



TÜRKİYE CUMHURİYETİ
KARADENİZ TEKNİK ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

BİYOİSTATİSTİK VE TIP BİLİŞİMİ ANABİLİM DALI

PULMONER EMBOLİ KLİNİK TANİ SİSTEMİ

Özlem KARADENİZ

YÜKSEK LİSANS TEZİ

Yrd. Doç. Dr. Burçin KURT

TRABZON-2018



TÜRKİYE CUMHURİYETİ
KARADENİZ TEKNİK ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

BİYOİSTATİSTİK VE TIP BİLİŞİMİ ANABİLİM DALI

PULMONER EMBOLİ KLİNİK TANİ SİSTEMİ

Özlem KARADENİZ

YÜKSEK LİSANS TEZİ

Yrd. Doç. Dr. Burçin KURT

TRABZON-2018

ONAY

Bu tez Yüksek Lisans/Doktora Tezi Standartlarına Uygun Bulunmuştur

Prof. Dr. Kemal TURHAN

Biyoistatistik ve Tıp Bilişimi Anabilim Dalı Başkanı

Karadeniz Teknik Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik ve Tıp Bilişimi Anabilim Dalı Yüksek Lisans öğrencisi Özlem KARADENİZ'in hazırladığı **“Pulmoner Emboli Klinik Tanı Sistemi”** başlıklı tez KTÜ Lisansüstü Eğitim - Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca kapsam ve bilimsel kalite yönünden değerlendirilerek Yüksek Lisans Tezi olarak kabul edilmiştir.

Danışman: Yrd. Doç. Dr. Burçin KURT

Yüksek Lisans Sınavı Jüri Üyeleri

Prof. Dr. Kemal TURHAN

Yrd. Doç. Dr. Burçin KURT

Doç. Dr. Ali Rıza AKYÜZ

Tarih: 26/01/2018

Bu tez KTÜ Sağlık Bilimleri Enstitüsü Yönetim Kurulu'nun/.../.... tarih ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Ali Osman KILIÇ
Sağlık Bilimleri Enstitüsü Müdürü

BEYAN

Bu tez çalışmasının KTÜ Sağlık Bilimleri Enstitüsü tez yazım kılavuzu standartlarına uygun olarak yazıldığını, tezin akademik ve etik kurallara bağlı kalınarak gerçekleştirilmiş özgün bir bilimsel araştırma eserim olduğunu, tezde yer alan ve bu tez çalışmasıyla elde edilmeyen tüm bilgi ve yorumlara kaynak gösterdiğimi ve kaynakların kaynaklar listesinde yer aldığını, tezin çalışılması ve yazımı aşamalarda patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

26/02/2018

Özlem KARADENİZ

TEŞEKKÜR

Bu tez çalışmasında yardımlarını esirgemeyen, bilgi ve tecrübelerini benimle paylaşan ve bana yol gösteren danışman hocam Yrd. Doç. Dr. Burçin KURT' a, teşekkür ederim. Bu süreçte büyük bir sabırla beni destekleyen her zaman yanımda olan başta annem ve babam olmak üzere, gösterdikleri yardım ve fedakarlık için abim Uzm. Dr. Serkan KARADENİZ ve eşi Uzm. Dr. Zehra Merve KARADENİZ' e ayrıca tezimi bitirmemde büyük katkıları olan sevgili kuzenim Uzm. Dr. Burcu KEMAL OKATAN' a ve bu tez süreci boyunca beni hiç yalnız bırakmayan kıymetli arkadaşlarım Handan YARDIM ve Büşra TARAKÇI başta olmak üzere tüm arkadaşlarıma sevgi, saygı ve teşekkürlerimi sunarım.

Özlem KARADENİZ

İÇİNDEKİLER

İç kapak sayfası	Sayfa
KABUL ve ONAY	
BEYAN	
TEŞEKKÜR	
TABLolar DİZİNİ	viii
ŞEKİLLER DİZİNİ	ix
KISALTMA, SİMGE ve FORMÜLLER	x
1. ÖZET	1
2. SUMMARY	2
3. GİRİŞ ve AMAÇ	3
4. GENEL BİLGİLER	5
4.1. Pulmoner Emboli	5
4.1.1. Kronik Pulmoner Emboli	5
4.1.2. Akut Pulmoner Emboli	5
4.2. Pulmoner Emboli Tanı ve Tedavi Yöntemleri	6
4.2.1. Pulmoner Emboli Laboratuvar Bulguları	7
4.3. Pulmoner Emboli'yi Tetikleyen Risk Faktörleri	8
4.4. Makine Öğrenmesi	8
4.4.1. Danışmanlı Öğrenme	9
4.4.2. Danışmansız Öğrenme	9
5. GEREÇ ve YÖNTEMLER	10
5.1. Yapay Sinir Ağı	10
5.1.2. Yapay Sinir Ağının Sınıflandırılması	12
5.1.2.1. İleri Beslemeli Yapay Sinir Ağı	13
5.1.2.2. Geri Beslemeli Yapay Sinir Ağı	14
5.1.3. Yapay Sinir Ağının Kullanım Alanları	14
5.1.4. Yapay Sinir Ağının Avantajları ve Dezavantajları	15
5.2. Karar Ağaçları	16
5.2.1. Karar Ağacı Algoritmaları	17
5.2.2. Karar Ağaçları Kullanım Alanları	18
5.2.3. Karar Ağaçlarının Avantajları ve Dezavantajları	18

5.3. Destek Vektör Makineleri	18
5.3.1. Doğrusal Ayrılabilen Veriler için DVM	19
5.3.2. Doğrusal Olarak Ayrılamayan Veriler için DVM	19
5.3.3. Destek Vektör Makinelerinin Kullanım Alanları	20
5.3.4. Destek Vektör Makinelerinin Avantajları ve Dezavantajları	20
5.4. Çalışmada Kullanılan Programlar	21
5.5. Verilerin Hazırlanması	21
6. BULGULAR	22
6.1. Veri Analizi	22
6.2. Karar Ağacı ile Modellemede Elde Edilen Bulgular	26
6.3. Destek Vektör Makineleri ile Modellemede Elde Edilen Bulgular	28
6.4. Yapay Sinir Ağları ile Modellemede Elde Edilen Bulgular	29
6.5. Geliştirilen Modellerin Değerlendirilmesi	30
7. TARTIŞMA ve SONUÇ	32
8. KAYNAKLAR	34
9. ETİK KURUL ONAY	38
10. ÖZGEÇMİŞ	39

TABLOLAR DİZİNİ

Tablo	Sayfa
Tablo 1. Pulmoner emboli tanısı alan hasta verilerinin cinsiyetine göre dağılımları	23
Tablo 2. Pulmoner emboli tanısı almayan hasta verilerinin cinsiyete göre dağılımları	23
Tablo 3. Pulmoner emboli tanısı alan hasta verilerinin tanımlayıcı istatistikleri	23
Tablo 4. Pulmoner emboli tanısı almayan hasta verilerinin tanımlayıcı istatistikleri	24
Tablo 5. Pulmoner emboli tanısı alan ve almayan hasta verilerinin karşılaştırılması	24
Tablo 6. Cinsiyet değişkenleri arasında Ki-Kare testi sonucu	25
Tablo 7. Geliştirilen karar ağacı modeline ait performans sonuçları	27
Tablo 8. DVM çekirdek fonksiyonlarının sonuçları	28
Tablo 9. Geliştirilen DVM modeline ait performans sonuçları	28
Tablo 10. Geliştirilen YSA modeline ait performans sonuçları	30
Tablo 11. Geliştirilen modellerin karşılaştırılması	31

ŞEKİLLER DİZİNİ

Şekil	Sayfa
Şekil 1. Pulmoner emboli tanısının basitçe gösterimi	7
Şekil 2. Makine öğrenmesi işlem akışı	9
Şekil 3. Biyolojik sinir hücresi yapısı	10
Şekil 4. Yapay sinir hücresi genel yapısı	11
Şekil 5. Çok katmanlı yapay sinir ağı	13
Şekil 6. Karar ağacı yapısı	16
Şekil 7. Doğrusal olarak ayrılabilen veriler için DVM	19
Şekil 8. Doğrusal ayrılamayan veriler için DVM	20
Şekil 9. Pulmoner emboli tanısı alan hastaların veri seti örneği	22
Şekil 10. Pulmoner emboli tanısı almayan hastaların veri seti örneği	23
Şekil 11. Eğitim seti örneği	25
Şekil 12. Test seti örneği	26
Şekil 13. Karar ağacı modeli	26
Şekil 14. Geliştirilen karar ağacı modeli için ROC eğrisi	27
Şekil 15. Geliştirilen DVM modeli için ROC eğrisi	29
Şekil 16. Yapay sinir ağı modeli	29
Şekil 17. Geliştirilen YSA modeli için ROC eğrisi	30

KISALTMA, SİMGE ve FORMÜLLER**Kısaltmalar**

AMI	Akut Miyokard Enfarktüsü
ANN	Artificial Neural Network
AUC	Area Under Curve
C&RT	Classification and Regression Tree
CHAID	Chi-Square Automatic Interaction Detector
DN	Doğru Negatif
DP	Doğru Pozitif
DVM	Destek Vektör Makinesi
EKG	Elektrokardiyografi
KTÜ	Karadeniz Teknik Üniversitesi
LVQ	Learning Vector Quantization
MLP	Multilayered Perceptron
PE	Pulmoner Emboli
ROC	Receiver Operating Characteristic
SLIQ	Supervised Learning in Quest
SVM	Support Vector Machine
YN	Yanlış Negatif
YP	Yanlış Pozitif
YSA	Yapay Sinir Ağı

1. ÖZET

Pulmoner Emboli Klinik Tanı Sistemi

Pulmoner emboli (PE) vücudun herhangi bir yerinde oluşan pıhtının koparak akciğerleri tıkanması durumudur. PE damarın tıkanıklığına göre ölüme neden olabilir. Genellikle sık karşılaşılan ölüm oranı yüksek ve kısa sürede müdahale edilmesi gereken bir hastalıktır. Ölüm oranı % 7 ile % 12 arasında değişmektedir. Bu yüzden PE’de erken ve doğru teşhis tedavide büyük rol oynamaktadır.

Son yıllarda makine öğrenmesi ve veri madenciliği yöntemleriyle özellikle tıp alanında birçok hastalığın sınıflandırılması, tahminlenmesi, risk hesaplanması gibi önemli çalışmalar yapılmaktadır. Bu tez çalışmasında makine öğrenmesi yöntemleri kullanılarak pulmoner emboli teşhisinde hekime yardımcı olabilecek bir klinik tanı sistemi geliştirilmesi hedeflenmiştir. Sistemin geliştirilmesi için yapay sinir ağı (YSA), destek vektör makinesi (DVM) ve karar ağacı yöntemleri kullanılmıştır.

Bu tez çalışmasında 2012-2017 yılları arasında KTÜ Tıp Fakültesi Farabi Hastanesi Acil Tıp Servisine nefes darlığı şikayeti ile başvuran ve pulmoner emboli tanısı alan ve pulmoner emboli tanısı almayan 240 hasta verisi kullanılmıştır. Kullanılan hasta verilerinin biyokimya laboratuvar test sonuçları uzman hekim yardımıyla incelenerek elde edilmiştir. Cinsiyet, yaş, D_dimer, PaCO₂, PaO₂ ve lokosit girdi verileri kullanılarak hekime yardımcı bir klinik tanı sistemi geliştirilmiştir. Geliştirilen DVM modeline göre % 91 duyarlılık, % 77 özgüllük, karar ağacı modeline göre % 84 duyarlılık, % 88 özgüllük ve YSA modeline göre % 93 duyarlılık, % 85 özgüllük değerleri elde edilmiştir. Bu tez çalışmasında geliştirilen YSA, DVM ve karar ağacı modellerine göre en iyi sonucu YSA modelinin verdiği görülmüştür.

Anahtar Sözcükler: Destek vektör makinesi, karar ağacı, klinik tanı sistemi, nefes darlığı, pulmoner emboli, yapay sinir ağı

2. SUMMARY

Pulmonary Embolism Clinical Support System

Pulmonary embolism (PE) is a condition in which the clot that forms anywhere in the body breaks and blocks the lungs. PE can cause death according to occlusion of the vein. Generally, it is common disease which has high mortality rate and must be cured in short time. Mortality rate ranges from 7 % to 12 %. So early and correct diagnosis of PE plays a major role in the treatment.

In recent years, important studies such as classification, estimation and risk calculation of many diseases have made especially in the field of medicine with machine learning and data mining methods. This thesis study, it is aimed to develop a clinical diagnosis system which can help the diagnosis pulmonary embolism using machine learning methods. Artificial neural network (ANN), support vector machine (SVM) and decision tree methods are used for system development.

In this thesis study, 240 patients who applied with the complaint of dyspnea to Farabi Hospital Emergency Medicine Department of KTU Medical Faculty between the years of 2012-2017 and who did and did not receive a pulmonary embolism were used. Patient data obtained were obtained by examining the biochemical laboratory test results with help of specialist physician. Clinical diagnostic system has developed to aid in the treatment of patients using sex, age, D-dimer, PaCO₂, PaO₂ and leukocyte input data. According to developed SVM model, 91 % sensitivity, 77 % specificity, according to decision tree model, 84 % sensitivity, 88 % specificity and and 93 % sensitivity, 85 % specificity according to ANN model have obtained. In this thesis study, the best result was given by ANN model to ANN, SVM and decision tree models have developed.

Key Words: Artificial neural network, clinical diagnosis system, decision tree, dyspnea, pulmonary embolism, support vector machine

3. GİRİŞ ve AMAÇ

Bilimsel ve teknolojik gelişmeler insan hayatını kolaylaştırmaktadır. Bu sayede insanlar özellikle sağlık alanındaki gelişmelerle daha rahat bir hayat sürdürmektedir. Son yıllarda makine öğrenmesi ve veri madenciliği yöntemleriyle özellikle tıp alanında hekimlere yardımcı olmak amacıyla birçok hastalığın sınıflandırılması, tahminlenmesi, risk hesaplanması gibi önemli çalışmalar yapılmaktadır.

Pulmoner emboli sıklıkla karşılaşılan kardiyovasküler bir durumdur. Pulmoner emboli tanısını koymak oldukça güçtür ve belirli bir klinik tablo ile sonuçlanmadığı için tanı koyulması gözden kaçabilmektedir. Bu yüzden, erken tedavi oldukça önemli olması nedeniyle, erken tanı konulması da önem arz etmektedir (1). Pulmoner emboli yüzünden meydana gelen ölüm oranı tüm hastane ölüm oranlarının % 10'unu oluşturmaktadır. Tedavi edilemezse bu ölüm oranı % 25 ile % 30 arasında, tedavi edildiği halde ölüm oranı % 2 ile % 8 oranında değişmektedir (2). PE nedeniyle ölen hastaların çoğu tanı konulamamış ve bu yüzden tedavi başlanılamamış hastalardır. Bu nedenle yapılan çalışmada pulmoner emboli ele alınmıştır.

Serpen ve arkadaşları tarafından 2008 yılında yapılan bir çalışmada pulmoner emboli teşhisi için yapay sinir ağı modeli geliştirilmiştir. 30 test örneği üzerinde 10 çapraz doğrulama yöntemi kullanılmıştır. Bu yöntem sonucunda naive Bayes ile % 73.3, Bayesian belief network ile % 50.0, çok katmanlı perceptron ile % 60.0, c4.5 karar ağacı ile % 63.3, karar destekli uyarlanabilir yükseltme m1 algoritması ile % 33.3 ve rastgele ağaç kullanarak torbalama (bagging) meta sınıflayıcı ile % 56.7 doğruluk değeri elde edilmiştir (3).

Agharezaei ve arkadaşları tarafından (4) 2014 yılında yapılan bir çalışmada farklı parametreler kullanılarak pulmoner emboli mortalite oranını azaltma amacıyla risk faktörlerini belirleyen bir karar destek sistemi geliştirilmiştir. Ergün ve arkadaşları tarafından 2012 yılında yapılan bir başka çalışmada hastanın demografik bulguları, klinik laboratuvar sonuçları ve bu sonuçların emboli yerleşim yeri ile ilişkisini belirlemede istatistiksel analiz yöntemleri kullanılarak tedavi yanıtları değerlendirilmiştir. Bunun sonucunda mortaliteyi azaltmada izlenecek yollara katkı sağlanmıştır (5).

Bu tez çalışmasında karar ağacı, yapay sinir ağı, destek vektör makinesi gibi bazı makine öğrenmesi yöntemleri kullanılarak pulmoner emboli klinik tanı sistemi geliştirilmiştir. Günümüzde acil servislere nefes darlığı şikayetiyle gelen hastalardan ilk önce biyokimya laboratuvar test bulguları istenmektedir. Bu test bulgularından olan D_dimer pulmoner emboli tanısında çok önemli bir parametredir. D_dimer'ın yüksek çıkması halinde pulmoner emboli hastalığından şüphelenilmektedir. Bunun yanı sıra PaCO₂ ile PaO₂ değerlerinin düşüklüğü ve lokosit değerinin yüksekliği de D_dimer yüksekliği ile birlikte pulmoner emboli olasılığını oldukça kuvvetlendirmektedir. Ancak bazı durumlarda pulmoner emboli tanısında çok önemli bir parametre olan D_dimer düşük çıktığı halde pulmoner emboli tanısı alan hastalar da mevcuttur ve kesin tanı bilgisayarlı tomografi görüntüleme yöntemi ile konulmaktadır. Pulmoner embolide erken tanı oldukça hayati önem taşımaktadır. Bu yüzden yapılan tez çalışmasında pulmoner emboli tanısı için bilgisayarlı tomografiyi kullanmadan biyokimya laboratuvar test bulgularından D_dimer, PaCO₂, PaO₂ ve lokosit verileri ile demografik veri olan cinsiyet ve yaş verileri kullanılarak erken tanı için hekime yardımcı olabilecek akıllı bir pulmoner emboli klinik tanı sistemi geliştirilmiştir.

4. GENEL BİLGİLER

4.1. Pulmoner Emboli

Pulmoner emboli, kalp krizi ve inmeden sonra gelen kardiyovasküler hastalıkların en yaygın üçüncü ölüm nedenidir (26). Pulmoner emboli, pulmoner arteriyel yatağın tam veya kısmi tıkanmasından dolayı oluşan ve hayatı tehdit edebilen önemli bir hastalıktır. Bundan dolayı pulmoner emboli şüphesi olan hastalarda klinik, laboratuvar sonuçları ve ilgili görüntüleme teknikleriyle elde edilen sonuçlar dikkatle incelenmelidir (5).

ABD (Amerika Birleşik Devletleri) bazlı bir çalışmaya göre 20 yıllık süreçte 42 milyondan fazla ölüm gerçekleşmiş ve bu ölümlerin yaklaşık altı yüz bini (% 1.5) PE tanısı almıştır. Bu sayılar elbette ki gerçek PE insidansını göstermemektedir, çünkü muhtemelen PE vakalarının yarısı tanı alamamaktadır (27). Türkiye’de pulmoner emboli de ölüm oranı % 7 ile % 12 arasında değişmektedir. Bu yüzden PE de erken ve doğru teşhis tedavide büyük rol oynamaktadır. Pulmoner emboli akut ve kronik olarak sınıflandırılmıştır.

4.1.1. Kronik Pulmoner Emboli

Kronik PE’ de ise yıllar sonra gelişen pulmoner hipertansiyon sonucu progresif solunum sıkıntısı ile kendini gösterir (27). Kronik pulmoner, pulmoner embolinin nadir görülen ancak önemli ölüm oranı olan bir komplikasyonudur. Ana tedavi yöntemi cerrahidir (28).

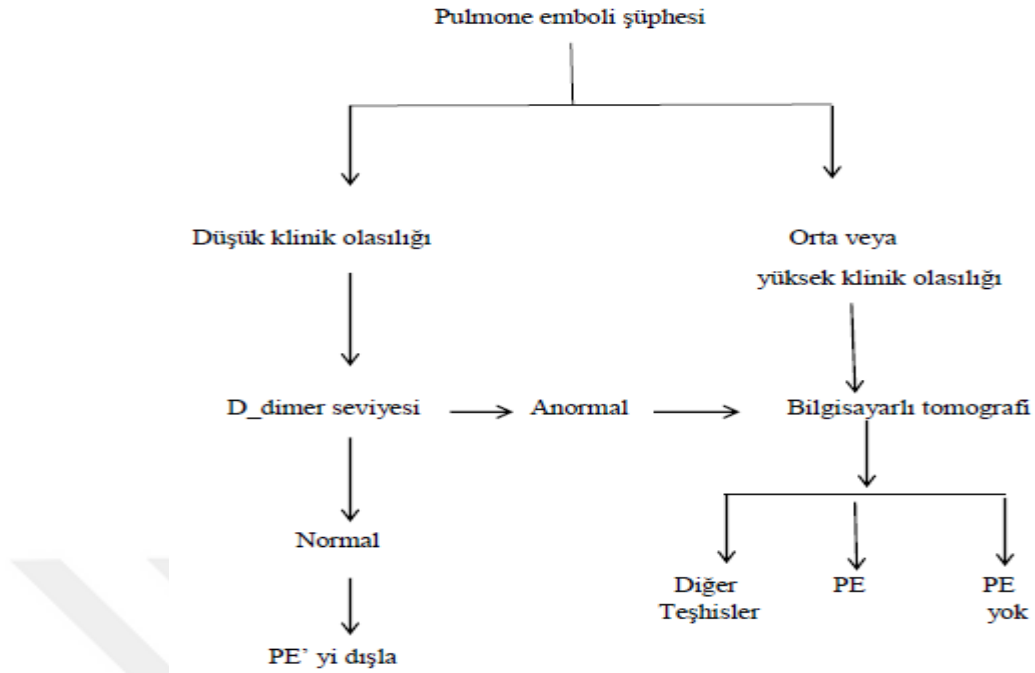
4.1.2. Akut Pulmoner Emboli

Akut pulmoner emboli sıklıkla meydana gelen ve ölümcül sonuçlar doğuran bir hastalıktır. Erken teşhis ve tedavi ile bu ölüm oranında azalma sağlanabilir. Tedavi ile % 2 ile % 8 arası olan ölüm oranı, tedavisiz yaklaşık % 30'lara kadar çıkmaktadır. Bu ölümlerin çoğu olayın başlamasından sonraki ilk birkaç saatte olan tekrarlayan pulmoner emboli kaynaklıdır. Klinik semptomlar, bulgular ve laboratuvar testleri pulmoner emboli teşhisinde düşük duyarlılık ve özgüllüğe sahiptirler. Bundan dolayı pulmoner emboli teşhisini koymak veya dışlamak için tanı testlerine ihtiyaç duyulmaktadır (29).

4.2. Pulmoner Emboli Tanı ve Tedavi Yöntemleri

Pulmoner emboli acil serviste sık karşılaşılan ve tanıdan sonraki ilk üç aylık süreçte ortalama % 10 ile % 15 arasında ölümle sonuçlanan kardiyovasküler ölümlerin önde gelen nedenleri arasında gösterilmektedir. Öncelikle pulmoner emboli tanısında yaşanan güçlükler, hayati riskin en başta gelen nedenidir. Çünkü tanı konulmamış veya yanlış bir tanı üzerinde durulurken kaybedilen hastalar hayli çoktur. Tanı aşamasında yaşanan güçlüklerin de başında hastanın her zaman tanıyı sağlayabilecek teknik ekipmana sahip bir merkeze başvurmaması veya bulunduğu bölgede o tür imkanlara sahip kliniklerin olmamasıdır. PE de akut miyokard enfarktüsü (AMI) hastası gibi acildir ve hayati risk içermektedir. Bu nedenle hızlı tanı ve tedavinin hastanın prognozunda önemli yere sahip olduğu PE den şüphelenilen hastalarda tanısal amaçlı teknikler uygun algoritmalar kullanarak değerlendirilmelidir (36). PE şüphesi olan hastada ilk değerlendirmede elektrokardiyografi, akciğer filmi ve ekokardiyografi ile birlikte arter kan gazı, beyin natriüretik peptid, troponin ve D-dimer testleri yapılmalıdır. Bu tetkikler ayırıcı tanı için önemlidir. PE'de kesin tanı için kullanılan görüntüleme yöntemleri arasında ventilasyon-perfüzyon sintigrafisi, pulmoner kateter anjiyografi, bilgisayarlı tomografi (BT) pulmoner anjiyografi, manyetik rezonans (MR) pulmoner anjiyografi yer almaktadır. BT pulmoner anjiyografi PE şüphesi olan hastalarda pulmoner damarları göstermede günümüzde en çok kullanılan ve tercih edilen tanı yöntemidir. (30).

Klinik olarak pulmoner emboli olasılığı orta veya yüksek ise tanısal işlemler devam ederken bir an önce tedaviye başlanılmalıdır. Hastaya ait özel bir durum yoksa hastanın tedavisi en az üç ay sürmelidir (32). PE tanısına ilişkin şema aşağıdaki şekilde verilmiştir.



Şekil 1. Pulmoner emboli tanısının basitçe gösterimi (34)

4.2.1 Pulmoner Emboli Laboratuvar Bulguları

D_dimer: Spesifik bir fibrin yıkım ürünüdür. Kanseri, hepatik ve renal yetmezlik, sepsis, inme, travma durumlarında da yükselir. Klinik olasılığın yüksek olmadığı düşük veya orta olduğu durumda pulmoner emboli şüpheli olgularda D_dimer'in negatif çıkması sonucu hastalığı dışlar. Yüksek olasılık durumunda D_dimer testlerinin hiçbirisi PE'nin dışlanmasında kullanılamaz (31).

PaCO₂ ve PaO₂: PE olgularında düşük PaO₂ ve PaCO₂ değerleri saptanmaktadır. Hastalarda % 10 ile % 25 değerleri arasında arter kan gazları normal şekilde bulunmaktadır. PaO₂'si normal bulunan hastaların çoğunda alveolar-arteriyel O₂ farkı genellikle 20 mmHg'nin üzerindedir. Nedeni açıklanamayan hipoksemi ile birlikte normo/hipokapni saptandığında PE açısından ileri tetkik yapılmalıdır. Ayrıca arteriyel kan gazı PE şiddetini belirlemede ve tedaviye cevabı izlemede yararlanılır (31).

Lökosit: Pulmoner emboli lökosit hızında belirli bir artış olmaktadır. Bu durumda lökositin pulmoner emboli tanısında bir nebze etkili olduğu görülmektedir. Normal lökosit değeri 4800-10800 değer aralığındadır.

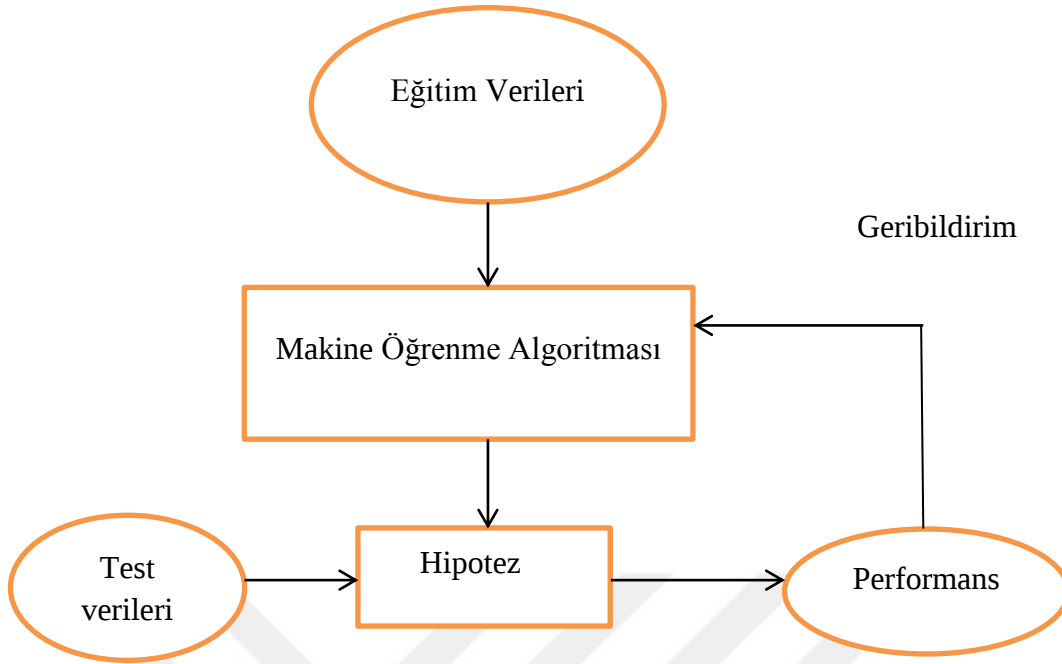
4.3. Pulmoner Emboli'yi Tetikleyen Risk Faktörleri

Pulmoner emboliye zemin hazırlayan faktörler hastalıkla ve koşullarla alakalı olmak üzere iki gruba ayrılır. Hasta ile alakalı risk faktörleri kronik kalp ya da solunum yetersizliği, hormon replasman tedavisi, malignite, oral kontraseptif kullanımı, paralitik inme, venöz tromboemboli, ileri yaş, obezite ve varislerdir. Koşullarla ilgili risk faktörleri ise kalça yada bacak kırığı, kalça veya diz eklemi replasmanı, majör genel cerrahi, majör travma, artroskopik diz cerrahisi, santral venöz kateter, kemoterapi, gebelik/doğum sonrasıdır (33).

4.4. Makine Öğrenmesi

Bilgisayar sistemlerinde, bilgisayarları programlamak meydana gelen problemlere bazen çözüm getirmeyebilmektedir. Girdi verilerine karşılık, çıkışı elde etmenin belli bir yöntemi olmayabilir. Gerekli olan yöntem açık bir şekilde ifade edilmediği zaman, örnekler bilindik programlama yaklaşımının yerine deneyimlerden faydalanarak çözülmeye çalışılmaktadır. Bilgisayarın öğrenmesi ise giriş ve çıkış işlevselliğini örneklerden öğrenmesidir. Bundan dolayı programları sentezlemek için örnek kullanımına makine öğrenme yöntemi denir (6).

Yapay zekanın bir türü olan makine öğrenmesi, verileri akıllıca hareket ettirmek için bilgisayar algoritmalarının geliştirilmesi ile ilgilenen çalışma alanıdır. Bu alan mevcut verilerin istatistiksel yöntemlerin ve bilgi işlem gücünün hızla ve eş zamanlı evrildiği bir ortamdan gelmektedir. Makine öğrenmesi yaygın olarak kullanılmasına ve çok büyük bir potansiyele sahip olmasına rağmen, sınırlarını anlamak önemlidir. Makinelerin öğrendiği parametrelerin dışına çıkarmak için çok az esneklik payı vardır. Makine öğrenmesi, şu an insan beyninin yerini alması söz konusu değildir. İnsan ya da makine olmasına bakılmaksızın öğrenme süreci ikisinde de benzerlik göstermektedir. Öğrenme süreci birbiriyle ilişkili öğelerden oluşmaktadır (7). Aşağıdaki şekilde makine öğrenmesine ait işlem akış şeması verilmiştir.



Şekil 2. Makine öğrenmesi işlem akışı (37)

Makine öğrenmesi çeşitleri amaçlarına göre kategorilere ayrılır (7). Girdi ve çıktı değerleri önceden biliniyorsa danışmanlı öğrenme (Supervised learning), girdi verilerine karşı herhangi bir çıktı verisi yoksa yani girdi verisi önceden bilinmiyorsa danışmansız öğrenme (Unsupervised learning) çeşitleri kullanılmaktadır (6).

4.4.1. Danışmanlı Öğrenme

Danışmanlı öğrenme, eğitilmiş veriler üzerinden gerçekleştirilir. Danışmanlı öğrenme yönteminde girdi ve çıktı verileri önceden belli olmaktadır. Bundan dolayı girdi ve çıktı verileri arasındaki ilişki ile makinenin işleyişi gerçekleştirilmektedir. Bu öğrenme yapısının sonucuna göre de makineye yeni girdi verisi geldiğinde çıktı verisi olabilecek tahminlerde bulunarak, sınıflandırma yapılabilmektedir (6).

4.4.2. Danışmansız Öğrenme

Danışmansız öğrenmede sadece girdi verileri vardır (8). Girdi verilerine karşılık herhangi bir çıktı verisi yoktur. Hiçbir veri girişinin herhangi bir sınıfa ait olduğu bilinmemektedir. Bu öğrenme yöntemi, sınıfları bilinmeyen girdi kümesini parçalayarak girdi verilerinin gruplandırılmasını veya kümelenmesinde rol oynamaktadır (6).

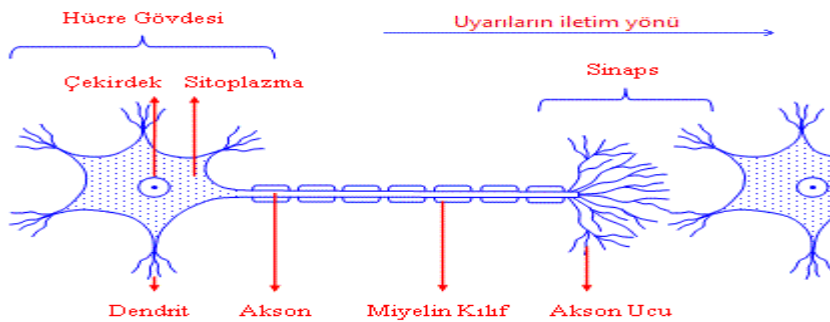
Tez çalışmasında modelin geliştirilmesi için makine öğrenmesi yöntemlerinden YSA, DVM ve karar ağacı yöntemleri kullanılmıştır.

5. GEREÇ ve YÖNTEMLER

5.1. Yapay Sinir Ağı

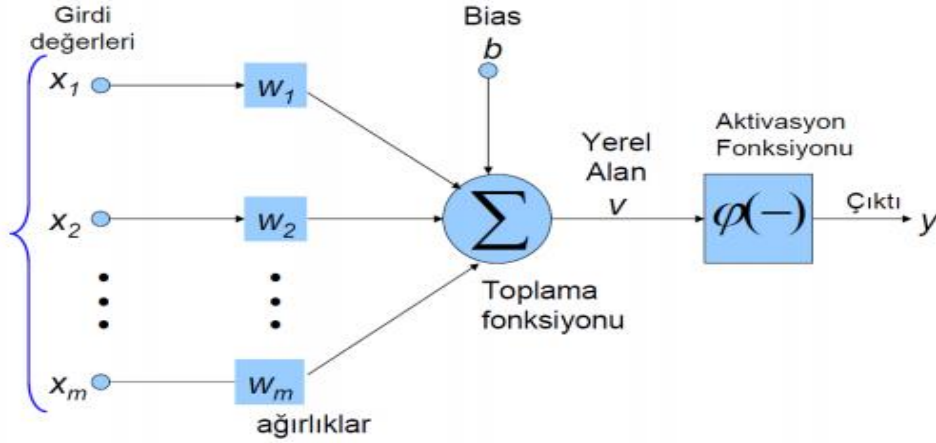
Yapay sinir ağı (YSA), insan beynine benzer özelliklerinden oluşan öğrenme yolu ile yeni bilgiler üretebilme, oluşturabilme ve keşfedebilme gibi yetenekleri herhangi bir yardım olmaksızın otomatik olarak gerçekleştirmek amacı ile geliştirilmiştir. Yapay sinir ağı, öğrenmenin haricinde bilgiler arasında ilişkiler kurma yeteneğine de sahiptir (9).

Yapay sinir ağı bilinen hesaplama yöntemlerinin haricinde farklı bir hesaplama yöntemi sunmaktadır. Bulundukları ortama uyum sağlayan, eksik bilgi ile de çalışabilen, belirsizlikler durumlar altında karar verebilen, hatalara karşı tolerans sahibi olan bu hesaplama türünün hayatın neredeyse her alanında başarılı uygulamalarını görmek mümkündür. Oluşturulacak olan ağın yapısının belirlenmesinde, ağ parametrelerinin tercihinde, belirli bir standardın olmaması durumunda, problemlerin sadece nümerik bilgiler ile gösterilebilmesi, eğitimin nasıl bitirileceğinin bilinmemesi ve ağın davranışlarını açıklayamamasına rağmen bu ağlara olan ilgi her geçen gün artmaktadır. Özellikle, sınıflandırma, örüntü tanıma, sinyal filtreleme, veri sıkıştırma ve optimizasyon gibi birçok çalışmalarda yapay sinir ağı en güçlü teknikler arasında gösterilmektedir. Veri madenciliği, optik karakter taşıma, optimum rota belirleme, parmak izi tanıma, malzeme analizi, kalite kontrol ve tıbbi analizler gibi pek çok alanda başarılı örneklerine rastlamak mümkündür (10). Bir yapay sinir ağı hücresi insan beyindeki sinir hücrelerinden esinlenerek geliştirilmiştir.



Şekil 3. Biyolojik Sinir Hücresi Yapısı (38)

Biyolojik sinir hücrelerinde aksonlar gelen bilgiyi başka hücrelere gönderirler. Bu sayede hücreler birbirleri ile iletişim kurarlar. Yapay sinir hücreleri de buna benzer yapı göstermektedir.



Şekil 4. Yapay sinir hücresi genel yapısı (11)

$x_1, x_2, \dots, x_m$ girdi verileri sisteme verilen örnek bilgilerdir. Bu bilgiler başka bir sinir hücresinden ya da dış ortamdan gelebilmektedir. $w_1, w_2, \dots, w_m$ nöronlar arasındaki bağlantıların matematiksel katsayı değerleridir. Her bir girdi değerinin kendine özgü bir ağırlığı bulunmaktadır. YSA'ya örnekler tanıtıldıkça ağırlık değerleri değişmektedir. YSA'nın çalışma mantığına göre ağırlık değeri verilen örnek girdiler için en doğru çıktıları üretebilecek ağırlıkların bulunması esastır. Eşik değeri, eğitilecek ağırlık ezberlemesini önlemek için kullanılan değerdir. Toplama fonksiyonu, nöronlara gelen her bir girdinin sahip olduğu ağırlık değerleri ile çarpılarak toplanmasıyla ve eşik değerinin eklenmesiyle elde edilen fonksiyondur. Bu sayede nöronlar için net girdi hesaplanmaktadır (11).

$$y(x) = \sum_{i=1}^m (w_{ij} x_i + b) \quad (5.1)$$

Aktivasyon fonksiyonu, net girdinin aktivasyon fonksiyonundan geçirilmesiyle net çıktı hesaplanmaktadır. YSA'da çeşitli aktivasyon fonksiyonları bulunmaktadır. İncelenen problem türü ve kurulacak ağ modeline göre uygun aktivasyon fonksiyonu seçilerek kullanılmalıdır. Çıktı, önceki katmanlardan gelen bilgi işlenerek dış ortama gönderilen değerleridir. Bir nöronun çıkışı yalnızca bir tane çıktı değeri bulunmaktadır (11).

Aktivasyon fonksiyonunun doğru seçilmesi, ağın performansını etkileyen bir unsurdur. Seçilen aktivasyon fonksiyonu doğrusal değilse, eğim parametresinin belirlenmesi lazımdır. Eğim parametresi, uygun sonuca yeterli derecede ulaşmasında rol oynayan önemli bir faktördür (12). Günümüzde çoğunlukla aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmaktadır. Sigmoid fonksiyonu doğrusal olmayan problemlerin çözümünde kullanılan yapay sinir ağına kullanılması ile birlikte türev bilgisine ihtiyaç duyulan durumlar için de olanak sağlamaktadır. Bu fonksiyon sürekli ve türevi alınabilir bir fonksiyondur. Doğrusal olmadığından YSA’ da sıklıkla kullanılır. Fonksiyon girdi verilerinin her biri için sıfır ve bir arasında değer üretmektedir (13).

Yapay sinir ağına en ayırt edici özellik öğrenme yeteneğidir. Yapay sinir ağına öğrenme, istenilen bir durumu yerine getirecek bir şekilde ağırlıkların kullanılması ile yapılmaktadır. Ağırlıkların ayarlanması, probleme ait örnekler yardımı ile yapılmaktadır. Ağ, gördüğü her örnek problem için ağırlıklarını değiştirmektedir. İşlem sonucunda ağı doğru ağırlık değerlerine ulaşması, başka ifadeyle, örneklerin temsil ettikleri olay ile ilgili genelleme yapabilecek duruma gelmesiyle son bulmaktadır. Ağırlıkların değiştirilmesi, kullanılacak olan öğrenme stratejisine ve öğrenme kuralları ile yapılmaktadır. Yapay sinir ağına öğrenme algoritmaları, danışmanlı, danışmansız ve takviyeli stratejiler olmak üzere üç gruba ayrılır. Danışmanlı öğrenmede, ağa belirli girdiler için çıktı değerleri de verilir. Ağ, girdiyi süreçten geçirerek kendi çıktı değerini üretir ve gerçek çıktı değeri ile karşılaştırır. Danışmansız öğrenmede, çıktıların istenen değerleri ağa gösterilmiyorsa, bu tip öğrenme şekli danışmansız öğrenmedir. Danışmansız öğrenmede, girdiler aynı zamanda çıktı görevi görmektedir. Danışmansız öğrenmede, danışmanlı öğrenmedeki gibi arzu edilen çıkış değerleri bilinmediğinde girdi değerlerinin verdiği cevaplar gözlemlenerek başarıya ulaşabilmektedir. Takviyeli öğrenme ise, genelde danışmanlı öğrenme yöntemine benzer ama ağa hedef çıktılar yerine ağı çıktılarının ne ölçüde doğru olduğunu belirten bir derece verilir (13).

5.1.2. Yapay Sinir Ağına Sınıflandırılması

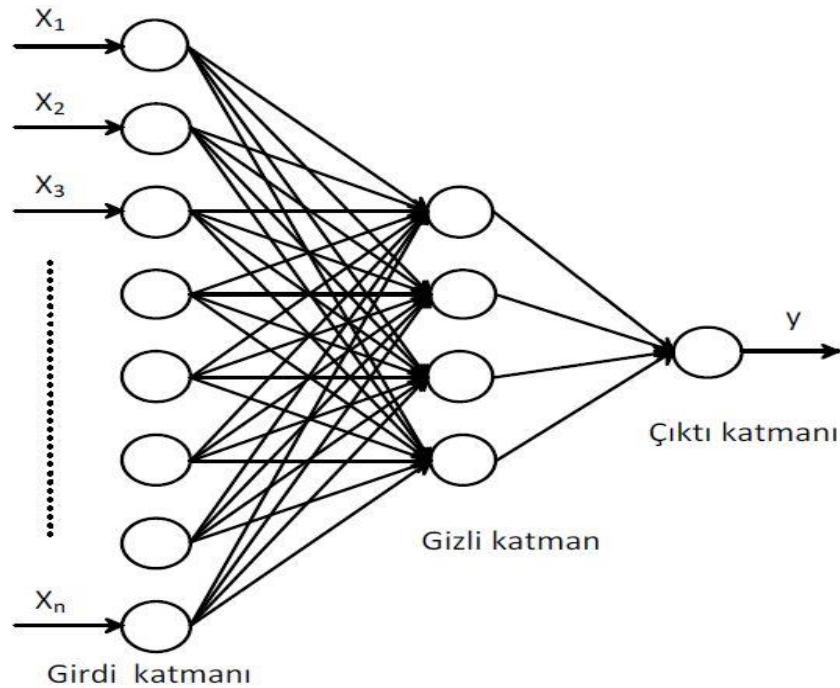
Yapay sinir ağı, sinir hücrelerinin birbirlerine değişik şekillerde bağlanmalarından oluşmaktadır. Hücre çıkışları, ağırlıklar üzerinden başka hücelere veya kendisine girdi olarak bağlanabilir ve bağlantılarda gecikme de kullanılabilir. Hücrelerin bağlantı şekillerine, öğrenme kurallarına ve aktivasyon fonksiyonlarına göre çeşitli modeller

geliştirilmiştir. İçerisinde birçok nöron barındıran yapay sinir ağı, belli topolojiler ile tanımlanabilir (14).

Yapay sinir ağı, ileri beslemeli (feedforward) ve geri beslemeli (feedback) olmak üzere iki farklı yapıda incelenebilir (14).

5.1.2.1. İleri Beslemeli Yapay Sinir Ağı

İleri beslemeli ağda, işlemci elemanlar çoğunlukla katmanlara ayrılmışlardır. İşaretler, girdi katmanından çıktı katmanına doğru tek yönlü bağlantılar ile iletilmektedir. İşlemci elemanlar, bir katmandan diğerine bağlantı kurarken, aynı katman içerisinde bağlantıları bulunamaz. İleri beslemeli ağlara, çok katmanlı algılayıcılar (MLP) ve sayısallaştırılmış öğrenme vektörü (LVQ) ağları örnek olarak verilebilir. Blok olarak ifade edilen ve oldukça yaygın kullanım alanı bulunan çok katmanlı algılayıcı sinir ağı (MLP); bir giriş katmanı, bir veya birden çok gizli katman ve bir çıkış katmanından oluşmaktadır. Her katmanında, bir veya birden fazla işlemci eleman bulunabilmektedir (14).



Şekil 5. Çok katmanlı yapay sinir ağı (39)

5.1.2.2. Geri Beslemeli Yapay Sinir Ağı

Geri beslemeli yapay sinir ağında, en az bir işlemci elemanın çıktısı, kendisine veya diğer işlemci elemanlara girdi olarak verilmektedir ve çoğunlukla geri besleme, bir geciktirme elemanı üstünden yapılmaktadır. Geri besleme, bir katmandaki işlemciler arasında olabileceği gibi katmanlar arasındaki işlemciler arasında da olabilmektedir. Bu yapısı sayesinde geri beslemeli yapay sinir ağı, doğrusal olmayan dinamik bir davranış göstermektedir. Geri beslemenin yapılış durumuna göre farklı yapı ve davranışta geri beslemeli yapay sinir ağı elde edilebilmektedir. Geri beslemeli yapay sinir ağı, karmaşık bir çalışma düzeneğine sahip olmasına rağmen, dinamik hafızaları nedeniyle ön tahmin uygulamalarında başarılı sonuçlar vermektedir. Geri besleme genellikle bir geciktirme işlemci elemanı üzerinden yapılır veya bir katmandaki hücreler arasında olduğu gibi katmanlar arasındaki hücreler arasında da olabilmektedir (14).

5.1.3. Yapay Sinir Ağının Kullanım Alanları

Yapay sinir ağı problem çözümünde matematiksel bir model veya algoritma olmaması durumlarında genellikle kullanılmaktadır. YSA'lar doğrusal olmayan, çok boyutlu ve karmaşık, gürültülü veya eksik veriler içeren veri kümelerinin incelendiği analizlerde de başarılı sonuçlar üretmesi sebebiyle tercih edilen bir yöntemdir. Genel olarak teşhis, sınıflandırma, örüntü tanıma, tahmin, zaman serileri analizi, optimizasyon ve doğrusal olmayan sistemlerin modellenmesinde kullanılan YSA'nın bir çok alanda kullanımı yaygındır (11).

- Sağlık: İlaç etkileri analizi, kalp krizi erken teşhis ve tedavisi vb.
- Finans: Pazar performans analizi.
- Bankacılık: Müşteri analizi, kredi başvurularını değerlendirme.
- Öngörü ve Tahmin Çalışmaları: Hava durumu tahminleme, bir ürünün gelecekteki satış tahmini.
- Robotik: Optimum rota belirleme, uzaktan kumandalı sistemler vb.
- Ekonomi: Ekonomik değişkenlerin modellenerek tahmini.
- Güvenlik: Parmak izi, retina vb. tanıma, yüz eşleştirme gibi uygulamalar.
- Telekomünikasyon: Ses ve görüntü işleme.
- Dil: Sözcük tanıma, dil tercümesi vb.

5.1.4. Yapay Sinir Ağı'nın Avantajları ve Dezavantajları

Yapay sinir ağı kullanılarak elde edilebilecek avantajlardan bazıları aşağıdaki gibi sıralanmaktadır (14) :

- Uygulamaları hem pratiktir ve maliyet açısından ucuzdur.
- Zaman bakımından verimli çalışmaktadır.
- Yapay sinir ağı yeni bilgilerin ortaya çıkması durumunda ve ortamda bazı değişikliklerin olması durumunda tekrardan eğitilebilir.
- Yapay sinir ağı modeli, her biri problemin farklı parçası ile ilgilenen çok sayıda işlemci elemanından oluşması ve bağlantı ağırlıklarının ayarlanabilir olması gibi niteliklerinden dolayı da esnek bir yapıya sahiptir. Bu yapısından ötürü, ağı bir kısmı zarar görse dahi model yine de çalışmaya devam eder. Bu durum sadece performans düşüklüğüne sebep olur, fakat modelin işlevini tamamen yitirmesi durumu olmamaktadır.
- Yapay sinir ağı, giriş bilgilerine göre kendi ilişkilerini oluşturmaktadır. Ateorik modeller olduklarından denklem içermemektedir. Bununla birlikte modelin içinde ne olduğu tam olarak tahmin edilemez. Yapay sinir ağı, bu karışık özelliklerinden dolayı, bazıları için akıllı bir “kara kutu” olarak tanımlanmaktadır.

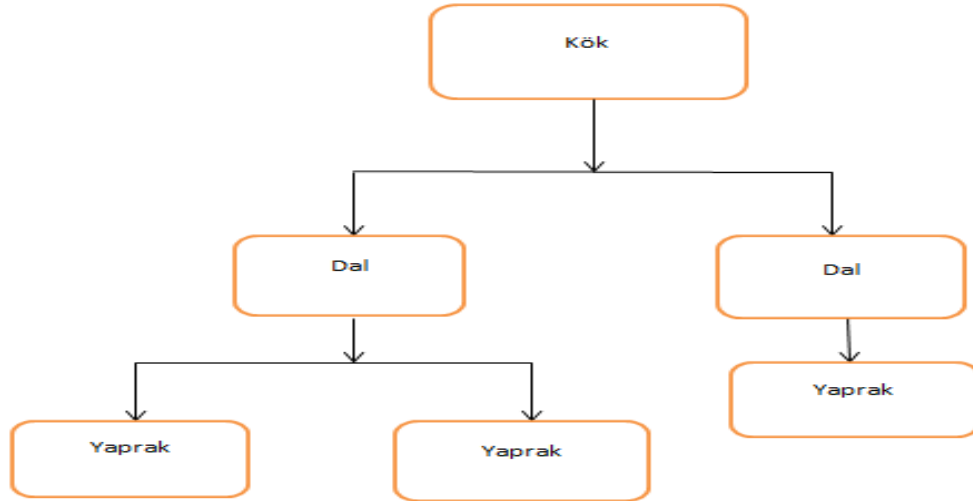
Yapay sinir ağı'nın bazı dezavantajları aşağıdaki gibidir (14):

- Ağı eğitiminin ne zaman bitirileceğine karar vermek için de geliştirilmiş bir yöntem bulunmamaktadır.
- Ağı davranışlarının açıklanamaması durumu vardır. Bir probleme çözüm üretildiği zaman bunun nasıl ve neden üretildiği konusunda bir bilgi bulmak mümkün olmamaktadır. Bu durum ağı ürettiği çözüme karşılık güven azalmaktadır.
- Problemlere optimum sonuçları garanti etmez.
- Yapay sinir ağı'nın öğrenme süresi uzun olmaktadır. Yapay sinir ağını eğitebilmek için çok fazla denemeye ihtiyaç duyulmaktadır. Eğitme zamanının kısaltılması ise kritik bir durum oluşturmaktadır. Bundan dolayı bir araştırmacı sınırlı bir zaman diliminde ne kadar fazla deneme yaparsa sonuçtan o kadar emin olacaktır.

- Yapay sinir ağı, kabul edilebilir çözümler üretmektedir fakat en iyi çözümü garanti etmemektedir.
- Ağın öğreneceği problemin ağa gösterimi de önemli bir dezavantajdır.

5.2. Karar Ağaçları

Veri madenciliği ile tahmin modellerinin geliştirilmesinde, yorumlanmasının kolay olmasında, veri tabanı sistemlerine kolayca entegre edilebilmesinde, güvenilirliklerinin daha iyi olmasından dolayı sınıflama modelleri arasında en yaygın ve etkin olarak kullanılan yöntemlerden biridir (15). Bir karar ağacı, basit karar verme adımları kullanarak, büyük miktarlardaki kayıtları, çok küçük kayıt gruplarına ayırarak kullanılan bir yöntemdir. Büyük veri tabanlarının kullanıldığı pek çok sınıflama sorunlarında karmaşık veya hata içeren bilgilerde karar ağaçları yararlı bir çözüm olabilmektedir (16).



Şekil 6. Karar ağacı yapısı (18)

Karar ağacı, ağaç şeklinde görünen bir modeldir. Ağacın her dalı bir sınıflandırma sorusu ve yaprakları da veri setinin bu sınıflandırmaya ait parçalarıdır (17). Karar ağacının yapısında bulunan her bir düğüm belirli bir niteliği temsil eder. Dallar ve yapraklar ise ağaç yapısının birer elemanlarıdır. En üstteki eleman kök, en alttaki eleman yaprak ve bunların arasında kalan elemanlar ise dal olarak nitelendirilir. Karar ağacında kurallar kökten yaprağa doğru inilerek yazılmaktadır. Karar ağacı yönteminde bir olayın sonuçlandırılması için sorunun cevabına göre hareket edilmektedir (18).

5.2.1. Karar Ağacı Algoritmaları

Veri madenciliğinde karar ağacı oluşturmak için çeşitli algoritmalar geliştirilmiştir (16). Geliştirilen bazı algoritmalara CHAID, C&RT, C5.0 ve C4.5, SLIQ, SPRINT örnek verilebilir.

C&RT: Gini'ye dayalı ikili bölme işlemidir. Son veya uç olmayan her bir düğümde iki adet dal bulunmaktadır. Budama işlemi ağacın karmaşıklık ölçüsüne dayanır. Sınıflandırma ve regresyonu destekleyici bir yapıdadır. Sürekli hedef değişkenleri ile çalışır. Verinin hazırlanmasına gereksinim duyar (15).

CHAID: Ki-kare testleri kullanarak bölme işlemini gerçekleştirmektedir. Dalların sayısı iki ile tahmin edicinin kategori sayısı arasında değişmektedir (15).

C5.0 ve C4.5: ID3 karar ağacı algoritmasının ileri versiyonudur. Her düğümde çıkan çoklu dallar ile ağaç oluşturur. Dalların sayısı tahmin edicinin kategori sayısına eşittir. Tek bir sınıflayıcıda birden çok karar ağacını birleştirir. Ayırma işlemi için bilgi kazancı kullanır. Budama işlemi her yapraktaki hata oranına dayanır (15).

SLIQ: Hızlı ölçeklenebilir bir sınıflayıcı ve ağaç budama algoritması mevcuttur (15).

SPRINT: Tek bir niteliğin değerine dayanır. Tüm bellek sınırlamaları üzerinde nitelik listesi veri yapısı kullanarak işlem yapar (15).

5.2.2. Karar Ağaçları Kullanım Alanları

Karar ağacının yaygın olarak kullanıldığı alanlar şunlardır (15):

- Belirli bir sınıfın muhtemel üyesi olacak elemanları belirlemesidir.
- Çeşitli vakaların yüksek, orta veya düşük risk durumları gibi kategorilere ayrılmasıdır.
- Parametrik modellerin kurulmasında kullanılmak şartıyla birçok değişken arasından en önemli olanlarının seçilmesidir.
- Gelecekteki olayların tahmin edilebilir olması için kurallar oluşturulmasıdır.
- Sadece belli bir alt gruplara özgü olan ilişkilerin tanımlanmasıdır.
- Kategorilerin birleştirilmesinde ve sürekli değişken yapılarının kesikli olan değişken yapılarına dönüştürülmesinde kullanılır.

5.2.3. Karar Ağaçlarının Avantajları ve Dezavantajları

Karar ağaçları, sınıflandırma ve tahmin için oldukça sık kullanılan veri madenciliği yaklaşım modelidir. Karar ağaçları, kolay yorumundan dolayı karar vericiler için avantaj sağlayabilmektedir. Bunun yanı sıra maliyetlinin az olması, anlaşılmasının, yorumlanmasının ve veri tabanlarıyla entegrasyonun kolay olması, güvenilirliklerinin iyi olmasından dolayı bazı avantajlar sağlamaktadır (19).

Karar ağaçlarının sağladığı avantajlarının yanı sıra dezavantajları da mevcuttur. Ağaç ayrılan bölümleri ile birlikte genişledikçe sınıflandırma sonucunda bölüm düğümlerinde olduğundan daha az bilgi kalacaktır. Karar ağaçları veriyi fazla sayıda takıma ayırır. Ayrılan takımlar daha spesifik bir durum kazandıkça küçülmeye başlar. Bakılması gereken farklı ifadelerin sayısı arttıkça eğitim kümelerinin her biri daha da küçülür. Rakamlardaki küçülme yüzünden sınıflandırılmanın düzgün bir şekilde ifade edilmesinde daha az güvenilirlik olabilmektedir (20).

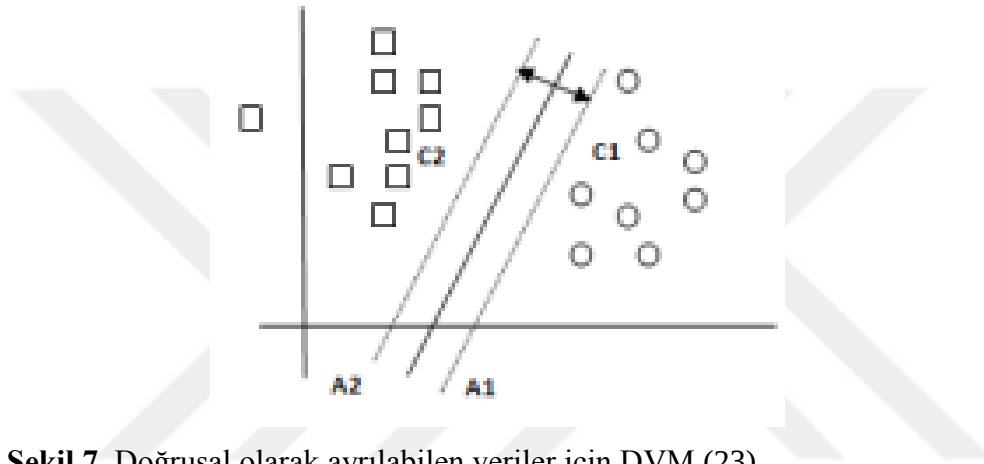
5.3. Destek Vektör Makineleri

Destek vektör makineleri, örüntü tanıma ve sınıflandırma problemlerinin çözümü açısından Vapnik tarafından geliştirilmiştir (21). Destek vektör makineleri, öğrenme, sınıflandırma, kümeleme, yoğunluk tahmini ve veriden regresyon kuralları geliştirmek için kullanılan bir eğitim algoritmasıdır. Destek vektör makineleri, istatistiksel öğrenme teorisi üzerine inşa edilmiştir (22). DVM iki ve ikiden çok sınıflı sınıflandırma probleminin çözümü için tasarlanmış makine öğrenmesi algoritmasıdır. Destek vektör makinelerinde amaç, sınıfları birbirinden ayıracak hiper düzleminin elde edilmesidir. Yani, farklı sınıflara ait destek vektörleri arasındaki uzaklığı maksimize etmektir (21).

Destek vektör makineleri, sınıflandırma ve regresyon sorunlarında çalıştırılabilir. Destek vektör makineleri regresyon metodundaki ana fikir, eldeki eğitim verilerinin karakterini mümkün olduğunca gerçeğe yakın bir şekilde yansıtan ve istatistiksel öğrenme teorisine uyan doğrusal ayırıcı fonksiyonun bulunmasıdır. Regresyonda doğrusal olmayan durumların işlenebilmesi için çekirdek fonksiyonları kullanılır. DVM’de karşılaşılabilecek iki durum vardır. Bunlar, verilerin doğrusal olarak ayrılacakları bir yapıda olması ya da doğrusal olarak ayrılamayan bir yapıda olmasıdır (23).

5.3.1. Doğrusal Ayrılabilen Veriler için DVM

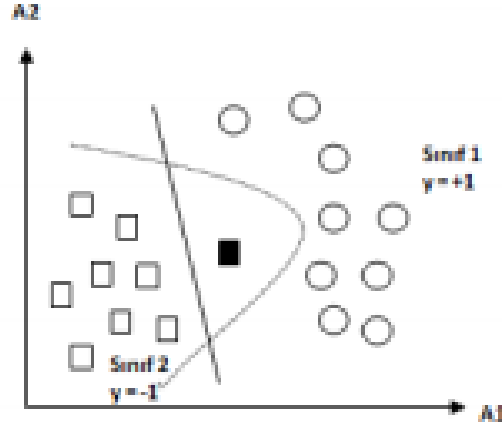
Eğitim için kullanılacak N elemandan oluşan verinin $Q = \{x_i, y_i\}, i = 1, 2, \dots, N$ olduğu varsayarsak $y_i \in \{-1, 1\}$ etiket değerleri ve $x_i \in R^N$ özellikler vektörünü göstermektedir. Lineer olarak ayrılabilme durumunda, bu iki değerli veriler direk olarak aşırı düzlem ile ayrılabilir. Buradaki aşırı düzleme ayırıcı aşırı düzlem denir. Doğrusal olarak ayrılabilen veriler için DVM amacı aşırı düzlemin iki ayrı sınıfta bulunan örnek grubuna eşit uzaklıkta olmasını sağlamaktır (23).



Şekil 7. Doğrusal olarak ayrılabilen veriler için DVM (23)

5.3.2. Doğrusal Ayrılamayan Veriler için DVM

Doğrusal ayrılabilen durumlarda veriler iki sınıfa doğrusal bir düzlem ile ayrılabilir. Uygulamada bu durum her zaman geçerli değildir. Yani bu durumda doğrusal bir düzlemle veriler birbirinden ayrılmayabilir. Bundan dolayı verilerin doğrusal olarak ayrılamadığı yerlerde doğrusal olan sınıflandırıcılar yerine doğrusal olmayan sınıflandırıcılar kullanılabilir. Doğrusal olmayan ayrılabilirlik durumu olduğu zaman, eğitim örnekleri orijinal giriş uzayında doğrusal olarak ayrılamazlar. Bu durumlarda destek vektör makineler, doğrusal olmayan haritalama fonksiyonu ile orijinal giriş uzayından doğrusal olarak kolayca sınıflandırma yapabileceği yüksek boyutlu nitelik uzayına dönüşüm yapabilmektedir. Sonuç olarak çekirdek fonksiyonları kullanılarak bütün değerlerin yeniden çarpım değerlerinin hesaplanarak bulunması yerine doğrudan çekirdek fonksiyonunda değerinin yerine koyularak nitelik uzayındaki değerinin bulunması sağlanmaktadır. Bundan dolayı da, son derece yüksek boyutlu bir nitelik uzayı ile uğraşma durumu kalmaz (23).



Şekil 8. Doğrusal ayrılamayan veriler için DVM (23)

5.3.3. Destek Vektör Makinelerinin Kullanım Alanları

Destek vektör makineleri, yüksek boyutlu fakat sayıca az veri içeren durumlarda da başarı sağlamaktadır. Bu özelliklerinden dolayı DVM; veri madenciliği, müşterilerin dolandırıcılık tespiti ve görüntü sınıflandırma gibi birçok uygulama alanında kullanılmıştır (24). Örüntü tanıma, yüz bulma ve tanıma, veri madenciliği, dil yapısını inceleyen mantıksal programlamalarda, uçak alt basınç profillerinin modellenmesi, biyoloji ve diğer biyoinformatik uygulamalarda, gen analizlerinde ve proteinlerin sınıflandırılmasında destek vektör makineleri kullanılmaya başlanmıştır. DVM'nin uygulama alanlarına, el yazısı tanıma, yüz tanıma, 3-boyutlu nesne tanıma, ses tanıma, konuşmacı tanıma, metin sınıflandırma örnek verilebilir (25).

5.3.4. Destek Vektör Makinelerinin Avantajları ve Dezavantajları

Destek vektör makinesi kullanılarak elde edilebilecek avantajlardan bazıları aşağıdaki gibidir:

- Yüksek doğruluk ve yüksek profilinden ötürü veri madenciliği kazanımları yüksektir.
- Yapay sinir ağlarından kullanımı daha kolay olabilir.
- Gürültü verilerinden aşırı derecede etkilenmez ve uyum problemi yoktur.
- Sayısal tahmin problemleri veya sınıflandırma için kullanılabilir (7).

Destek vektör makinesinin dezavantajlarından bazıları aşağıdaki gibidir:

- Girdi verisi kümesinde çok sayıda örnek ve özellik varsa eğitmek zor olabilir.
- Karmaşık kara kutu modelinden dolayı yorumlamak imkansız olmasa da zordur.

- En iyi modeli bulmak için kernel ve model parametreleri ile çeşitli kombinasyonlar yapmak gerebilir (7).

5.4. Çalışmada Kullanılan Programlar

Bu tez çalışmasında kullanılan laboratuvar verileri KTÜ Tıp Fakültesi Farabi Hastanesi Acil Tıp servisinden uzman hekim danışmanlığında alınmıştır. Alınan hasta verileri tanı alan ve tanı almayan şeklinde iki sınıfa ayrılmıştır. Verilerin istatistiksel analizi SPSS Statistics 22 programı yardımıyla yapılmıştır. Klinik tanı sistemi modelleri makine öğrenmesi yöntemlerinden YSA, DVM ve karar ağacı ile Rstudio 3.2 programı kullanılarak geliştirilmiştir.

5.5. Verilerin Hazırlanması

Çalışmada kullanılan veriler KTÜ Tıp Fakültesi Farabi Hastanesi Acil Tıp Servisine nefes darlığı şikayeti ile gelen 240 hasta verisinden elde edilmiştir. Kullanılan veriler pulmoner emboli tanısı alan ve almayan şeklinde iki sınıfa ayrılmıştır.

6. BULGULAR

6.1. Veri Analizi

Bu tez çalışmasında 142'si kadın 98'i erkek olmak üzere 240 hasta verisi kullanılmıştır. Bu hasta verilerinin 149'u pulmoner emboli tanısı alan ve 91'i pulmoner emboli tanısı almayan veriler olarak iki sınıfa ayrılmıştır.

Cinsiyet değişkeni ile pulmoner emboli tanısı alan ve almayan hastaların ilişkisini değerlendirmek için Ki-Kare testi uygulanmıştır (Tablo 7). Pulmoner emboli tanısı alan ve almayan hasta verileri arasında yaş, D_dimer, PaCO₂, PaO₂ ve lokosit değerleri açısından tanımlayıcı istatistiklerine bakılmıştır (Tablo 3) (Tablo 4). Bu değerlerin normal dağılıma uygun olup olmadığını belirlemek için Shapiro-Wilk testi uygulanmıştır ($p < 0.05$). Normal dağılım göstermeyen sürekli değişkenler olan yaş, D_dimer, PaCO₂, PO₂ ve lokosit değerleri açısından anlamlı bir fark olup olmadığını belirlemede Mann Whitney U testi yapılmıştır (Tablo 5). İstatistiksel anlamlılık düzeyi $p < 0.05$ olarak kabul edilmiştir.

Pulmoner emboli tanısı alan ve almayan hastalarla ilgili veri seti örneği Şekil 9 ve Şekil 10'da gösterilmiştir. Pulmoner emboli tanısı alan ve almayan hastaların cinsiyete göre dağılımları Tablo 1 ve Tablo 2'de gösterilmiştir.

Cinsiyet	Yaş	D_dimer	Paco2	Pao2	Lokosit	Tani
Kadin	75	1,92	46,00	21,00	14,20	Var
Kadin	81	0,97	31,70	61,80	11,70	Var
Kadin	71	0,52	31,70	59,60	12,30	Var
Kadin	29	0,24	42,90	88,70	8,20	Var
Kadin	70	3,92	24,90	63,90	12,90	Var
Erkek	85	1,50	42,50	50,20	13,80	Var
Kadin	51	0,55	30,10	77,60	11,90	Var
Erkek	53	15,70	44,10	30,60	13,80	Var
Kadin	86	12,10	29,60	56,60	12,50	Var
Erkek	84	39,31	26,30	78,30	11,40	Var
Kadin	89	9,39	35,10	93,20	16,40	Var
Kadin	67	0,56	40,30	57,60	10,10	Var
Erkek	73	4,85	28,70	61,60	11,60	Var
Erkek	81	12,82	25,60	47,90	2,40	Var
Kadin	68	7,16	42,30	65,50	12,30	Var
Kadin	81	10,68	29,80	82,30	14,00	Var
Kadin	79	1,52	79,70	40,90	13,80	Var
Kadin	77	0,89	31,20	59,10	13,30	Var
Kadin	74	3,34	76,00	63,10	13,80	Var
Kadin	87	19,79	24,80	48,10	11,60	Var
Kadin	59	3,21	46,10	53,70	17,20	Var

Şekil 9. Pulmoner emboli tanısı alan hastaların veri seti örneği

Cinsiyet	Yas	D_dimer	Paco2	Pao2	Lokosit	Tani
Erkek	55	0,80	54,70	102,80	12,80	Yok
Erkek	74	0,67	29,40	53,40	6,00	Yok
Kadin	47	11,40	20,30	259,00	13,90	Yok
Kadin	60	5,48	33,40	158,00	12,00	Yok
Erkek	78	5,63	33,00	46,50	5,00	Yok
Erkek	68	6,02	47,30	300,00	15,80	Yok
Kadin	86	1,18	36,80	82,70	10,10	Yok
Erkek	46	2,13	38,80	52,60	28,10	Yok
Erkek	23	0,40	42,10	66,10	14,60	Yok
Erkek	68	0,20	22,70	71,40	32,70	Yok
Kadin	80	0,30	36,30	71,20	8,00	Yok
Kadin	55	0,44	33,20	82,60	4,30	Yok
Kadin	87	8,20	42,40	139,00	9,30	Yok
Kadin	72	0,25	45,20	120,10	3,74	Yok
Erkek	55	1,05	30,30	79,10	10,75	Yok
Erkek	60	0,57	38,60	112,00	10,12	Yok
Erkek	83	0,36	37,00	107,00	2,55	Yok
Erkek	73	0,80	42,80	54,90	5,91	Yok
Kadin	87	0,87	30,40	79,30	7,63	Yok
Erkek	69	0,65	87,50	34,10	4,55	Yok
Erkek	78	0,55	45,10	24,50	17,08	Yok

Şekil 10. Pulmoner emboli tanısı almayan hastaların veri seti örneği

Tablo 1. Pulmoner emboli tanısı alan hasta verilerinin cinsiyete göre dağılımları

	Frequency	Percent	Valid Percent	Cumulative Percent
Kadin	94	63.1	63.1	63.1
Erkek	55	36.9	36.9	100.0
Toplam	149	100.0	100.0	

Tablo 2. Pulmoner emboli tanısı almayan hasta verilerinin cinsiyete göre dağılımları

	Frequency	Percent	Valid Percent	Cumulative Percent
Kadin	48	52.7	52.7	52.7
Erkek	43	47.3	47.3	100.0
Toplam	91	100.0	100.0	

Tablo 3. Pulmoner emboli tanısı alan hasta verilerinin tanımlayıcı istatistikleri

	N	Minimum	Maksimum	Mean	Std Deviation
Yas	149	20	98	70.28	15.802
D_dimer	149	.10	39.31	7.3424	7.32948
PaCO2	149	15.50	88.90	36.6181	12.61531
PaO2	149	10.40	140.00	55.1971	25.03437
Lokosit	149	2.20	27.91	12.2340	4.06901

Tablo 4. Pulmoner emboli tanısı almayan hasta verilerinin tanımlayıcı istatistikleri

	N	Minimum	Maksimum	Mean	Std Deviation
Yas	91	20	91	65.13	17.863
D_dimer	91	.05	12.59	1.6639	2.63157
PaCO2	91	20.30	104.00	46.3310	16.97105
PaO2	91	10.80	315.00	75.9023	49.01468
Lokosit	91	2.55	32.70	10.3529	4.75306

Tablo 5. Pulmoner emboli tanısı alan ve almayan hasta verilerinin karşılaştırılması

	Tani	N	Mean Rank	Sum of Rank	Mann Whitney U	z	p
Yas	Var	149	128.24	19108.50	5625.500	2.212	.027
	Yok	91	107.82	9811.50			
	Toplam	240					
D_dimer	Var	149	158.80	27766.50	1967.500	9.222	.000
	Yok	91	67.62	6153.50			
	Toplam	240					
PaCO2	Var	149	101.15	15071.50	3896.500	5.525	.000
	Yok	91	152.18	13848.50			
	Toplam	240					
PaO2	Var	149	180.23	16126.50	4951.500	3.503	.000
	Yok	91	140.59	12793.50			
	Toplam	240					
Lokosit	Var	149	135.21	20146.00	4588.500	4.200	.000
	Yok	91	96.42	8774.00			
	Toplam	240					

Tablo 5’te uygulanan Mann Whitney U sonucuna göre pulmoner emboli tanısı alan ve almayan hasta verilerinde yaş, D_dimer, PaCO2, PaO2 ve lokosit değerleri açısından anlamlı bir farklılık olduğu görülmüştür.

Tablo 6. Cinsiyet değişkenleri arasında Ki-Kare testi sonucu

	Tanı		χ^2	p
	Var n(%)	Yok n(%)		
Cinsiyet 2				
Erkek	55 (56.1)	43 (43.9)	2.500	.114
Kadın	94 (66.2)	48 (33.8)		

Tablo 6'daki ki-kare testi sonucuna göre p değeri 0.114 (>0.05) çıkmıştır. Bu sonuca göre cinsiyet değişkeninin pulmoner emboli tanısı açısından anlamlı bir fark oluşturmadığı görülmektedir ve bu nedenle geliştirilen tanı modellerinde bu parametre kullanılmamıştır.

Yapılan istatistiksel analiz sonucunda geliştirilecek model için yaş, D_dimer, PaCO₂, PaO₂ ve lokosit verileri kullanılmıştır. Toplam 240 verinin; rastgele seçilerek 168'i eğitim seti ve 72'si ise test seti olarak kullanılmıştır. Bu veri setleri kullanılarak R programlama dili yardımıyla DVM, YSA ve karar ağacı tanı modelleri geliştirilmiştir. Eğitim seti ve test seti örnekleri Şekil 11 ve Şekil 12'de gösterilmiştir.

Yaş	D dimer	PaCO ₂	PaO ₂	Lokosit	Tanı
86	14.41	41.1	20.2	3.13	1
86	17.91	61	90.4	26.44	1
87	0.87	30.4	79.3	7.63	0
87	3.16	38.6	61.4	16.57	1
87	8.2	42.4	139	9.3	0
87	10.9	41.2	45	23.2	1
87	19.79	24.8	48.1	11.6	1
88	4.02	31.9	56.3	9.07	1
88	5.02	36.3	11	11.5	1
90	0.42	43.5	81.6	6.75	0
91	3.32	73.1	140	12.06	1
91	3.79	18.2	92.8	13.7	1
92	6.05	27.6	63.4	14.57	1
92	10.47	48.1	24.7	13.2	1
95	3.47	54.5	36.7	7.7	1
98	6.57	38.4	99	8.3	1
20	2.42	49.2	48.8	12.55	1
23	0.4	42.1	66.1	14.6	0
39	2.38	46.4	37.7	14.42	1
41	20	80.3	31.2	8.34	1
41	26.41	18.9	76.5	14.37	1
42	11.42	33.1	35.85	11.71	1
45	0.8	33.3	64.8	5.7	0
46	0.25	31	65.5	3.6	0
46	2.13	38.8	52.6	28.1	0
50	4.41	22.9	16	13.48	1

Şekil 11. Eğitim seti örneği

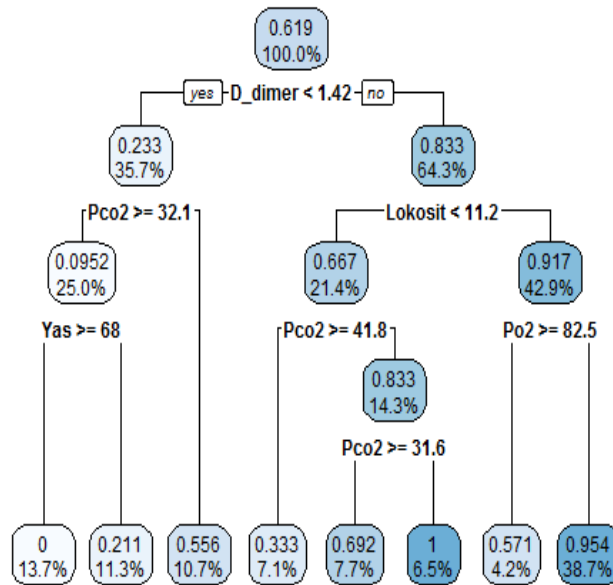
Yas	D_dimer	PaCO2	PaO2	Lokosit	Tani
77	0.4	35	89	4.8	0
78	3.02	30.8	70.4	12.42	1
78	5.4	44	49.4	20.6	1
79	5.49	35.2	77.1	12.7	1
79	5.71	43.1	25.9	9.86	1
80	0.3	36.3	71.2	8	0
80	3.55	29.6	80.6	15.17	1
81	1.41	37.9	20.4	11.21	1
83	0.2	36.4	112	9	0
83	2.45	63.1	23.5	7.37	0
83	4.47	58.9	50.8	6.2	0
84	0.6	40.5	89.8	8.2	0
84	8.52	44.2	86.4	4.4	1
86	12.1	29.6	56.6	12.5	1
89	9.39	35.1	93.2	16.4	1
90	7.24	27.2	73.4	12.89	1
21	0.4	35	315	16.9	0
27	0.16	32.1	75.3	12.08	0
42	6.56	27.4	95.7	9.71	1
46	4.74	47.8	28.1	11.2	1
50	3.4	28.2	59.9	10.21	1
51	0.4	58.3	98.7	8.9	0
53	0.26	40.3	93.2	9.9	0
55	0.37	58.3	67.9	6.35	0
58	24.64	40.7	18	11.17	1

Şekil 12. Test seti örneği

Eğitim ve test seti örneklerinde tanı kısmında 0 Yok (pulmoner emboli tanısı almayan) ve 1 Var (pulmoner emboli tanısı alan) şeklinde kodlanmıştır.

6.2. Karar Ağacı ile Modellemede Elde Edilen Bulgular

Eğitim veri seti kullanılarak karar ağacı yöntemiyle PE tanı modeli geliştirilmiştir. Geliştirilen bu karar ağacı modeli Şekil 13’te gösterilmiştir.



Şekil 13. Karar ağacı modeli

Geliştirilen karar ağacı modeli 72 test verisi üzerinde test edilmiş ve aşağıdaki sonuçlar elde edilmiştir.

Tablo 7. Geliştirilen karar ağacı modeline ait performans sonucu

	Gerçek Durum	
Tanı	Pozitif(hasta)	Negatif(hasta değil)
Pozitif(hasta)	38(DP)	3(YP)
Negatif(hasta değil)	7(YN)	24(DN)

Duyarlılık ve özgüllük aşağıdaki formülle hesaplanmıştır(35):

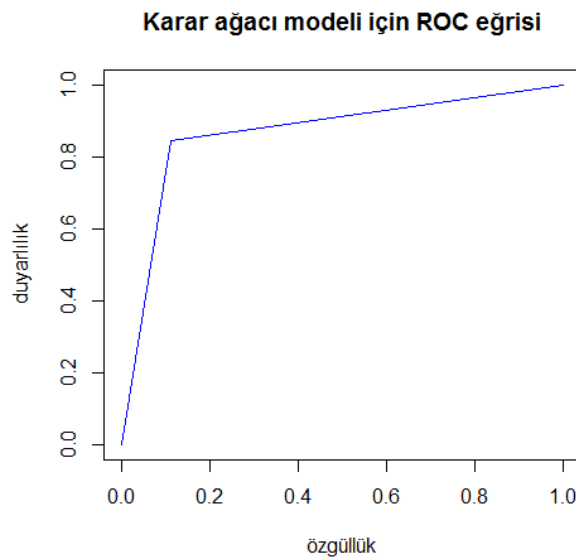
Duyarlılık (sensitivity): Test, gerçek hastalar içinden hastaları ayırmasıdır.

$$\text{Duyarlılık} = \text{DP}/(\text{DP}+\text{YN}) \quad (6.1)$$

Özgüllük (specificity): Test, gerçek sağlamlar içinden sağlamları ayırmasıdır.

$$\text{Özgüllük} = \text{DN}/(\text{DN}+\text{YP}) \quad (6.2)$$

Bu sonuca göre geliştirilen karar ağacı modeli için duyarlılık % 84 ve özgüllük % 88 olarak elde edilmiştir. Karar ağacı ile yapılan modellemelerde ROC eğrisi Şekil 14’te gösterilmiştir.



Şekil 14. Geliştirilen karar ağacı modeline için ROC eğrisi

ROC eğrisi altında kalan alan AUC (Are Under Curve) değeri geliştirilen karar ağacı modeli için 0.86 bulunmuştur.

6.3. Destek Vektör Makineleri ile Modellemede Elde Edilen Bulgular

DVM modelinde eğitim seti kullanılarak sigmoid, linear, radial ve polynomial çekirdek fonksiyonları ile dört farklı model geliştirilmiştir. Geliştirilen bu modellerin performans sonuçları Tablo 8’de gösterilmiştir.

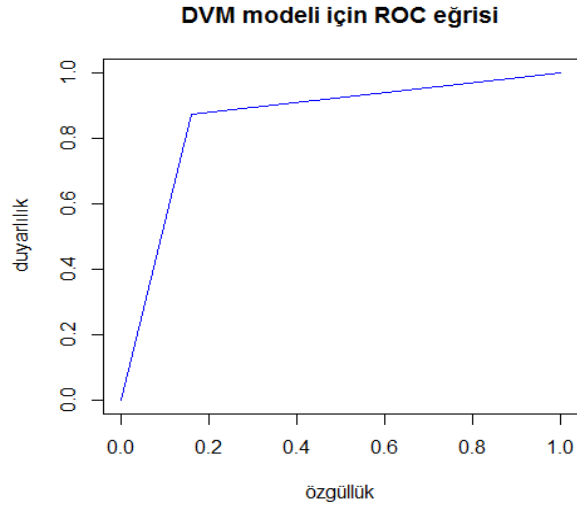
Tablo 8. DVM çekirdek fonksiyonlarının sonuçları

	Duyarlılık	Özgüllük
Sigmoid	% 80	% 59
Radial	% 91	% 77
Polynomial	% 91	% 48
Linear	% 91	% 70

Tablo 8’de geliştirilen DVM modelleri için en iyi sonucu radial çekirdek fonksiyonunun verdiği görülmektedir. En iyi DVM modelinde (radial kernel fonksiyonu) test verisi üzerinden % 91 duyarlılık % 77 özgüllük elde edilmiştir. Geliştirilen DVM modeli test seti üzerinde test edilmiş ve elde edilen değerler Tablo 9’da gösterilmiştir. Ayrıca DVM modelinin performansına ilişkin ROC eğrisi Şekil 15’te verilmiştir.

Tablo 9. Geliştirilen DVM modeline ait performans sonucu

	Gerçek Durum	
Tanı	Pozitif(hasta)	Negatif(hasta değil)
Pozitif(hasta)	41(DP)	6(YP)
Negatif(hasta değil)	4(YN)	21(DN)

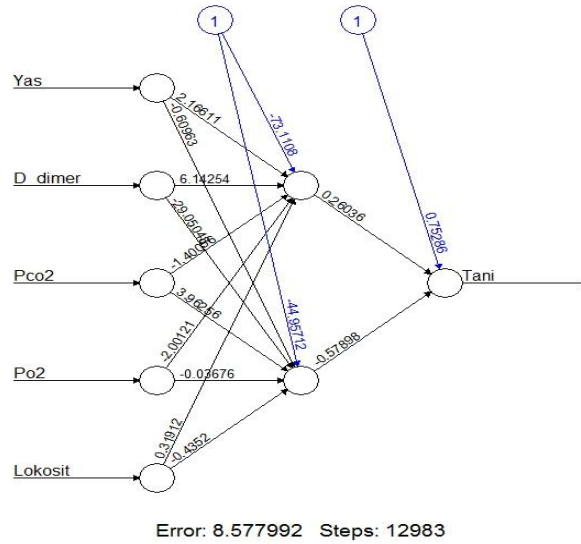


Şekil 15. Geliştirilen DVM modeli için ROC eğrisi

ROC eğrisi altında kalan AUC değeri geliştirilen DVM modeli için 0.85 olarak bulunmuştur.

6.4. Yapay Sinir Ağları ile Modellemede Elde Edilen Bulgular

Eğitim seti kullanılarak iki katmanlı bir yapay sinir ağı modeli geliştirilmiştir. Geliştirilen yapay sinir ağı modeli Şekil 16’da verilmiştir.

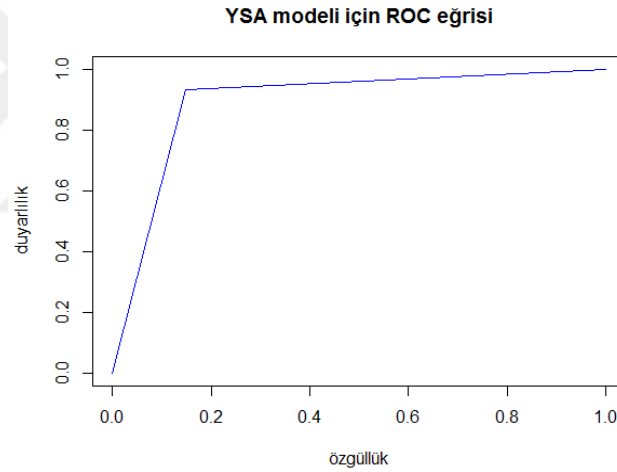


Şekil 16. Yapay sinir ağı modeli

Geliştirilen YSA modeli için test seti üzerinde % 93 duyarlılık ve % 85 özgüllük değerleri elde edilmiştir. YSA sonucunda elde edilen değerler Tablo 10.'de gösterilmiştir. Ayrıca YSA ile yapılan modellemede ROC eğrisi Şekil 17'de gösterilmiştir.

Tablo 10. Geliştirilen YSA modeline ait performans sonuçları

Tanı	Gerçek Durum	
	Pozitif(hasta)	Negatif(hasta değil)
Pozitif(hasta)	42(DP)	4(YP)
Negatif(hasta değil)	3(YN)	23(DN)



Şekil 17. Geliştirilen YSA modeli için ROC eğrisi

ROC eğrisi altında kalan AUC değeri geliştirilen YSA modeli için 0.89 elde bulunmuştur.

6.5. Geliştirilen Modellerin Değerlendirilmesi

Pulmoner emboli klinik tanı sistemi için karar ağacı, DVM ve YSA modelleri geliştirilmiştir. Bu modellerin sonucuna bakıldığında geliştirilen karar ağacı modeli için % 84 duyarlılık, % 88 özgüllük ve 0.86 AUC değeri, geliştirilen DVM modeli için % 91 duyarlılık, % 77 özgüllük ve 0.85 AUC değeri, geliştirilen YSA modelinde ise % 93 duyarlılık, % 85 özgüllük ve 0.89 AUC değeri elde edilmiştir. Geliştirilen bu üç

modelin sonucuna bakıldığında pulmoner emboli klinik tanı sistemi için en iyi model % 93 duyarlılık ve % 85 özgüllük değerleri ile YSA modeli olduğu görülmektedir.

Tablo 11. Geliştirilen modellerin karşılaştırılması

	Duyarlılık (%)	Özgüllük (%)	AUC Değeri
Karar Ağacı	84	88	0.86
DVM	91	77	0.85
YSA	93	85	0.89

7. TARTIŞMA ve SONUÇ

Son yıllarda yapılan çalışmalarda makine öğrenmesi ve veri madenciliği yöntemleri kullanılarak birçok alanda önemli çalışmalar yapılmıştır. Özellikle tıp alanında yapılan çalışmalar da birçok hastalığa doğru teşhis ve hızlı müdahale olanağı sağlamıştır. Bu nedenle yapılan tez çalışmasında makine öğrenmesi yöntemleri kullanılarak hekime yardımcı bir pulmoner emboli klinik tanı sistemi geliştirilmiştir. PE klinik tanı sistemi için DVM, YSA ve karar ağacı yöntemleri kullanılarak 3 farklı model geliştirilmiş ve performans sonuçları karşılaştırılarak en iyi tanı modeli seçilmiştir.

Yapılan çalışmasında acil servise nefes darlığı ile başvuran pulmoner emboli tanısı alan ve pulmoner emboli tanısı almayan hasta verileri kullanılmıştır. Demografik veri olan cinsiyet, yaş ile biyokimya laboratuvar test bulgularından D_dimer, PaCO₂, PaO₂ ve lokosit parametreleri ele alınmıştır. Veri setinde 149 pulmoner emboli tanısı alan ve 91 pulmoner emboli tanısı almayan şeklinde toplam 240 hasta bulunmaktadır. İstatistiksel analiz sonucuna göre pulmoner emboli tanısı alan ve almayan hastalar arasında yaş, D_dimer, PaCO₂, PaO₂ ve lokosit değerleri açısından anlamlı bir farklılık bulunmuştur. Pulmoner emboli tanısı alan ve almayan hastalar arasında cinsiyetler açısından yapılan istatistiksel analiz sonucuna göre de anlamlı bir farklılık bulunmamıştır. Veri setinde bulunan 240 hasta verisini rastgele 168'i eğitim ve 72'si test seti olmak üzere iki sınıfa ayırdıktan sonra R programlama dili yardımıyla DVM, YSA ve karar ağacı yöntemleri kullanılarak üç farklı tanı model geliştirilmiştir. DVM yönteminde, 4 farklı çekirdek fonksiyonu için 4 farklı model geliştirilmiştir. En iyi sonucu veren DVM radial çekirdek fonksiyonu sonucuna göre 72 test verisi üzerinden % 91 duyarlılık, % 77 özgüllük ve 0.85 AUC değeri elde edilmiştir. Geliştirilen karar ağacı modeline göre 72 test verisi üzerinden % 84 duyarlılık, % 88 özgüllük ve 0.86 AUC değeri elde edilmiştir. Geliştirilen YSA modelinde ise % 93 duyarlılık, % 85 özgüllük ve 0.89 AUC değeri elde edilmiştir. Geliştirilen bu üç modelin sonucuna göre en iyi sonucu YSA modelinin verildiği görülmüştür.

Serpen ve arkadaşları tarafından 2008 yılında yapılan bir çalışmada pulmoner emboli teşhisi için yapay sinir ağı modeli geliştirilmiştir. 30 test örneği üzerinde 10 çapraz doğrulama yöntemi kullanılmıştır. Bu yöntem sonucunda naive Bayes ile % 73.3, Bayesian belief network ile % 50.0, çok katmanlı perceptron ile % 60.0, c4.5 karar

ağacı ile % 63.3, karar destekli uyarlanabilir yükseltme m1 algoritması ile % 33.3 ve rastgele ağaç kullanarak torbalama (bagging) meta sınıflayıcı ile % 56.7 doğruluk değeri elde edilmiştir (3). Yapılan başka bir çalışmada Agharezaei ve arkadaşları farklı parametreler kullanılarak pulmoner emboli mortalite oranını azaltma amacıyla risk faktörlerini belirleyen bir karar destek sistemi geliştirilmiştir (4).

Yapılan bu tez çalışmasının amacı acil servise nefes darlığı ile gelen hastaların pulmoner emboli olasılığını belirleyerek için bilgisayarlı tomografi görüntüleme yöntemini kullanmadan sadece yaş, D_dimer, PaCO₂, PaO₂ ve lökosit verilerini kullanarak erken tanıyı yardımcı olmaktır. Pulmoner embolide erken tanı çok önemli olduğu için bu parametreleri kullanarak hekime yardımcı bir klinik tanı sistemi geliştirilmiştir. Geliştirilen klinik tanı sistemi yeni bir hasta geldiği zaman veri değerleri sisteme girildiğinde hastanın pulmoner emboli olasılığının yüksek olup olmamasına ilişkin hızlıca bilgi vererek erken tanıya yardımcı olabilecektir. Bununla birlikte, tanının koyulamadığı gözden kaçtığı durumlar için de ikinci bir gözlemci olarak hekime yardımcı olabilecektir.

Literatürde benzer çalışmalar bulunmasına rağmen bu parametreler ve kullanılan yöntemler ile yapılan çalışmalar oldukça sınırlı sayıdadır. Bu bağlamda ileride geliştirilecek farklı bilimsel çalışmalara yol göstermesi beklenmektedir.

8. KAYNAKLAR

1. Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galiè N, Pruszczyk P, Bengel F, Brady A, Ferreira D, Janssens U, Klepetko W, Mayer E, Jardin M, Bassand J (2008). Acute Pulmonary Embolism (Diagnosis and Management of), 29: 2276-2315.
2. Cushman M (2007). Epidemiology and risk factors for venous thrombosis, 44: 62-9
3. Serpen G, Tekkedil D.K, Orra M.(2007). A knowledge artificial network classifier for pulmonary embolism diagnosis, 38:204-220.
4. Agharezaei Z., Bahaadinbeigy K., Tofighi S., Agharezaei L., Nemati A.(2014). Attitude of Iranian physicians and nurses towards a clinical decision support system for pulmonary embolism and deep vein thrombosis 115:95-101.
5. Duru S, Dilli A, Ergün R, Ardiç S (2012). Clinical, laboratoray and computed tomography results in pulmonary embolism: Retrospective evaluation of 205 patients. Anadolu Kardoloji Dergisi, 12:142-149.
6. Bayer H, Çoban T (2015). Web İstatistiklerinde Makine Öğrenmesi Algoritmaları ile Kritik Parametre Tespiti. Electronic Journal of Vocational Collages, 5: 23-41.
7. Lantz B (2015). Machine Learning with R Second Edition Packt Publishing Brimingham: UK, 1-300.
8. Alpaydın E (2010). İntroduction to Machine Learning Second Edition The MIT Press Cambridge, Massachusetts London, England, 1-40.
9. Öztemel E (2012). Yapay Sinir Ağları. Üçüncü Baskı. Papatya Yayıncılık Eğitim AŞ. İstanbul, No. 11/3, 29-60.
10. Dülger E (2016). Ekonomide Öncü Göstergeler ile Büyüme Tahmini Uygulaması. Yüksek Lisans Tezi. Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
11. Yazıcı Ü (2010). Yapay Sinir Ağları İle Banka Müşterisi Bekleme Süresi Tahmini. Yüksek Lisans Tezi. Atatürk Üniversitesi, Fen Bilimleri Enstitüsü, Erzurum.
12. Sarı M (2016). Yapay Sinir Ağları ve Bir Otomotiv Firmasında Satış Talep Tahmini Uygulaması. Yüksek Lisans Tezi. Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya.
13. Karahan M (2011). İstatistiksel Tahmin Yöntemleri: Yapay Sinir Ağları Metodu ile Ürün Talep Tahmini Uygulaması. Doktora Tezi. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, Konya.

14. Ballı M. T (2014). Yapay Sinir Ağları ile Talep Tahmini ve Gıda Sektöründe Uygulanması. Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
15. Emel G, Taşkın Ç (2005). Veri Madenciliğinde Karar Ağaçları Ve Bir Satış Analizi Uygulanması. Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 2-6.
16. Albayrak A. S, Yılmaz Ş (2009). Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 31-52.
17. Koyuncu A. S, Özgülbaş N (2008). İMKB’ de İşlem Gören Kobi’lerin Güçlü ve Zayıf Yönleri: CHAID Karar Ağacı Uygulanması. Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 1-21.
18. Daş B, Türkoğlu İ (2014). DNA Dizilimlerinin Sınıflandırılmasında Karar Ağacı Algoritmalarının Karşılaştırılması. Elektronik – Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu, 381-383
19. Çalış A, Kayapınar S, Çetinyokuş T (2014). Veri madenciliği karar ağacı algoritmaları ile bilgisayar ve internet güvenliği üzerine bir uygulama. Endüstri Mühendisliği Dergisi, 3-4: 2-19.
20. Altaş D, Gülpınar V (2012). Karar Ağaçları ve Yapay Sinir Ağlarının Sınıflandırma Performanslarının Karşılaştırılması: Avrupa Birliği Örneği. Trakya Üniversitesi Sosyal Bilimler Dergisi, 1: 1-22.
21. Erdoğan Ş, Ayhan S (2014). Destek Vektör Makineleri ile Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi. Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 9: 175-198.
22. Çomak E (2004). Destek Vektör Makineleri Çoklu Sınıf Problemleri İçin Çözüm Önerileri. Yüksek Lisans Tezi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
23. Yakut E, Elmas B, Yavuz S (2014). Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini. Süleyman Demirel Üniversitesi İİBF Dergisi, 1: 139-157.
24. Çomak E, Güner N (2011). Mühendislik Öğrencilerinin Matematik 1 Derslerindeki Başarısının Destek Vektör Makineleri Kullanılarak Tahmin Edilmesi. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 2: 87-96.

25. Hacıefendioğlu Ş (2012). Makine Öğrenmesi Yöntemleri ile Glokom Hastalığının Teşhisi. Yüksek Lisans Tez. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
26. Goldhaber S, Bounameaux H (2012). Pulmonary Embolism and Deep Vein Thrombosis, 379:1835-1846.
27. Salçın E (2015). Acil Servise Solunum Sıkıntısı ile Başvurup Pulmoner Emboli Tanısı Alan Hastalarda Bakılan C-reaktif Protein Düzeyinin Well's Skoru ve Pulmoner Emboli Ağırlık İndeksi Skorlama Sonuçlarıyla Klinik Korelasyonun İncelenmesi. Uzmanlık Tezi. Marmara Üniversitesi Tıp Fakültesi Acil Tıp Anabilim Dalı, İstanbul.
28. Çilingir B, Günbatır H (2016). Kronik Tromboembolik Pulmoner Hipertansiyon. Van Tıp Dergisi 23: 125-131.
29. Arslantaş U, Yeter E (2012). Akut Pulmoner Emboli. Türkiye Klinikleri Journal of Cardiology Special Topics 5: 66-70.
30. Balcı M (2015). Akut Pulmoner Embolinin Klinik ve Radyolojik Bulguları. Türkiye Klinikleri J Cardiol-Special Topics, 8:42-6.
31. Yumrutepe S (2012). Acil Serviste Dinamik Bilgisayarlı Tomografi ile Pulmoner Emboli Tanısı Konulan Hastaların Difüzyon Manyetik Rezonans Görüntüleme Bulgularının Değerlendirilmesi. Uzmanlık Tezi. İnönü Üniversitesi Tıp Fakültesi Acil Tıp Anabilim Dalı, Malatya.
32. Sade L (2014). Akut Pulmoner Emboli Tanı ve Tedavi Klavuzu. Türk Kardiyoloji Derneği Arşivi 43: 7-11.
33. Fındık M (2016). Başkent Üniversitesi Ankara Hastanesi Erişkin Acil Servisinde 2011-2015 Yıllarında Pulmoner Tromboemboli Tespit Edilen Hastalarda Risk Faktörlerinin Araştırılması ve Risk Faktörlerinin Mortaliteye Katkılarının Belirlenmesi. Uzmanlık Tezi.
34. Makkar S (2001). Diagnostic Approaches to Possible Pulmonary Embolism. American Family Physician 64: 844-848.
35. Dirican A (2001). Tanı Testi Performanslarının Değerlendirilmesi ve Kıyaslanması. Cerrahpaşa Tıp Dergisi 32: 25-30.
36. Uyanık E, Yılmaz A, Eliçabuk H, Beydilli H, Dağlı B, Avcil M (2014). Acil Serviste Pulmoner Emboli Tanısı Alan Hastaların Sosyodemografik- Klinik

Özellikleri ve Mortalite Üzerine Etkisi. Gaziosmanpaşa Üniversitesi Tıp Fakültesi Dergisi, 6:18-26.

37. Machine Learning Model (2017) [online].

https://commons.wikimedia.org/wiki/File:Machine_Learning_Technique.JPG.

38. Denetleme ve Düzenleme Sistemi, Sinir Sistemi ve Endokrin Sistem (2017) [online].

<http://www.biyolojidersnotlari.com/denetleme-ve-duzenleme-sistemi-sinir-sistemi-ve-endokrin-sistem.html>.

39. Ysa Nedir (2017) [online].

<http://fanntool.blogspot.com.tr/p/ysa-nedir.html>.

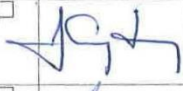
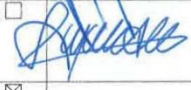
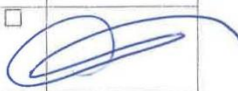


9. ETİK KURUL ONAYI

KTÜ TIP FAKÜLTESİ BİLİMSEL ARAŞTIRMALAR ETİK KURULU KARAR FORMU

KARAR BİLGİLERİ	Karar No: 4	Tarih: 02/10/2017
	Y.Doç.Dr.Burçin KURT'un sorumluluğunda yürütülmesi planlanan Yük.Lis.Öğr.Özlem KARADENİZ'e ait "Pulmoner Emboli Klinik Tanı Sistemi" başlıklı 2017/163 no.lu ve yukarıda başvuru bilgileri verilen araştırma/tez başvuru dosyası ile ilgili belgeler araştırmanın gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş, gerçekleştirilmesinde etik sakınca bulunmadığına; toplantıya katılan etik kurul üyelerinin oy birliği ile karar verilmiştir.	

KTÜ TIP FAKÜLTESİ BİLİMSEL ARAŞTIRMALAR ETİK KURULU KARAR FORMU	
ÇALIŞMA ESASI	Klinik Araştırmalar Hakkında Yönetmelik, İyi Klinik Uygulamaları Kılavuzu
BAŞKANIN UNVANI / ADI / SOYADI:	Prof.Dr.Faruk AYDIN

Unvanı/Adı/Soyadı	Uzmanlık Alanı	Kurumu	Cinsiyet	İlişki *	Katılım **	İmza
Prof.Dr.Faruk AYDIN Başkan:	Tıbbi Mikrobiyoloji	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☒ H ☐	
Prof.Dr.Gamze ÇAN Başkan Yrd.	Halk Sağlığı	KTÜ Tıp Fakültesi	E ☐ K ☒	E ☐ H ☒	E ☒ H ☐	
Prof.Dr.S.Caner KARAHAN Üye:	Tıbbi Biyokimya	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☐ H ☒	İZİNLİ
Prof.Dr.S. Murat KESİM Raporör:	Tıbbi Farmakoloji	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☒ H ☐	
Prof.Dr.Yılmaz BÜLBÜL Üye:	Göğüs Hastalıkları	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☒ H ☐	
Prof.Dr. Murat LİVAOĞLU Üye:	Plastik, Rekons. ve Estetik Cer.	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☒ H ☐	
Prof.Dr.Şafak ERSÖZ Üye:	Tıbbi Patoloji	KTÜ Tıp Fakültesi	E ☐ K ☒	E ☐ H ☒	E ☒ H ☐	
Y.Doç.Dr.Demet SAĞLAM AYKUT Üye:	Ruh Sağlığı ve Hastalıkları	KTÜ Tıp Fakültesi	E ☐ K ☒	E ☐ H ☒	E ☐ H ☒	İZİNLİ
Prof.Dr.Murat ÇAKIR Üye:	Çocuk Sağlığı ve Hastalıkları	KTÜ Tıp Fakültesi	E ☒ K ☐	E ☐ H ☒	E ☒ H ☐	

* :Araştırma ile İlişki

** :Toplantıda Bulunma

10. ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, Adı: KARADENİZ, Özlem

Uyruğu: T.C.

Doğum Tarihi: 31/10/1989

Telefon: 5058921231

E-posta: krdnzozm@gmail.com

EĞİTİM BİLGİLERİ

Derece	Mezun Olduğu Kurum	Yılı
Yüksek Lisans	KTÜ/ Biyoistatistik ve Tıp Bilişimi	2015-
Lisans	KTÜ/ Bilgisayar Mühendisliği	2010-2014
	Gazi Üniversitesi/ Mimarlık	2007-2009
Lise	Trabzon Fatih Lisesi	2005-2007
	Beşikdüzü İMKB Öğretmen Lisesi	2003-2005

YABANCI DİL

İngilizce