans tezi). Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Stella, G. M., Kolling, S., Benvenuti, S. ve Bortolotto, C.** (2019). Lung-Seeking Metastases, *Cancers*, 11 (7), s.y.
- Şahiner, S.** (2000). En Küçük Kareler Yöntemi ile Doğrusal Regresyon Modeli Oluşturmanın Temel Prensipleri, *MKÜ Ziraat Fakültesi Dergisi*, 5 (1-2), 57-73.
- Tang, T. C. & Chi, L. C.** (2005). Neural networks analysis in business failure prediction of chinese importers: a between-countries approach, *Expert Systems with Applications*, 29 (2), 244-255.
- Vishwanathan, S. V. N. & Murty M. N.** (2002). SSVM : A Simple SVM Algorithm, *Proceedings of the 2002 International Joint Conference on Neural Networks*, 3, 2393-2398
- Wald, N. J., Nanchahal, K., Thompson S. G. ve Cuckle, H.** (1986). Does breathing other people's tobacco smoke cause lung cancer?, *British Medical Journal*, 293, 1217-1222.
- Wang, S. & Summers, R. M.** (2012). Machine learning and radiology, *Medical Image Analysis*. 16 (5), 55-62.



EKLER

EK A: Python Kodu

```
import numpy as np

from os import listdir

import nrrd

import pandas as pd

from skimage.feature import greycomatrix, greycoprops

from sklearn.neighbors import KNeighborsClassifier

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler

from sklearn.metrics import accuracy_score, confusion_matrix, recall_score,
precision_score, f1_score

from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier

from sklearn.linear_model import LogisticRegression

from sklearn.decomposition import PCA

from sklearn.svm import SVC

from keras.models import Sequential

from keras.layers import Dense, Dropout


def normalize(img):

    x=(img-np.min(img))/(np.max(img)-np.min(img))

    return np.transpose((x*255).astype(np.uint8), (2,1,0))


def bolge_df(e,deger):
```

```

path= 'D:/Metastaz/Slicer/'
bolgeler= listdir(path)
bolge_hastalar=listdir(path+bolgeler[e])
bolge_hastalar_path=[listdir(path+bolgeler[e]+'/' +i) for i in bolge_hastalar]
bolge_tam_resim=[bolge_hastalar_path[i][0] for i in
range(len(bolge_hastalar_path))]
bolge_boyali=[]
for i in range(len(bolge_hastalar_path)):
    x=len(bolge_hastalar_path[i])
    y=int((x+1)/2)
    z=bolge_hastalar_path[i][y:]
    bolge_boyali.append(z)
bolge_butun_resim=[]
bolge_butun_bilgi=[]
for i in range(len(bolge_tam_resim)):
    x, y=nrrd.read(path+bolgeler[e]+'/' +bolge_hastalar[i]+'/' +bolge_tam_resim[i])
    x=normalize(x)
    bolge_butun_resim.append(x)
    bolge_butun_bilgi.append(y)

bolge_isaretli_resim=[]
for i in range(len(bolge_boyali)):
    bos=[]
    for j in range(len(bolge_boyali[i])):
        x, y=nrrd.read(path+bolgeler[e]+'/' +bolge_hastalar[i]+'/' +bolge_boyali[i][j])
        x=normalize(x)

```

```

        bos.append(x)

    bolge_isaretli_resim.append(bos)


bolge_kesilmis_resim_grup=[]

for i in range(len(bolge_butun_resim)):

    bos=[]

    for j in range(len(bolge_isaretli_resim[i])):

        x=np.copy(bolge_butun_resim[i])

        x[np.where(bolge_isaretli_resim[i][j]==0)]=0

        bos.append(x)

    bolge_kesilmis_resim_grup.append(bos)


bolge_kesilmis_resim=[]

bolge_nodul_sayisi=[]

for i in range(len(bolge_kesilmis_resim_grup)):

    for j in range(len(bolge_kesilmis_resim_grup[i])):

        x= bolge_kesilmis_resim_grup[i][j]

        bolge_kesilmis_resim.append(x)

        bolge_nodul_sayisi.append(len(bolge_kesilmis_resim_grup[i]))


bolge_hedef_slice_ort=[]

bolge_kesitler=[]

for i in range(len(bolge_kesilmis_resim)):

    bos=[]

```

```

x=np.zeros((bolge_kesilmis_resim[i].shape[1],bolge_kesilmis_resim[i].shape[2])).ast
ype(np.uint8)

    for j in range(len(bolge_kesilmis_resim[i])):

        if np.any(x!=bolge_kesilmis_resim[i][j]):

            if len(np.where(bolge_kesilmis_resim[i][j]!=0)[0])>5:

                bos.append(j)

    bolge_kesitler.append(bos)

    mean=np.mean(bos)

    bolge_hedef_slice_ort.append(int(mean))

bolge_tumor_array=[]
for n,i in enumerate(bolge_kesitler):

    x = np.array([np.array(bolge_kesilmis_resim[n][j]) for j in i])

    bolge_tumor_array.append(x)

bolge_tumor_ort = [np.average(i,axis=0) for i in bolge_tumor_array]

bolge_tumor_sapma = [np.std(i,axis=0) for i in bolge_tumor_array]

bolge_tumor_ort = [i.astype(np.uint8) for i in bolge_tumor_ort]

bolge_tumor_sapma = [i.astype(np.uint8) for i in bolge_tumor_sapma]

bolge_tumor_alan=[]

bolge_x_cap=[]

bolge_y_cap=[]

for i in range(len(bolge_kesilmis_resim)):

    bosx=[]

    bosy=[]

    x,y=np.where(bolge_kesilmis_resim[i][bolge_hedef_slice_ort[i]]!=0)

    x_cap=np.max(x)-np.min(x)

```

```

y_cap=np.max(y)-np.min(y)

bosx.append([np.min(x),np.max(x)])

bosy.append([np.min(y),np.max(y)])

bolge_tumor_alan.append([bosx,bosy])

bolge_x_cap.append(x_cap)

bolge_y_cap.append(y_cap)

bolge_haar=[]

for i in range(len(bolge_kesilmis_resim)):

    grayco_prop_list = ['contrast', 'dissimilarity', 'homogeneity', 'energy',
'correlation', 'ASM']

    glcm_buyuk =
greycomatrix(bolge_kesilmis_resim[i][bolge_hedef_slice_ort[i]][bolge_tumor_alan[i]
][0][0][0]:bolge_tumor_alan[i][0][0][1],bolge_tumor_alan[i][1][0][0]:bolge_tumor_a
lan[i][1][0][1]],

[3], [0], 256, symmetric=True, normed=True)

    glcm_ort =
greycomatrix(bolge_tumor_ort[i][bolge_tumor_alan[i][0][0][0]:bolge_tumor_alan[i][
0][0][1],bolge_tumor_alan[i][1][0][0]:bolge_tumor_alan[i][1][0][1]],

[3], [0], 256, symmetric=True, normed=True)

    glcm_sapma =
greycomatrix(bolge_tumor_sapma[i][bolge_tumor_alan[i][0][0][0]:bolge_tumor_ala
n[i][0][0][1],bolge_tumor_alan[i][1][0][0]:bolge_tumor_alan[i][1][0][1]],

[3], [0], 256, symmetric=True, normed=True)

    x=np.copy(bolge_kesilmis_resim[i][bolge_hedef_slice_ort[i]])

    out_row = dict(

buyuk_ort=np.mean(bolge_kesilmis_resim[i][bolge_hedef_slice_ort[i]][np.where(x>
0)]),buyuk_std=np.std(bolge_kesilmis_resim[i][bolge_hedef_slice_ort[i]][np.where(
x>0)]),

```

```
ortalama_ort=np.mean(bolge_tumor_ort[i][np.where(x>0)]),ortalama_std=np.std(bolge_tumor_ort[i][np.where(x>0)]),
```

```
sapma_ort=np.mean(bolge_tumor_sapma[i][np.where(x>0)]),sapma_std=np.std(bolge_tumor_sapma[i][np.where(x>0)]),
```

```
label=deger,x_cap=bolge_x_cap[i],y_cap=bolge_y_cap[i])
```

```
",nodul_sayisi=bolge_nodul_sayisi[i]"
```

```
for c_prop in grayco_prop_list:
```

```
    out_row[c_prop+'_buyuk'] = greycoprops(glc_m_buyuk, c_prop)[0, 0]
```

```
    out_row[c_prop+'_ort'] = greycoprops(glc_m_ort, c_prop)[0, 0]
```

```
    out_row[c_prop+'_sapma'] = greycoprops(glc_m_sapma, c_prop)[0, 0]
```

```
    bolge_haar += [out_row]
```

```
bolge_df=pd.DataFrame(index=out_row.keys())
```

```
for i in range(len(bolge_haar)):
```

```
    x=pd.DataFrame.from_dict(bolge_haar[i],orient='index')
```

```
    bolge_df=pd.concat([x,bolge_df],axis=1)
```

```
bolge_df=bolge_df.T.reset_index(drop=True)
```

```
return bolge_df
```

```
###
```

```
import time
```

```
start_time=time.time()
```

```

bobrek_df=bolge_df(0,0)

kolon_df=bolge_df(1,0)

meme_df=bolge_df(2,0)

akciger_df=bolge_df(3,1)


print((time.time()-start_time))

#%%%


threshold_otlier = -10

test_size = 0.30

random_state = 8


birlesmis_df =
pd.concat([bobrek_df,meme_df,kolon_df,akciger_df],axis=0).reset_index(drop=True
)

etiket=birlesmis_df.label

son_df=birlesmis_df.drop("label",axis=1)

sutunlar = son_df.columns.tolist()


clf = LocalOutlierFactor()

y_pred = clf.fit_predict(son_df)

outlier_score = pd.DataFrame()

outlier_score["score"] = clf.negative_outlier_factor_


outlier_index = outlier_score[outlier_score["score"] < threshold_otlier].index.tolist()

x=son_df.drop(outlier_index)

y=etiket.drop(outlier_index)

```

```
x_train , x_test, y_train, y_test = train_test_split(x, y, test_size=test_size,
random_state=random_state )
```

```
scaler = StandardScaler()
```

```
x_train = scaler.fit_transform(x_train)
```

```
x_test = scaler.transform(x_test)
```

```
####
```

```
knn_sonuc=[]
```

```
for i in range(1,30):
```

```
    for j in ["uniform","distance"]:
```

```
        for w in range(1,10):
```

```
            knn = KNeighborsClassifier(n_neighbors=i, weights=j, p=w)
```

```
            knn.fit(x_train, y_train)
```

```
            acc_test = accuracy_score(y_test, knn.predict(x_test))
```

```
            acc_train = accuracy_score(y_train, knn.predict(x_train))
```

```
            x = ["knn", i, j, w, acc_test, acc_train,
```

```
precision_score(y_test,knn.predict(x_test)), recall_score(y_test,knn.predict(x_test)),
```

```
f1_score(y_test,knn.predict(x_test))]
```

```
            knn_sonuc.append(x)
```

```
knn_df=pd.DataFrame(knn_sonuc,columns=['method','n_neighbors','weights','p','acc
_test','acc_train','precision_score','recall_score','f1_score'])
```

```
knn_df.to_excel(r"D:\Masaüstü\knn5.xlsx", sheet_name='knn', index = True)
```

```
#%%%
```

```
rfc_sonuc=[]
```

```
for i in range(1,20):
```

```
    for j in ['entropy','gini']:
```

```
        rfc = RandomForestClassifier(n_estimators=i, criterion=j)
```

```
        rfc.fit(x_train, y_train)
```

```
        acc_test = accuracy_score(y_test, rfc.predict(x_test))
```

```
        acc_train = accuracy_score(y_train, rfc.predict(x_train))
```

```
        x = ["rfc", i, j, acc_test, acc_train, precision_score(y_test,rfc.predict(x_test)),  
recall_score(y_test,rfc.predict(x_test)), f1_score(y_test,rfc.predict(x_test))]
```

```
        rfc_sonuc.append(x)
```

```
rfc_df=pd.DataFrame(rfc_sonuc,columns=['method','n_estimators','criterion','acc_test',  
'acc_train','precision_score','recall_score','f1_score'])
```

```
rfc_df.to_excel(r"D:\Masaüstü\rfc.xlsx", sheet_name='rfc', index = True)
```

```
#%%%
```

```
lr_sonuc=[]
```

```
for j in ['l1', 'l2']:
```

```
    for i in ['newton-cg', 'lbfgs', 'liblinear', 'sag', 'saga']:
```

```
        for w in [0.001,0.01,0.1,1,10,100,1000]:
```

```
            try:
```

```

lr = LogisticRegression(penalty=j,solver=i,C=w)

lr.fit(x_train, y_train)

acc_test = accuracy_score(y_test, lr.predict(x_test))

acc_train = accuracy_score(y_train, lr.predict(x_train))

x = ["lr", j,i,w, acc_test, acc_train, precision_score(y_test,
lr.predict(x_test)), recall_score(y_test,lr.predict(x_test)),
f1_score(y_test,lr.predict(x_test))]

lr_sonuc.append(x)

except:

    pass

lr_df=pd.DataFrame(lr_sonuc,columns=['method','penalty','solver','C',
'acc_test','acc_train','precision_score','recall_score','f1_score'])

lr_df.to_excel(r"D:\Masaüstü\lr3.xlsx", sheet_name='lr', index = True)

#%%

svm_sonuc =[]

for i in ["scale","auto"]:

    for j in ['linear', 'poly', 'rbf', 'sigmoid']:

        for z in range(1,10):

            for w in [0.001,0.01,0.1,1,10,100,1000]:

                svm = SVC(gamma=i, kernel=j, degree=z, C=w)

                svm.fit(x_train,y_train)

                acc_test = accuracy_score(y_test, svm.predict(x_test))

                acc_train = accuracy_score(y_train, svm.predict(x_train))

```

```

        x = ["svm", i, j, z, w, acc_test, acc_train,
precision_score(y_test,svm.predict(x_test)), recall_score(y_test,svm.predict(x_test)),
f1_score(y_test,svm.predict(x_test))]

        svm_sonuc.append(x)

```

```

svm_df=pd.DataFrame(svm_sonuc,columns=['method','gamma','kernel','degree','C','a
cc_test','acc_train','precision_score','recall_score','f1_score'])

```

```

svm_df.to_excel(r"D:\Masaüstü\svm2.xlsx", sheet_name='svm', index = True)

```

```

#%%%

```

```

ann_sonuc=[]

```

```

for i in range(40,49,8):

```

```

    for j in np.arange(0,0.5,0.2):

```

```

        for k in ['adam','RMSprop']:

```

```

            for z in ['binary_crossentropy']:

```

```

                for w in range(200,401,100):

```

```

                    ann = Sequential()

```

```

                    ann.add(Dense(output_dim=i,init='uniform',
activation='relu',input_dim=len(x_train[1])))

```

```

                    ann.add(Dropout(p=j))

```

```

                    ann.add(Dense(output_dim=i,init='uniform', activation='relu'))

```

```

                    ann.add(Dropout(p=j))

```

```

ann.add(Dense(output_dim=i,init='uniform', activation='relu'))

ann.add(Dropout(p=j))


ann.add(Dense(output_dim=1, init='uniform', activation='sigmoid'))


ann.compile(optimizer=k, loss=z, metrics=['accuracy'])


ann.fit(x_train, y_train, batch_size=10, nb_epoch=w)

ara1=(ann.predict(x_test)>0.5)
acc_test = accuracy_score(y_test,ara1)
ara2=(ann.predict(x_train)>0.5)
acc_train = accuracy_score(y_train, ara2)

x = ["ann", i, j, z, k, w, acc_test, acc_train, precision_score(y_test,ara1),
recall_score(y_test,ara1), f1_score(y_test,ara1)]
ann_sonuc.append(x)


ann_df=pd.DataFrame(ann_sonuc,columns=['method','output','dropout', 'loses',
'optimizer','epoch_size',
'acc_test','acc_train','precision_score','recall_score','f1_score'])


ann_df.to_excel(r"D:\Masaüstü\ann5.xlsx", sheet_name='ann', index = True)

#%%


gb_sonuc=[]

for i in range(5,200,5):

```

```

for j in np.arange(0.1,1,0.05):

    for k in ['deviance','exponential']:

        gb = GradientBoostingClassifier(learning_rate=j, n_estimators=i , loss=k)

        gb.fit(x_train, y_train)

        acc_test = accuracy_score(y_test, gb.predict(x_test))

        acc_train = accuracy_score(y_train, gb.predict(x_train))

        x = ["gb", i, j, k, acc_test, acc_train,
precision_score(y_test,gb.predict(x_test)), recall_score(y_test,gb.predict(x_test)),
f1_score(y_test,gb.predict(x_test))]

        gb_sonuc.append(x)

gb_df=pd.DataFrame(gb_sonuc,columns=['method','n_estimators','learning_rate',
'loss', 'acc_test','acc_train','precision_score','recall_score','f1_score'])

gb_df.to_excel(r"D:\Masaüstü\gb.xlsx", sheet_name='gb', index = True)

#%%

pca = PCA(0.95)

pca.fit(x_train)

x_train_pca = pca.transform(x_train)

x_test_pca = pca.transform(x_test)

#%%

knn = KNeighborsClassifier(n_neighbors=4, weights='distance',p=2)

```

```

knn.fit(x_train_pca, y_train)

acc_test = accuracy_score(y_test, knn.predict(x_test_pca))

acc_train = accuracy_score(y_train, knn.predict(x_train_pca))

x = ["knn_pca", '4', 'distance', '2', acc_test, acc_train,
precision_score(y_test, knn.predict(x_test_pca)),
recall_score(y_test, knn.predict(x_test_pca)),
f1_score(y_test, knn.predict(x_test_pca))]

knn_sonuc.append(x)


knn_df=pd.DataFrame(knn_sonuc,columns=['method','n_neighbors','weights','p','acc
_test','acc_train','precision_score','recall_score','f1_score'])

knn_df.to_excel(r"D:\Masaüstü\knn_pca2.xlsx", sheet_name='knn', index = True)

###


gb_sonuc=[]

for i in range(5,200,5):

    for j in np.arange(0.1,1,0.05):

        for k in ['deviance','exponential']:

            gb = GradientBoostingClassifier(learning_rate=j, n_estimators=i , loss=k)

            gb.fit(x_train_pca, y_train)

            acc_test = accuracy_score(y_test, gb.predict(x_test_pca))

            acc_train = accuracy_score(y_train, gb.predict(x_train_pca))

            x = ["gb_pca", i, j, k, acc_test, acc_train,
precision_score(y_test, gb.predict(x_test_pca)),
recall_score(y_test, gb.predict(x_test_pca)), f1_score(y_test, gb.predict(x_test_pca))]

            gb_sonuc.append(x)

```

```
gb_df=pd.DataFrame(gb_sonuc,columns=['method','n_estimators','learning_rate',
'loss','acc_test','acc_train','precision_score','recall_score','f1_score'])
```

```
gb_df.to_excel(r"D:\Masaüstü\gb_pca.xlsx", sheet_name='gb', index = True)
```

```
#%%%
```

```
rfc_sonuc=[]
```

```
for i in range(1,20):
```

```
    for j in ['entropy','gini']:
```

```
        rfc = RandomForestClassifier(n_estimators=i, criterion=j)
```

```
        rfc.fit(x_train_pca, y_train)
```

```
        acc_test = accuracy_score(y_test, rfc.predict(x_test_pca))
```

```
        acc_train = accuracy_score(y_train, rfc.predict(x_train_pca))
```

```
        x = ["rfc", i, j, acc_test, acc_train,
```

```
precision_score(y_test,rfc.predict(x_test_pca)),
```

```
recall_score(y_test,rfc.predict(x_test_pca)), f1_score(y_test,rfc.predict(x_test_pca))]
```

```
        rfc_sonuc.append(x)
```

```
rfc_df=pd.DataFrame(rfc_sonuc,columns=['method','n_estimators','criterion','acc_test',
',acc_train','precision_score','recall_score','f1_score'])
```

```
rfc_df.to_excel(r"D:\Masaüstü\rfc2.xlsx", sheet_name='rfc', index = True)
```

EK B: KNN Parametre Ayarlamaları (En iyi ilk 250 parametre)

	k değeri	ağırlık	p değeri	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
1	6	uzaklık	7	0,8824	1	0,95	0,7308	0,8261
2	6	uzaklık	3	0,8676	1	0,9048	0,7308	0,8085
3	6	uzaklık	6	0,8676	1	0,9474	0,6923	0,8
4	8	uzaklık	3	0,8676	1	0,9474	0,6923	0,8
5	4	uzaklık	2	0,8529	1	0,8636	0,7308	0,7917
6	5	tekdüze	4	0,8529	0,7722	0,9	0,6923	0,7826
7	5	tekdüze	5	0,8529	0,7658	0,9	0,6923	0,7826
8	5	tekdüze	6	0,8529	0,7595	0,9	0,6923	0,7826
9	5	uzaklık	4	0,8529	1	0,9	0,6923	0,7826
10	5	uzaklık	5	0,8529	1	0,9	0,6923	0,7826
11	5	uzaklık	6	0,8529	1	0,9	0,6923	0,7826
12	6	uzaklık	2	0,8529	1	0,9	0,6923	0,7826
13	6	uzaklık	4	0,8529	1	0,9	0,6923	0,7826
14	6	uzaklık	5	0,8529	1	0,9444	0,6538	0,7727
15	7	uzaklık	3	0,8529	1	0,9444	0,6538	0,7727
16	8	uzaklık	2	0,8529	1	0,9444	0,6538	0,7727
17	4	uzaklık	5	0,8382	1	0,8571	0,6923	0,766
18	4	uzaklık	6	0,8382	1	0,8571	0,6923	0,766
19	4	uzaklık	7	0,8382	1	0,8571	0,6923	0,766
20	5	tekdüze	2	0,8382	0,7848	0,8571	0,6923	0,766
21	5	uzaklık	2	0,8382	1	0,8571	0,6923	0,766
22	7	uzaklık	4	0,8529	1	1	0,6154	0,7619
23	7	uzaklık	5	0,8529	1	1	0,6154	0,7619
24	9	uzaklık	2	0,8529	1	1	0,6154	0,7619
25	1	tekdüze	7	0,8235	1	0,7917	0,7308	0,76
26	1	tekdüze	8	0,8235	1	0,7917	0,7308	0,76
27	1	tekdüze	9	0,8235	1	0,7917	0,7308	0,76
28	1	uzaklık	7	0,8235	1	0,7917	0,7308	0,76
29	1	uzaklık	8	0,8235	1	0,7917	0,7308	0,76
30	1	uzaklık	9	0,8235	1	0,7917	0,7308	0,76
31	2	uzaklık	7	0,8235	1	0,7917	0,7308	0,76
32	2	uzaklık	8	0,8235	1	0,7917	0,7308	0,76
33	2	uzaklık	9	0,8235	1	0,7917	0,7308	0,76
34	3	tekdüze	2	0,8235	0,8481	0,7917	0,7308	0,76
35	3	uzaklık	2	0,8235	1	0,7917	0,7308	0,76
36	4	uzaklık	3	0,8235	1	0,7917	0,7308	0,76
37	5	tekdüze	7	0,8382	0,7658	0,8947	0,6538	0,7556
38	5	uzaklık	7	0,8382	1	0,8947	0,6538	0,7556
39	7	tekdüze	3	0,8382	0,7975	0,8947	0,6538	0,7556
40	1	tekdüze	5	0,8088	1	0,7407	0,7692	0,7547
41	1	uzaklık	5	0,8088	1	0,7407	0,7692	0,7547

42	2	uzaklık	5	0,8088	1	0,7407	0,7692	0,7547
43	4	uzaklık	4	0,8235	1	0,8182	0,6923	0,75
44	1	tekdüze	6	0,8088	1	0,76	0,7308	0,7451
45	1	uzaklık	6	0,8088	1	0,76	0,7308	0,7451
46	2	uzaklık	6	0,8088	1	0,76	0,7308	0,7451
47	6	uzaklık	8	0,8382	1	0,9412	0,6154	0,7442
48	6	uzaklık	9	0,8382	1	0,9412	0,6154	0,7442
49	7	tekdüze	2	0,8382	0,7722	0,9412	0,6154	0,7442
50	7	tekdüze	4	0,8382	0,7848	0,9412	0,6154	0,7442
51	7	tekdüze	5	0,8382	0,7848	0,9412	0,6154	0,7442
52	7	uzaklık	2	0,8382	1	0,9412	0,6154	0,7442
53	8	uzaklık	1	0,8382	1	0,9412	0,6154	0,7442
54	8	uzaklık	4	0,8382	1	0,9412	0,6154	0,7442
55	8	uzaklık	5	0,8382	1	0,9412	0,6154	0,7442
56	1	tekdüze	4	0,7941	1	0,7143	0,7692	0,7407
57	1	uzaklık	4	0,7941	1	0,7143	0,7692	0,7407
58	2	uzaklık	4	0,7941	1	0,7143	0,7692	0,7407
59	4	uzaklık	8	0,8235	1	0,85	0,6538	0,7391
60	4	uzaklık	9	0,8235	1	0,85	0,6538	0,7391
61	6	uzaklık	1	0,8235	1	0,85	0,6538	0,7391
62	6	tekdüze	4	0,8382	0,7532	1	0,5769	0,7317
63	9	uzaklık	1	0,8382	1	1	0,5769	0,7317
64	10	uzaklık	1	0,8382	1	1	0,5769	0,7317
65	1	tekdüze	2	0,7794	1	0,6897	0,7692	0,7273
66	1	tekdüze	3	0,7794	1	0,6897	0,7692	0,7273
67	1	uzaklık	2	0,7794	1	0,6897	0,7692	0,7273
68	1	uzaklık	3	0,7794	1	0,6897	0,7692	0,7273
69	2	uzaklık	2	0,7794	1	0,6897	0,7692	0,7273
70	2	uzaklık	3	0,7794	1	0,6897	0,7692	0,7273
71	3	tekdüze	5	0,8088	0,8228	0,8095	0,6538	0,7234
72	3	tekdüze	6	0,8088	0,8228	0,8095	0,6538	0,7234
73	3	tekdüze	7	0,8088	0,8228	0,8095	0,6538	0,7234
74	3	tekdüze	8	0,8088	0,8165	0,8095	0,6538	0,7234
75	3	uzaklık	5	0,8088	1	0,8095	0,6538	0,7234
76	3	uzaklık	6	0,8088	1	0,8095	0,6538	0,7234
77	3	uzaklık	7	0,8088	1	0,8095	0,6538	0,7234
78	3	uzaklık	8	0,8088	1	0,8095	0,6538	0,7234
79	1	tekdüze	1	0,7794	1	0,7037	0,7308	0,717
80	1	uzaklık	1	0,7794	1	0,7037	0,7308	0,717
81	2	uzaklık	1	0,7794	1	0,7037	0,7308	0,717
82	2	tekdüze	3	0,8235	0,7911	0,9375	0,5769	0,7143
83	5	tekdüze	8	0,8235	0,7658	0,9375	0,5769	0,7143
84	5	tekdüze	9	0,8235	0,7722	0,9375	0,5769	0,7143
85	5	uzaklık	8	0,8235	1	0,9375	0,5769	0,7143
86	5	uzaklık	9	0,8235	1	0,9375	0,5769	0,7143

87	8	uzaklık	6	0,8235	1	0,9375	0,5769	0,7143
88	8	uzaklık	7	0,8235	1	0,9375	0,5769	0,7143
89	8	uzaklık	8	0,8235	1	0,9375	0,5769	0,7143
90	10	uzaklık	2	0,8235	1	0,9375	0,5769	0,7143
91	10	uzaklık	3	0,8235	1	0,9375	0,5769	0,7143
92	14	uzaklık	3	0,8235	1	0,9375	0,5769	0,7143
93	5	tekdüze	3	0,8088	0,7658	0,8421	0,6154	0,7111
94	5	uzaklık	3	0,8088	1	0,8421	0,6154	0,7111
95	7	tekdüze	1	0,8088	0,7785	0,8421	0,6154	0,7111
96	7	uzaklık	1	0,8088	1	0,8421	0,6154	0,7111
97	2	tekdüze	6	0,8235	0,7911	1	0,5385	0,7
98	6	tekdüze	1	0,8235	0,7468	1	0,5385	0,7
99	6	tekdüze	3	0,8235	0,7595	1	0,5385	0,7
100	7	tekdüze	7	0,8235	0,7848	1	0,5385	0,7
101	7	tekdüze	8	0,8235	0,7848	1	0,5385	0,7
102	7	uzaklık	6	0,8235	1	1	0,5385	0,7
103	7	uzaklık	7	0,8235	1	1	0,5385	0,7
104	7	uzaklık	8	0,8235	1	1	0,5385	0,7
105	8	tekdüze	2	0,8235	0,7342	1	0,5385	0,7
106	9	tekdüze	1	0,8235	0,7468	1	0,5385	0,7
107	9	tekdüze	2	0,8235	0,7278	1	0,5385	0,7
108	9	tekdüze	4	0,8235	0,7468	1	0,5385	0,7
109	9	uzaklık	4	0,8235	1	1	0,5385	0,7
110	9	uzaklık	7	0,8235	1	1	0,5385	0,7
111	9	uzaklık	8	0,8235	1	1	0,5385	0,7
112	10	uzaklık	7	0,8235	1	1	0,5385	0,7
113	10	uzaklık	8	0,8235	1	1	0,5385	0,7
114	10	uzaklık	9	0,8235	1	1	0,5385	0,7
115	11	tekdüze	2	0,8235	0,7405	1	0,5385	0,7
116	11	uzaklık	2	0,8235	1	1	0,5385	0,7
117	11	uzaklık	3	0,8235	1	1	0,5385	0,7
118	12	uzaklık	9	0,8235	1	1	0,5385	0,7
119	14	uzaklık	4	0,8235	1	1	0,5385	0,7
120	15	uzaklık	3	0,8235	1	1	0,5385	0,7
121	15	uzaklık	4	0,8235	1	1	0,5385	0,7
122	3	tekdüze	9	0,7941	0,8038	0,8	0,6154	0,6957
123	3	uzaklık	9	0,7941	1	0,8	0,6154	0,6957
124	5	tekdüze	1	0,7941	0,8038	0,8	0,6154	0,6957
125	5	uzaklık	1	0,7941	1	0,8	0,6154	0,6957
126	2	tekdüze	4	0,8088	0,7975	0,9333	0,5385	0,6829
127	2	tekdüze	5	0,8088	0,7911	0,9333	0,5385	0,6829
128	4	tekdüze	2	0,8088	0,7785	0,9333	0,5385	0,6829
129	7	tekdüze	6	0,8088	0,7848	0,9333	0,5385	0,6829
130	8	uzaklık	9	0,8088	1	0,9333	0,5385	0,6829
131	9	tekdüze	3	0,8088	0,7342	0,9333	0,5385	0,6829

132	9	uzaklık	3	0,8088	1	0,9333	0,5385	0,6829
133	12	uzaklık	2	0,8088	1	0,9333	0,5385	0,6829
134	3	tekdüze	4	0,7794	0,8228	0,7619	0,6154	0,6809
135	3	uzaklık	4	0,7794	1	0,7619	0,6154	0,6809
136	4	uzaklık	1	0,7647	1	0,7083	0,6538	0,68
137	3	tekdüze	3	0,7647	0,8291	0,7273	0,6154	0,6667
138	3	uzaklık	3	0,7647	1	0,7273	0,6154	0,6667
139	4	tekdüze	1	0,7941	0,8038	0,875	0,5385	0,6667
140	2	tekdüze	7	0,8088	0,7848	1	0,5	0,6667
141	2	tekdüze	8	0,8088	0,7911	1	0,5	0,6667
142	2	tekdüze	9	0,8088	0,7911	1	0,5	0,6667
143	6	tekdüze	2	0,8088	0,7532	1	0,5	0,6667
144	6	tekdüze	6	0,8088	0,7468	1	0,5	0,6667
145	7	tekdüze	9	0,8088	0,7658	1	0,5	0,6667
146	7	uzaklık	9	0,8088	1	1	0,5	0,6667
147	8	tekdüze	3	0,8088	0,7342	1	0,5	0,6667
148	9	tekdüze	7	0,8088	0,7595	1	0,5	0,6667
149	9	tekdüze	8	0,8088	0,7532	1	0,5	0,6667
150	9	uzaklık	5	0,8088	1	1	0,5	0,6667
151	9	uzaklık	6	0,8088	1	1	0,5	0,6667
152	9	uzaklık	9	0,8088	1	1	0,5	0,6667
153	10	tekdüze	2	0,8088	0,7215	1	0,5	0,6667
154	10	uzaklık	6	0,8088	1	1	0,5	0,6667
155	11	tekdüze	3	0,8088	0,7278	1	0,5	0,6667
156	11	tekdüze	7	0,8088	0,7089	1	0,5	0,6667
157	11	uzaklık	5	0,8088	1	1	0,5	0,6667
158	11	uzaklık	7	0,8088	1	1	0,5	0,6667
159	12	uzaklık	1	0,8088	1	1	0,5	0,6667
160	12	uzaklık	6	0,8088	1	1	0,5	0,6667
161	12	uzaklık	7	0,8088	1	1	0,5	0,6667
162	12	uzaklık	8	0,8088	1	1	0,5	0,6667
163	13	uzaklık	2	0,8088	1	1	0,5	0,6667
164	13	uzaklık	3	0,8088	1	1	0,5	0,6667
165	13	uzaklık	4	0,8088	1	1	0,5	0,6667
166	14	uzaklık	6	0,8088	1	1	0,5	0,6667
167	15	tekdüze	3	0,8088	0,7152	1	0,5	0,6667
168	15	uzaklık	2	0,8088	1	1	0,5	0,6667
169	16	uzaklık	2	0,8088	1	1	0,5	0,6667
170	16	uzaklık	3	0,8088	1	1	0,5	0,6667
171	17	uzaklık	3	0,8088	1	1	0,5	0,6667
172	3	tekdüze	1	0,7353	0,8671	0,6538	0,6538	0,6538
173	3	uzaklık	1	0,7353	1	0,6538	0,6538	0,6538
174	4	tekdüze	4	0,7941	0,7595	0,9286	0,5	0,65
175	10	uzaklık	4	0,7941	1	0,9286	0,5	0,65
176	10	uzaklık	5	0,7941	1	0,9286	0,5	0,65

177	11	uzaklık	4	0,7941	1	0,9286	0,5	0,65
178	12	uzaklık	4	0,7941	1	0,9286	0,5	0,65
179	12	uzaklık	5	0,7941	1	0,9286	0,5	0,65
180	14	uzaklık	2	0,7941	1	0,9286	0,5	0,65
181	11	tekdüze	1	0,7794	0,7405	0,8667	0,5	0,6341
182	11	uzaklık	1	0,7794	1	0,8667	0,5	0,6341
183	2	tekdüze	2	0,7941	0,8101	1	0,4615	0,6316
184	6	tekdüze	5	0,7941	0,7468	1	0,4615	0,6316
185	6	tekdüze	7	0,7941	0,7468	1	0,4615	0,6316
186	8	tekdüze	4	0,7941	0,7468	1	0,4615	0,6316
187	9	tekdüze	5	0,7941	0,7658	1	0,4615	0,6316
188	9	tekdüze	6	0,7941	0,7722	1	0,4615	0,6316
189	9	tekdüze	9	0,7941	0,7468	1	0,4615	0,6316
190	10	tekdüze	3	0,7941	0,7278	1	0,4615	0,6316
191	11	tekdüze	5	0,7941	0,7215	1	0,4615	0,6316
192	11	tekdüze	6	0,7941	0,7089	1	0,4615	0,6316
193	11	tekdüze	8	0,7941	0,7215	1	0,4615	0,6316
194	11	tekdüze	9	0,7941	0,7152	1	0,4615	0,6316
195	11	uzaklık	6	0,7941	1	1	0,4615	0,6316
196	11	uzaklık	8	0,7941	1	1	0,4615	0,6316
197	11	uzaklık	9	0,7941	1	1	0,4615	0,6316
198	13	tekdüze	1	0,7941	0,7278	1	0,4615	0,6316
199	13	uzaklık	1	0,7941	1	1	0,4615	0,6316
200	13	uzaklık	7	0,7941	1	1	0,4615	0,6316
201	13	uzaklık	8	0,7941	1	1	0,4615	0,6316
202	14	uzaklık	1	0,7941	1	1	0,4615	0,6316
203	14	uzaklık	5	0,7941	1	1	0,4615	0,6316
204	14	uzaklık	7	0,7941	1	1	0,4615	0,6316
205	14	uzaklık	8	0,7941	1	1	0,4615	0,6316
206	15	tekdüze	2	0,7941	0,7152	1	0,4615	0,6316
207	15	uzaklık	1	0,7941	1	1	0,4615	0,6316
208	17	uzaklık	1	0,7941	1	1	0,4615	0,6316
209	17	uzaklık	2	0,7941	1	1	0,4615	0,6316
210	17	uzaklık	4	0,7941	1	1	0,4615	0,6316
211	4	tekdüze	5	0,7794	0,7532	0,9231	0,4615	0,6154
212	4	tekdüze	6	0,7794	0,7468	0,9231	0,4615	0,6154
213	12	uzaklık	3	0,7794	1	0,9231	0,4615	0,6154
214	2	tekdüze	1	0,7647	0,8101	0,8571	0,4615	0,6
215	4	tekdüze	3	0,7647	0,7722	0,8571	0,4615	0,6
216	6	tekdüze	8	0,7794	0,7405	1	0,4231	0,5946
217	6	tekdüze	9	0,7794	0,7405	1	0,4231	0,5946
218	8	tekdüze	1	0,7794	0,7342	1	0,4231	0,5946
219	8	tekdüze	5	0,7794	0,7532	1	0,4231	0,5946
220	8	tekdüze	6	0,7794	0,7532	1	0,4231	0,5946
221	8	tekdüze	7	0,7794	0,7595	1	0,4231	0,5946

222	8	tekdüze	8	0,7794	0,7532	1	0,4231	0,5946
223	10	tekdüze	1	0,7794	0,7342	1	0,4231	0,5946
224	10	tekdüze	5	0,7794	0,7342	1	0,4231	0,5946
225	10	tekdüze	6	0,7794	0,7342	1	0,4231	0,5946
226	10	tekdüze	7	0,7794	0,7342	1	0,4231	0,5946
227	12	tekdüze	1	0,7794	0,7278	1	0,4231	0,5946
228	12	tekdüze	2	0,7794	0,7278	1	0,4231	0,5946
229	13	tekdüze	2	0,7794	0,7278	1	0,4231	0,5946
230	13	tekdüze	3	0,7794	0,7342	1	0,4231	0,5946
231	13	tekdüze	7	0,7794	0,7215	1	0,4231	0,5946
232	13	tekdüze	8	0,7794	0,7152	1	0,4231	0,5946
233	13	uzaklık	5	0,7794	1	1	0,4231	0,5946
234	13	uzaklık	6	0,7794	1	1	0,4231	0,5946
235	13	uzaklık	9	0,7794	1	1	0,4231	0,5946
236	14	uzaklık	9	0,7794	1	1	0,4231	0,5946
237	15	tekdüze	1	0,7794	0,7405	1	0,4231	0,5946
238	15	tekdüze	6	0,7794	0,7215	1	0,4231	0,5946
239	15	uzaklık	5	0,7794	1	1	0,4231	0,5946
240	15	uzaklık	6	0,7794	1	1	0,4231	0,5946
241	16	tekdüze	1	0,7794	0,7342	1	0,4231	0,5946
242	16	uzaklık	1	0,7794	1	1	0,4231	0,5946
243	16	uzaklık	4	0,7794	1	1	0,4231	0,5946
244	16	uzaklık	5	0,7794	1	1	0,4231	0,5946
245	16	uzaklık	6	0,7794	1	1	0,4231	0,5946
246	16	uzaklık	7	0,7794	1	1	0,4231	0,5946
247	17	tekdüze	1	0,7794	0,7405	1	0,4231	0,5946
248	18	uzaklık	2	0,7794	1	1	0,4231	0,5946
249	18	uzaklık	3	0,7794	1	1	0,4231	0,5946
250	18	uzaklık	4	0,7794	1	1	0,4231	0,5946

EK C: Destek Vektö Makineleri parametre ayarlamaları

	çekirdek	derece	C	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
1	polinom	4	100	0,83824	0,91139	0,85714	0,69231	0,76596
2	rbf	4	10	0,83824	0,91772	0,89474	0,65385	0,75556
3	polinom	1	100	0,82353	0,84177	0,9375	0,57692	0,71429
4	rbf	4	1	0,82353	0,76582	1	0,53846	0,7
5	polinom	4	1000	0,75	0,96203	0,65517	0,73077	0,69091
6	polinom	3	100	0,79412	0,93038	0,83333	0,57692	0,68182
7	doğrusal	4	10	0,79412	0,8481	0,83333	0,57692	0,68182
8	polinom	3	10	0,79412	0,84177	0,875	0,53846	0,66667
9	polinom	1	10	0,80882	0,81646	1	0,5	0,66667
10	polinom	2	100	0,73529	0,94304	0,64286	0,69231	0,66667
11	rbf	4	100	0,73529	0,98101	0,65385	0,65385	0,65385
12	doğrusal	4	1	0,79412	0,81646	1	0,46154	0,63158
13	rbf	4	1000	0,69118	1	0,59259	0,61538	0,60377
14	polinom	2	10	0,76471	0,82911	0,85714	0,46154	0,6
15	polinom	1	1	0,77941	0,75316	1	0,42308	0,59459
16	doğrusal	4	0,01	0,77941	0,74684	1	0,42308	0,59459
17	doğrusal	4	0,1	0,77941	0,77848	1	0,42308	0,59459
18	doğrusal	4	100	0,73529	0,87342	0,72222	0,5	0,59091
19	doğrusal	4	1000	0,73529	0,87342	0,72222	0,5	0,59091
20	polinom	3	1000	0,69118	1	0,6087	0,53846	0,57143
21	polinom	4	10	0,75	0,82911	0,84615	0,42308	0,5641
22	polinom	1	1000	0,72059	0,87342	0,70588	0,46154	0,55814
23	polinom	2	1000	0,60294	1	0,48387	0,57692	0,52632
24	polinom	1	0,1	0,75	0,72785	1	0,34615	0,51429
25	polinom	2	1	0,75	0,72785	1	0,34615	0,51429
26	polinom	4	1	0,72059	0,75316	0,81818	0,34615	0,48649
27	polinom	3	1	0,73529	0,74684	1	0,30769	0,47059
28	polinom	3	0,1	0,72059	0,73418	1	0,26923	0,42424
29	sigmoid	4	100	0,55882	0,60759	0,41667	0,38462	0,4
30	sigmoid	4	1000	0,55882	0,60759	0,41667	0,38462	0,4
31	sigmoid	4	1	0,57353	0,59494	0,42857	0,34615	0,38298
32	polinom	4	0,1	0,70588	0,72785	1	0,23077	0,375
33	sigmoid	4	10	0,54412	0,59494	0,3913	0,34615	0,36735
34	polinom	2	0,1	0,67647	0,72785	1	0,15385	0,26667
35	polinom	3	0,01	0,64706	0,70253	1	0,07692	0,14286
36	polinom	4	0,01	0,64706	0,70886	1	0,07692	0,14286
37	sigmoid	4	0,001	0,61765	0,65823	0	0	0
38	sigmoid	4	0,01	0,61765	0,65823	0	0	0
39	sigmoid	4	0,1	0,61765	0,65823	0	0	0
40	rbf	4	0,001	0,61765	0,65823	0	0	0
41	rbf	4	0,01	0,61765	0,65823	0	0	0
42	rbf	4	0,1	0,61765	0,65823	0	0	0

43	polinom	1	0,001	0,61765	0,65823	0	0	0
44	polinom	1	0,01	0,61765	0,65823	0	0	0
45	polinom	2	0,001	0,61765	0,65823	0	0	0
46	polinom	2	0,01	0,61765	0,65823	0	0	0
47	polinom	3	0,001	0,61765	0,65823	0	0	0
48	polinom	4	0,001	0,61765	0,66456	0	0	0
49	doğrusal	4	0,001	0,61765	0,65823	0	0	0



EK D : Rastgele Orman Yöntemi Parametre Ayarlaması

	n KA	ayırma endeksi	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
1	13	entropi	0,82353	0,99367	0,85	0,65385	0,73913
2	18	gini	0,80882	0,99367	0,80952	0,65385	0,7234
3	18	entropi	0,79412	0,99367	0,8	0,61538	0,69565
4	10	gini	0,77941	0,98101	0,73913	0,65385	0,69388
5	19	gini	0,76471	1	0,69231	0,69231	0,69231
6	5	gini	0,75	0,96835	0,65517	0,73077	0,69091
7	1	entropi	0,77941	0,87342	0,7619	0,61538	0,68085
8	7	gini	0,77941	0,97468	0,7619	0,61538	0,68085
9	16	gini	0,77941	1	0,7619	0,61538	0,68085
10	15	entropi	0,76471	1	0,70833	0,65385	0,68
11	17	gini	0,76471	1	0,72727	0,61538	0,66667
12	15	gini	0,73529	0,98734	0,65385	0,65385	0,65385
13	14	entropi	0,76471	0,98734	0,75	0,57692	0,65217
14	17	entropi	0,73529	1	0,66667	0,61538	0,64
15	7	entropi	0,75	0,98101	0,71429	0,57692	0,6383
16	11	gini	0,75	0,99367	0,71429	0,57692	0,6383
17	8	entropi	0,76471	0,98101	0,77778	0,53846	0,63636
18	13	gini	0,73529	0,98101	0,68182	0,57692	0,625
19	6	entropi	0,75	0,94937	0,73684	0,53846	0,62222
20	8	gini	0,76471	0,98101	0,8125	0,5	0,61905
21	9	entropi	0,72059	0,97468	0,65217	0,57692	0,61224
22	12	gini	0,72059	0,98734	0,65217	0,57692	0,61224
23	14	gini	0,73529	1	0,7	0,53846	0,6087
24	16	entropi	0,75	1	0,76471	0,5	0,60465
25	11	entropi	0,69118	0,99367	0,59259	0,61538	0,60377
26	12	entropi	0,73529	0,98734	0,72222	0,5	0,59091
27	3	entropi	0,73529	0,9557	0,75	0,46154	0,57143
28	19	entropi	0,73529	0,98101	0,75	0,46154	0,57143
29	10	entropi	0,70588	0,98734	0,65	0,5	0,56522
30	3	gini	0,63235	0,93671	0,51613	0,61538	0,5614
31	4	entropi	0,73529	0,96203	0,78571	0,42308	0,55
32	5	entropi	0,67647	0,93038	0,59091	0,5	0,54167
33	2	entropi	0,72059	0,88608	0,73333	0,42308	0,53659
34	4	gini	0,70588	0,93671	0,6875	0,42308	0,52381
35	9	gini	0,64706	0,98734	0,54167	0,5	0,52
36	2	gini	0,70588	0,87975	0,71429	0,38462	0,5
37	1	gini	0,63235	0,90506	0,52381	0,42308	0,46809
38	6	gini	0,66176	0,97468	0,58824	0,38462	0,46512

EK E: Gradyan Arttırma Yöntemi Parametre Ayarlaması (En iyi 250 parametre)

	aşama sayısı	kriter	kayıp fonksiyonu	doğruluk_test	doğruluk_eğt	kesinlik	duyarlılık	F1
1	35	0,7	sapma	0,8382	1	0,8261	0,7308	0,7755
2	80	0,2	üstel	0,8382	1	0,8261	0,7308	0,7755
3	150	0,2	üstel	0,8382	1	0,8261	0,7308	0,7755
4	170	0,1	sapma	0,8382	1	0,8261	0,7308	0,7755
5	110	0,1	sapma	0,8382	1	0,8571	0,6923	0,766
6	130	0,1	sapma	0,8382	1	0,8571	0,6923	0,766
7	160	0,1	sapma	0,8382	1	0,8571	0,6923	0,766
8	165	0,1	sapma	0,8382	1	0,8571	0,6923	0,766
9	70	0,75	sapma	0,8235	1	0,7917	0,7308	0,76
10	160	0,2	üstel	0,8235	1	0,7917	0,7308	0,76
11	15	0,5	üstel	0,8235	1	0,8182	0,6923	0,75
12	85	0,2	üstel	0,8235	1	0,8182	0,6923	0,75
13	100	0,1	sapma	0,8235	1	0,8182	0,6923	0,75
14	115	0,2	üstel	0,8235	1	0,8182	0,6923	0,75
15	145	0,1	sapma	0,8235	1	0,8182	0,6923	0,75
16	155	0,1	sapma	0,8235	1	0,8182	0,6923	0,75
17	40	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
18	55	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
19	85	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
20	125	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
21	150	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
22	185	0,75	sapma	0,8088	1	0,76	0,7308	0,7451
23	10	0,2	üstel	0,8235	0,9177	0,85	0,6538	0,7391
24	35	0,1	sapma	0,8235	0,981	0,85	0,6538	0,7391
25	135	0,1	sapma	0,8235	1	0,85	0,6538	0,7391
26	140	0,1	sapma	0,8235	1	0,85	0,6538	0,7391
27	165	0,9	sapma	0,8235	1	0,85	0,6538	0,7391
28	20	0,8	üstel	0,8088	1	0,7826	0,6923	0,7347
29	25	0,4	üstel	0,8088	1	0,7826	0,6923	0,7347
30	45	0,45	sapma	0,8088	1	0,7826	0,6923	0,7347
31	50	0,7	sapma	0,8088	1	0,7826	0,6923	0,7347
32	65	0,25	üstel	0,8088	1	0,7826	0,6923	0,7347
33	70	0,25	sapma	0,8088	1	0,7826	0,6923	0,7347
34	80	0,25	üstel	0,8088	1	0,7826	0,6923	0,7347
35	105	0,1	sapma	0,8088	1	0,7826	0,6923	0,7347
36	170	0,15	üstel	0,8088	1	0,7826	0,6923	0,7347
37	175	0,1	sapma	0,8088	1	0,7826	0,6923	0,7347
38	40	0,7	sapma	0,7941	1	0,7308	0,7308	0,7308
39	45	0,6	üstel	0,7941	1	0,7308	0,7308	0,7308
40	45	0,7	sapma	0,7941	1	0,7308	0,7308	0,7308
41	70	0,45	sapma	0,7941	1	0,7308	0,7308	0,7308

42	125	0,3	üstel	0,7941	1	0,7308	0,7308	0,7308
43	160	0,45	sapma	0,7941	1	0,7308	0,7308	0,7308
44	5	0,35	sapma	0,8088	0,9367	0,8095	0,6538	0,7234
45	15	0,15	üstel	0,8088	0,9177	0,8095	0,6538	0,7234
46	20	0,15	üstel	0,8088	0,9494	0,8095	0,6538	0,7234
47	30	0,25	sapma	0,8088	1	0,8095	0,6538	0,7234
48	30	0,9	sapma	0,8088	1	0,8095	0,6538	0,7234
49	40	0,3	sapma	0,8088	1	0,8095	0,6538	0,7234
50	45	0,35	üstel	0,8088	1	0,8095	0,6538	0,7234
51	50	0,9	sapma	0,8088	1	0,8095	0,6538	0,7234
52	55	0,25	sapma	0,8088	1	0,8095	0,6538	0,7234
53	75	0,1	sapma	0,8088	1	0,8095	0,6538	0,7234
54	90	0,1	sapma	0,8088	1	0,8095	0,6538	0,7234
55	90	0,15	üstel	0,8088	1	0,8095	0,6538	0,7234
56	100	0,15	üstel	0,8088	1	0,8095	0,6538	0,7234
57	100	0,9	sapma	0,8088	1	0,8095	0,6538	0,7234
58	115	0,1	sapma	0,8088	1	0,8095	0,6538	0,7234
59	115	0,9	sapma	0,8088	1	0,8095	0,6538	0,7234
60	125	0,9	sapma	0,8088	1	0,8095	0,6538	0,7234
61	150	0,1	sapma	0,8088	1	0,8095	0,6538	0,7234
62	155	0,25	sapma	0,8088	1	0,8095	0,6538	0,7234
63	195	0,2	üstel	0,8088	1	0,8095	0,6538	0,7234
64	195	0,25	sapma	0,8088	1	0,8095	0,6538	0,7234
65	25	0,45	sapma	0,7941	1	0,75	0,6923	0,72
66	50	0,15	sapma	0,7941	1	0,75	0,6923	0,72
67	55	0,2	üstel	0,7941	1	0,75	0,6923	0,72
68	55	0,7	sapma	0,7941	1	0,75	0,6923	0,72
69	60	0,4	üstel	0,7941	1	0,75	0,6923	0,72
70	65	0,75	sapma	0,7941	1	0,75	0,6923	0,72
71	70	0,25	üstel	0,7941	1	0,75	0,6923	0,72
72	70	0,7	sapma	0,7941	1	0,75	0,6923	0,72
73	75	0,2	sapma	0,7941	1	0,75	0,6923	0,72
74	75	0,25	sapma	0,7941	1	0,75	0,6923	0,72
75	75	0,35	üstel	0,7941	1	0,75	0,6923	0,72
76	85	0,35	üstel	0,7941	1	0,75	0,6923	0,72
77	90	0,25	üstel	0,7941	1	0,75	0,6923	0,72
78	95	0,1	üstel	0,7941	1	0,75	0,6923	0,72
79	95	0,2	üstel	0,7941	1	0,75	0,6923	0,72
80	100	0,15	sapma	0,7941	1	0,75	0,6923	0,72
81	105	0,25	üstel	0,7941	1	0,75	0,6923	0,72
82	105	0,35	üstel	0,7941	1	0,75	0,6923	0,72
83	105	0,7	sapma	0,7941	1	0,75	0,6923	0,72
84	110	0,35	üstel	0,7941	1	0,75	0,6923	0,72
85	110	0,7	sapma	0,7941	1	0,75	0,6923	0,72
86	110	0,75	sapma	0,7941	1	0,75	0,6923	0,72

87	115	0,7	sapma	0,7941	1	0,75	0,6923	0,72
88	120	0,15	üstel	0,7941	1	0,75	0,6923	0,72
89	120	0,35	üstel	0,7941	1	0,75	0,6923	0,72
90	125	0,35	üstel	0,7941	1	0,75	0,6923	0,72
91	125	0,7	sapma	0,7941	1	0,75	0,6923	0,72
92	130	0,15	üstel	0,7941	1	0,75	0,6923	0,72
93	145	0,15	üstel	0,7941	1	0,75	0,6923	0,72
94	145	0,2	üstel	0,7941	1	0,75	0,6923	0,72
95	150	0,35	üstel	0,7941	1	0,75	0,6923	0,72
96	150	0,75	üstel	0,7941	1	0,75	0,6923	0,72
97	165	0,35	üstel	0,7941	1	0,75	0,6923	0,72
98	170	0,2	üstel	0,7941	1	0,75	0,6923	0,72
99	190	0,7	sapma	0,7941	1	0,75	0,6923	0,72
100	195	0,75	üstel	0,7941	1	0,75	0,6923	0,72
101	35	0,4	üstel	0,7794	1	0,7037	0,7308	0,717
102	50	0,5	sapma	0,7794	1	0,7037	0,7308	0,717
103	100	0,45	sapma	0,7794	1	0,7037	0,7308	0,717
104	105	0,3	üstel	0,7794	1	0,7037	0,7308	0,717
105	110	0,45	sapma	0,7794	1	0,7037	0,7308	0,717
106	115	0,45	sapma	0,7794	1	0,7037	0,7308	0,717
107	135	0,15	sapma	0,7794	1	0,7037	0,7308	0,717
108	135	0,3	üstel	0,7794	1	0,7037	0,7308	0,717
109	145	0,15	sapma	0,7794	1	0,7037	0,7308	0,717
110	165	0,45	sapma	0,7794	1	0,7037	0,7308	0,717
111	5	0,5	üstel	0,8088	0,9304	0,8421	0,6154	0,7111
112	20	0,3	üstel	0,8088	1	0,8421	0,6154	0,7111
113	25	0,15	üstel	0,8088	0,9684	0,8421	0,6154	0,7111
114	30	0,15	üstel	0,8088	0,9937	0,8421	0,6154	0,7111
115	40	0,1	sapma	0,8088	0,9937	0,8421	0,6154	0,7111
116	55	0,1	sapma	0,8088	1	0,8421	0,6154	0,7111
117	60	0,1	sapma	0,8088	1	0,8421	0,6154	0,7111
118	65	0,1	sapma	0,8088	1	0,8421	0,6154	0,7111
119	5	0,4	sapma	0,7941	0,9367	0,7727	0,6538	0,7083
120	5	0,6	üstel	0,7941	0,9304	0,7727	0,6538	0,7083
121	15	0,2	üstel	0,7941	0,9304	0,7727	0,6538	0,7083
122	20	0,45	sapma	0,7941	1	0,7727	0,6538	0,7083
123	25	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
124	35	0,15	sapma	0,7941	1	0,7727	0,6538	0,7083
125	35	0,2	üstel	0,7941	1	0,7727	0,6538	0,7083
126	35	0,65	sapma	0,7941	1	0,7727	0,6538	0,7083
127	40	0,4	üstel	0,7941	1	0,7727	0,6538	0,7083
128	50	0,1	sapma	0,7941	1	0,7727	0,6538	0,7083
129	55	0,35	üstel	0,7941	1	0,7727	0,6538	0,7083
130	60	0,2	üstel	0,7941	1	0,7727	0,6538	0,7083
131	60	0,35	üstel	0,7941	1	0,7727	0,6538	0,7083

132	70	0,15	üstel	0,7941	1	0,7727	0,6538	0,7083
133	75	0,9	sapma	0,7941	1	0,7727	0,6538	0,7083
134	80	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
135	95	0,15	üstel	0,7941	1	0,7727	0,6538	0,7083
136	105	0,15	üstel	0,7941	1	0,7727	0,6538	0,7083
137	110	0,15	üstel	0,7941	1	0,7727	0,6538	0,7083
138	110	0,2	üstel	0,7941	1	0,7727	0,6538	0,7083
139	115	0,15	üstel	0,7941	1	0,7727	0,6538	0,7083
140	120	0,9	sapma	0,7941	1	0,7727	0,6538	0,7083
141	120	0,9	üstel	0,7941	1	0,7727	0,6538	0,7083
142	130	0,9	sapma	0,7941	1	0,7727	0,6538	0,7083
143	135	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
144	150	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
145	175	0,9	sapma	0,7941	1	0,7727	0,6538	0,7083
146	180	0,2	sapma	0,7941	1	0,7727	0,6538	0,7083
147	180	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
148	185	0,25	sapma	0,7941	1	0,7727	0,6538	0,7083
149	195	0,65	sapma	0,7941	1	0,7727	0,6538	0,7083
150	35	0,3	üstel	0,7794	1	0,72	0,6923	0,7059
151	40	0,8	sapma	0,7794	1	0,72	0,6923	0,7059
152	45	0,2	üstel	0,7794	1	0,72	0,6923	0,7059
153	60	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
154	65	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
155	65	0,2	üstel	0,7794	1	0,72	0,6923	0,7059
156	80	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
157	80	0,5	sapma	0,7794	1	0,72	0,6923	0,7059
158	85	0,5	sapma	0,7794	1	0,72	0,6923	0,7059
159	90	0,7	sapma	0,7794	1	0,72	0,6923	0,7059
160	95	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
161	105	0,2	üstel	0,7794	1	0,72	0,6923	0,7059
162	105	0,75	üstel	0,7794	1	0,72	0,6923	0,7059
163	110	0,3	üstel	0,7794	1	0,72	0,6923	0,7059
164	110	0,75	üstel	0,7794	1	0,72	0,6923	0,7059
165	115	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
166	120	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
167	120	0,25	üstel	0,7794	1	0,72	0,6923	0,7059
168	125	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
169	125	0,4	sapma	0,7794	1	0,72	0,6923	0,7059
170	140	0,7	sapma	0,7794	1	0,72	0,6923	0,7059
171	150	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
172	155	0,35	üstel	0,7794	1	0,72	0,6923	0,7059
173	155	0,5	sapma	0,7794	1	0,72	0,6923	0,7059
174	155	0,75	üstel	0,7794	1	0,72	0,6923	0,7059
175	160	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
176	165	0,2	sapma	0,7794	1	0,72	0,6923	0,7059

177	165	0,2	üstel	0,7794	1	0,72	0,6923	0,7059
178	165	0,75	üstel	0,7794	1	0,72	0,6923	0,7059
179	170	0,4	sapma	0,7794	1	0,72	0,6923	0,7059
180	175	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
181	175	0,2	üstel	0,7794	1	0,72	0,6923	0,7059
182	185	0,35	üstel	0,7794	1	0,72	0,6923	0,7059
183	185	0,4	üstel	0,7794	1	0,72	0,6923	0,7059
184	190	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
185	190	0,5	sapma	0,7794	1	0,72	0,6923	0,7059
186	190	0,75	üstel	0,7794	1	0,72	0,6923	0,7059
187	195	0,15	sapma	0,7794	1	0,72	0,6923	0,7059
188	195	0,35	üstel	0,7794	1	0,72	0,6923	0,7059
189	195	0,8	sapma	0,7794	1	0,72	0,6923	0,7059
190	15	0,7	sapma	0,7647	1	0,6786	0,7308	0,7037
191	90	0,45	sapma	0,7647	1	0,6786	0,7308	0,7037
192	120	0,45	sapma	0,7647	1	0,6786	0,7308	0,7037
193	130	0,15	sapma	0,7647	1	0,6786	0,7308	0,7037
194	130	0,3	üstel	0,7647	1	0,6786	0,7308	0,7037
195	155	0,45	sapma	0,7647	1	0,6786	0,7308	0,7037
196	190	0,3	üstel	0,7647	1	0,6786	0,7308	0,7037
197	190	0,45	sapma	0,7647	1	0,6786	0,7308	0,7037
198	195	0,45	sapma	0,7647	1	0,6786	0,7308	0,7037
199	5	0,25	sapma	0,7941	0,9051	0,8	0,6154	0,6957
200	10	0,4	üstel	0,7941	0,962	0,8	0,6154	0,6957
201	15	0,35	üstel	0,7941	0,9937	0,8	0,6154	0,6957
202	20	0,15	sapma	0,7941	0,9684	0,8	0,6154	0,6957
203	20	0,65	üstel	0,7941	1	0,8	0,6154	0,6957
204	30	0,2	üstel	0,7941	1	0,8	0,6154	0,6957
205	40	0,1	üstel	0,7941	0,981	0,8	0,6154	0,6957
206	45	0,25	sapma	0,7941	1	0,8	0,6154	0,6957
207	70	0,3	sapma	0,7941	1	0,8	0,6154	0,6957
208	70	0,9	sapma	0,7941	1	0,8	0,6154	0,6957
209	85	0,15	üstel	0,7941	1	0,8	0,6154	0,6957
210	85	0,9	sapma	0,7941	1	0,8	0,6154	0,6957
211	95	0,1	sapma	0,7941	1	0,8	0,6154	0,6957
212	125	0,65	üstel	0,7941	1	0,8	0,6154	0,6957
213	160	0,8	sapma	0,7941	1	0,8	0,6154	0,6957
214	175	0,25	sapma	0,7941	1	0,8	0,6154	0,6957
215	185	0,8	sapma	0,7941	1	0,8	0,6154	0,6957
216	15	0,15	sapma	0,7794	0,9367	0,7391	0,6538	0,6939
217	15	0,2	sapma	0,7794	0,962	0,7391	0,6538	0,6939
218	20	0,2	sapma	0,7794	0,9937	0,7391	0,6538	0,6939
219	20	0,35	üstel	0,7794	1	0,7391	0,6538	0,6939
220	25	0,2	üstel	0,7794	0,9937	0,7391	0,6538	0,6939
221	25	0,25	üstel	0,7794	1	0,7391	0,6538	0,6939

222	25	0,35	üstel	0,7794	1	0,7391	0,6538	0,6939
223	25	0,65	üstel	0,7794	1	0,7391	0,6538	0,6939
224	25	0,7	üstel	0,7794	1	0,7391	0,6538	0,6939
225	35	0,35	üstel	0,7794	1	0,7391	0,6538	0,6939
226	35	0,75	üstel	0,7794	1	0,7391	0,6538	0,6939
227	40	0,2	sapma	0,7794	1	0,7391	0,6538	0,6939
228	40	0,25	sapma	0,7794	1	0,7391	0,6538	0,6939
229	40	0,65	sapma	0,7794	1	0,7391	0,6538	0,6939
230	55	0,1	üstel	0,7794	1	0,7391	0,6538	0,6939
231	55	0,25	üstel	0,7794	1	0,7391	0,6538	0,6939
232	55	0,4	sapma	0,7794	1	0,7391	0,6538	0,6939
233	55	0,4	üstel	0,7794	1	0,7391	0,6538	0,6939
234	55	0,75	üstel	0,7794	1	0,7391	0,6538	0,6939
235	60	0,25	sapma	0,7794	1	0,7391	0,6538	0,6939
236	65	0,3	üstel	0,7794	1	0,7391	0,6538	0,6939
237	70	0,35	üstel	0,7794	1	0,7391	0,6538	0,6939
238	70	0,65	üstel	0,7794	1	0,7391	0,6538	0,6939
239	70	0,85	üstel	0,7794	1	0,7391	0,6538	0,6939
240	75	0,15	üstel	0,7794	1	0,7391	0,6538	0,6939
241	75	0,65	sapma	0,7794	1	0,7391	0,6538	0,6939
242	80	0,65	üstel	0,7794	1	0,7391	0,6538	0,6939
243	85	0,35	sapma	0,7794	1	0,7391	0,6538	0,6939
244	85	0,75	üstel	0,7794	1	0,7391	0,6538	0,6939
245	90	0,35	sapma	0,7794	1	0,7391	0,6538	0,6939
246	100	0,2	üstel	0,7794	1	0,7391	0,6538	0,6939
247	100	0,25	üstel	0,7794	1	0,7391	0,6538	0,6939
248	105	0,1	üstel	0,7794	1	0,7391	0,6538	0,6939
249	105	0,25	sapma	0,7794	1	0,7391	0,6538	0,6939
250	105	0,35	sapma	0,7794	1	0,7391	0,6538	0,6939

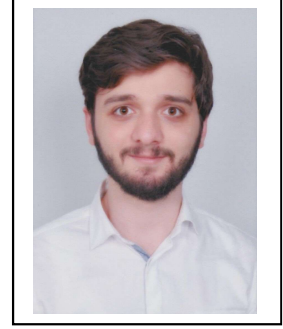
EK F: YSA parametre ayarlamaları

	nöron sayısı	dropout	en iyileştirici	epoch	doğruluk _test	doğruluk _eğt	kesinlik	duyarlılık	F1
1	32	0,2	RMSprop	100	0,8824	0,8797	0,9091	0,7692	0,8333
2	16	0,3	RMSprop	200	0,8824	0,8797	0,9091	0,7692	0,8333
3	24	0,3	adam	100	0,8824	0,8544	0,9091	0,7692	0,8333
4	24	0,2	RMSprop	200	0,8676	0,9241	0,8400	0,8077	0,8235
5	16	0,2	RMSprop	200	0,8676	0,8924	0,8696	0,7692	0,8163
6	40	0,2	RMSprop	200	0,8676	0,9557	0,8696	0,7692	0,8163
7	40	0,3	RMSprop	200	0,8676	0,9430	0,8696	0,7692	0,8163
8	40	0,4	RMSprop	200	0,8676	0,9304	0,8696	0,7692	0,8163
9	48	0,4	RMSprop	200	0,8676	0,9367	0,8696	0,7692	0,8163
10	40	0,2	adam	200	0,8676	0,9620	0,9048	0,7308	0,8085
11	40	0,4	adam	400	0,8529	0,9810	0,8077	0,8077	0,8077
12	40	0,1	adam	100	0,8529	0,9304	0,8333	0,7692	0,8000
13	48	0,4	adam	300	0,8676	0,9747	0,9474	0,6923	0,8000
14	48	0,4	RMSprop	400	0,8676	0,9937	0,9474	0,6923	0,8000
15	32	0,1	adam	200	0,8529	0,9557	0,8636	0,7308	0,7917
16	40	0,1	RMSprop	100	0,8529	0,9241	0,8636	0,7308	0,7917
17	32	0,2	adam	100	0,8529	0,8608	0,8636	0,7308	0,7917
18	40	0,2	RMSprop	100	0,8529	0,8861	0,8636	0,7308	0,7917
19	40	0,4	adam	100	0,8529	0,8418	0,8636	0,7308	0,7917
20	40	0,4	RMSprop	100	0,8529	0,8544	0,8636	0,7308	0,7917
21	40	0,4	adam	200	0,8529	0,9494	0,8636	0,7308	0,7917
22	48	0,4	adam	200	0,8529	0,9557	0,8636	0,7308	0,7917
23	48	0,4	RMSprop	300	0,8529	0,9684	0,8636	0,7308	0,7917
24	40	0,2	RMSprop	200	0,8529	0,9494	0,8636	0,7308	0,7917
25	16	0,1	adam	200	0,8382	0,9241	0,8000	0,7692	0,7843
26	40	0,3	RMSprop	100	0,8382	0,8924	0,8000	0,7692	0,7843
27	40	0,2	RMSprop	400	0,8382	0,9937	0,8000	0,7692	0,7843
28	40	0,4	RMSprop	200	0,8529	0,9557	0,9000	0,6923	0,7826
29	40	0,2	adam	300	0,8529	0,9684	0,9000	0,6923	0,7826
30	16	0,1	RMSprop	200	0,8235	0,9114	0,7500	0,8077	0,7778
31	24	0,1	adam	200	0,8382	0,9430	0,8261	0,7308	0,7755
32	40	0,1	RMSprop	200	0,8382	0,9620	0,8261	0,7308	0,7755
33	24	0,2	adam	200	0,8382	0,9177	0,8261	0,7308	0,7755
34	32	0,2	RMSprop	200	0,8382	0,9494	0,8261	0,7308	0,7755
35	24	0,3	RMSprop	200	0,8382	0,8924	0,8261	0,7308	0,7755
36	32	0,3	adam	200	0,8382	0,9430	0,8261	0,7308	0,7755
37	40	0,4	adam	200	0,8382	0,9367	0,8261	0,7308	0,7755
38	48	0,4	adam	400	0,8382	0,9873	0,8261	0,7308	0,7755
39	32	0,2	adam	200	0,8382	0,9430	0,8261	0,7308	0,7755
40	32	0,2	RMSprop	200	0,8382	0,9430	0,8261	0,7308	0,7755
41	40	0,2	RMSprop	300	0,8382	0,9747	0,8261	0,7308	0,7755

42	40	0,4	RMSprop	400	0,8529	0,9810	0,9444	0,6538	0,7727
43	8	0,1	adam	200	0,8235	0,8671	0,7692	0,7692	0,7692
44	32	0,4	RMSprop	200	0,8235	0,9114	0,7692	0,7692	0,7692
45	48	0,2	adam	200	0,8235	0,9684	0,7692	0,7692	0,7692
46	40	0,1	adam	200	0,8382	0,9684	0,8571	0,6923	0,7660
47	24	0,3	RMSprop	100	0,8382	0,8291	0,8571	0,6923	0,7660
48	32	0,3	RMSprop	200	0,8382	0,9241	0,8571	0,6923	0,7660
49	40	0,3	adam	200	0,8382	0,9557	0,8571	0,6923	0,7660
50	32	0,4	adam	200	0,8382	0,9241	0,8571	0,6923	0,7660
51	40	0,4	RMSprop	300	0,8382	0,9557	0,8571	0,6923	0,7660
52	16	0,2	adam	300	0,8382	0,9241	0,8571	0,6923	0,7660
53	24	0,2	RMSprop	400	0,8382	0,9620	0,8571	0,6923	0,7660
54	32	0,2	adam	400	0,8382	0,9873	0,8571	0,6923	0,7660
55	48	0,2	adam	300	0,8382	0,9937	0,8571	0,6923	0,7660
56	48	0,2	adam	400	0,8382	1,0000	0,8571	0,6923	0,7660
57	48	0,2	RMSprop	400	0,8382	0,9937	0,8571	0,6923	0,7660
58	24	0,1	RMSprop	200	0,8235	0,9177	0,7917	0,7308	0,7600
59	32	0,1	RMSprop	200	0,8235	0,9367	0,7917	0,7308	0,7600
60	24	0,2	adam	100	0,8235	0,8291	0,7917	0,7308	0,7600
61	40	0,2	adam	200	0,8235	0,9684	0,7917	0,7308	0,7600
62	16	0,2	adam	200	0,8235	0,9051	0,7917	0,7308	0,7600
63	24	0,2	adam	200	0,8235	0,9177	0,7917	0,7308	0,7600
64	24	0,2	RMSprop	300	0,8235	0,9430	0,7917	0,7308	0,7600
65	32	0,2	RMSprop	400	0,8235	0,9810	0,7917	0,7308	0,7600
66	48	0,2	RMSprop	200	0,8235	0,9620	0,7917	0,7308	0,7600
67	16	0,2	adam	400	0,8382	0,9430	0,8947	0,6538	0,7556
68	32	0,2	RMSprop	300	0,8382	0,9620	0,8947	0,6538	0,7556
69	24	0,4	adam	200	0,8235	0,8861	0,8182	0,6923	0,7500
70	16	0,2	RMSprop	400	0,8235	0,9304	0,8182	0,6923	0,7500
71	24	0,2	adam	300	0,8235	0,9684	0,8182	0,6923	0,7500
72	24	0,2	RMSprop	200	0,8235	0,8987	0,8182	0,6923	0,7500
73	40	0,2	adam	400	0,8235	0,9810	0,8182	0,6923	0,7500
74	24	0,3	adam	200	0,8235	0,9177	0,8500	0,6538	0,7391
75	32	0,2	adam	200	0,8088	0,9557	0,7826	0,6923	0,7347
76	32	0,4	adam	100	0,8088	0,8544	0,7826	0,6923	0,7347
77	32	0,2	adam	300	0,8088	0,9873	0,7826	0,6923	0,7347
78	48	0,2	RMSprop	300	0,7941	0,9937	0,7308	0,7308	0,7308
79	16	0,3	adam	200	0,8235	0,8671	0,8889	0,6154	0,7273
80	40	0,3	adam	100	0,8235	0,8228	0,8889	0,6154	0,7273
81	16	0,4	adam	200	0,8235	0,8544	0,8889	0,6154	0,7273
82	24	0,4	RMSprop	100	0,8235	0,8165	0,8889	0,6154	0,7273
83	16	0,2	RMSprop	200	0,8235	0,7911	0,8889	0,6154	0,7273
84	24	0,4	RMSprop	200	0,8088	0,8734	0,8095	0,6538	0,7234
85	16	0,2	RMSprop	300	0,8088	0,8544	0,8095	0,6538	0,7234
86	24	0,1	RMSprop	100	0,8235	0,7911	0,9375	0,5769	0,7143

87	16	0,2	adam	200	0,8235	0,8418	0,9375	0,5769	0,7143
88	24	0,1	adam	100	0,8088	0,8291	0,8421	0,6154	0,7111
89	32	0,1	RMSprop	100	0,8088	0,8291	0,8421	0,6154	0,7111
90	40	0,4	adam	300	0,8088	0,9810	0,8421	0,6154	0,7111
91	24	0,2	adam	400	0,7941	0,9747	0,7727	0,6538	0,7083
92	32	0,1	adam	100	0,8088	0,8544	0,8824	0,5769	0,6977
93	8	0,3	adam	200	0,8088	0,8165	0,8824	0,5769	0,6977
94	24	0,4	adam	100	0,8088	0,8291	0,8824	0,5769	0,6977
95	32	0,4	RMSprop	100	0,8088	0,7975	0,8824	0,5769	0,6977
96	16	0,2	adam	100	0,8088	0,7785	0,9333	0,5385	0,6829
97	24	0,2	RMSprop	100	0,8088	0,7848	0,9333	0,5385	0,6829
98	16	0,3	adam	100	0,8088	0,7658	0,9333	0,5385	0,6829
99	32	0,3	RMSprop	100	0,8088	0,7785	0,9333	0,5385	0,6829
100	8	0,2	RMSprop	200	0,7941	0,8101	0,8333	0,5769	0,6818
101	40	0,2	adam	100	0,7941	0,8418	0,8333	0,5769	0,6818
102	32	0,3	adam	100	0,7941	0,7975	0,8333	0,5769	0,6818
103	16	0,4	adam	100	0,7941	0,8038	0,8750	0,5385	0,6667
104	16	0,4	RMSprop	200	0,7941	0,8101	0,8750	0,5385	0,6667
105	8	0,2	adam	100	0,8088	0,7595	1,0000	0,5000	0,6667
106	8	0,3	RMSprop	100	0,8088	0,7595	1,0000	0,5000	0,6667
107	16	0,2	RMSprop	100	0,7941	0,7722	0,9286	0,5000	0,6500
108	8	0,3	RMSprop	200	0,7941	0,7532	0,9286	0,5000	0,6500
109	16	0,3	RMSprop	100	0,7941	0,7658	0,9286	0,5000	0,6500
110	16	0,4	RMSprop	100	0,7941	0,7658	0,9286	0,5000	0,6500
111	8	0,1	RMSprop	100	0,7794	0,7468	0,8667	0,5000	0,6341
112	16	0,1	RMSprop	100	0,7794	0,7532	0,8667	0,5000	0,6341
113	8	0,4	adam	100	0,7941	0,7532	1,0000	0,4615	0,6316
114	8	0,4	RMSprop	100	0,7941	0,7405	1,0000	0,4615	0,6316
115	16	0,1	adam	100	0,7794	0,7785	0,9231	0,4615	0,6154

ÖZGEÇMİŞ



Ad-Soyad : Abdullah Enes Taha

Doğum Tarihi ve Yeri : Bahçelievler, 1995

E-posta : taha@itu.edu.tr

ÖĞRENİM DURUMU:

- **Lisans** : 2017, İstanbul Teknik Üniversitesi, İşletme Fakültesi, Endüstri Mühendisliği