

**T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**DEMİR EKSİKLİĞİ ANEMİSİ HASTALIĞININ
TESPİTİNDE KULLANILAN SINIFLANDIRMA
ALGORİTMALARININ KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

Volkan SEYMEN

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ**
Tez Danışmanı : Yrd. Doç. Dr. Devrim AKGÜN

Aralık 2014

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DEMİR EKSİKLİĞİ ANEMİSİ HASTALIĞININ
TESPİTİNDE KULLANILAN SINIFLANDIRMA
ALGORİTMALARININ KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Volkan SEYMEN

Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ

Bu tez ~~24/12/2014~~ tarihinde aşağıdaki jüri tarafından oybirliği / oyçokluğu ile kabul edilmiştir.


Yrd.Doç.Dr. Deryim AKGÜN
Jüri Başkanı


Yrd.Doç.Dr. Ali GÜLBAG
Üye


Doç. Dr. Özdemir ÇETİN
Üye

TEŐEKKÜR

Tez alıőmamın gerekleőtirilmesinde byk bir sabırla yardım ve desteęini esirgemeyen, danıőman hocam Yrd. Do. Dr. Devrim AKGÜN'e ve bu srete bana bilgi ve deneyimleriyle katkı saęlayan tm hocalarıma katkılarından dolayı teőekkrlerimi sunarım.

Ayrıca tez alıőmam sresince her trl rahatsızlıklarıma sabır gsteren oda arkadaşlarıma anlayıőlarından dolayı teőekkrlerimi sunarım.

Ve son olarak bugnlere gelebilmemde ellerinden gelen hibir fedakrlıktan kaınmayan sevgili anneme, babama, abime ve kardeőime teőekkr bir bor bilirim.

İÇİNDEKİLER

TEŞEKKÜR	ii
İÇİNDEKİLER	iii
SİMGELER VE KISALTMALAR LİSTESİ	vi
ŞEKİLLER LİSTESİ	vii
TABLolar LİSTESİ	viii
ÖZET	ix
SUMMARY	x

BÖLÜM 1.

GİRİŞ	1
1.1. Literatür Özeti	3
1.2. Tez Çalışmasının Amacı	5
1.3. Tez Çalışmasının Katkısı	5
1.4. Tez Organizasyonu	6

BÖLÜM 2.

ANEMİ HASTALIĞI	7
2.1. Anemiye Genel Bakış	8
2.2. Anemi Hakkında	8
2.2.1. Anemilerin morfolojik olarak sınıflandırılması	9
2.3. Aneminin Nedenleri	10
2.3.1. Kan kaybı	10
2.3.2. Alyuvar hücrelerinin az üretilmesi	10
2.3.2.1. Beslenme	11
2.3.2.2. Hormonlar	11
2.3.2.3. Hastalıklar ve tedavileri	11
2.3.2.4. Hamilelik	12

2.3.2.5. Aplastik anemi	12
2.3.3. Yüksek oranda kırmızı kan hücresi yıkımı	12
2.4. Kimler Anemi Tarafından Risk Altındadır?	13
2.5. Aneminin Belirtileri	13
2.5.1. Anemide görülen komplikasyonları	14
2.6. Anemi Hastalığı Nasıl Tespit Edilir?	15
2.6.1. Tıbbi ve ailesel geçmiş	15
2.6.2. Fiziksel test	15
2.6.3. Tanı testleri ve işlemleri	16
2.6.3.1. Tam kan sayımı	16
2.6.3.2. Diğer test ve işlemler	16
2.7. Anemi Hastalığı Nasıl Tedavi Edilir?	17
2.7.1. Tedavinin hedefleri	17
2.7.2. Beslenme değişiklikleri ve takviyeleri	17
2.7.2.1. Demir	18
2.7.2.2. B12 vitamini	19
2.7.2.3. Folik asit	19
2.7.2.4. C vitamini	19
2.7.2.5. İlaçlar	20
2.7.3. Çeşitli tıbbi işlemler	20
2.7.3.1. Kan nakli	21
2.7.3.2. Kan ve kök hücre nakli	21
2.7.3.3. Ameliyat	21

BÖLÜM 3.

ANEMİ HASTALIĞININ TESPİTİ İÇİN KULLANILACAK ALGORİTMALAR 22

3.1. Yapay Sinir Ağları ve İleri Beslemeli Yapay Sinir Ağı	22
3.2. K-Ortalama Algoritması	28
3.3. Destek Vektör Makinesi (DVM)	29
3.4. Naive Bayes Algoritması	32
3.5. Regresyon Ağaçları	35
3.6. Holdout Yöntemi	36
3.7. ROC Analizi	36

BÖLÜM 4.

İSTATİSTİKSEL YÖNTEMLER YARDIMI İLE ANEMİ HASTALIĞININ TESPİTİ VE BU YÖNTEMLERİN ANALİZLERİ	38
4.1. Önışleme	39
4.2. İleri Beslemeli Yapay Sinir Ağı Testleri	40
4.3. Destek Vektör Makinesi Testleri	42
4.4. Regresyon Ağacı Testleri	44
4.5. Naive Bayes Testleri	46
4.6. K-ortalama algoritması Testleri	47

BÖLÜM 5.

SONUÇ VE DEĞERLENDİRMELER	50
5.1. Çalışmanın Faydaları	54
KAYNAKLAR	55
ÖZGEÇMİŞ	60

SİMGELER VE KISALTMALAR LİSTESİ

AIDS	: Acquired Immune Deficiency Syndrome
ANFIS	: Adaptive Neuro Fuzzy Inference System
CFN	: Cascade Forward Networks
DDN	: Distributed Delay Networks
DVM	: Destek Vektör Makinesi
FFN	: Feed Forward Network
FN	: False Negative
FP	: False Positive
HGB	: Hemoglobin
HCT	: Hematocrit
HIV	: Human Immunodeficiency Virus Infection
LVQ	: Learning Vector Quantization
MCH	: Mean Corpuscular Hemoglobin
MCHC	: Mean Corpuscular Hemoglobin Concentration
MCV	: Mean Corpuscular Volume
RBC	: Red Blood Cell
ROC	: Receiver Operating Characteristic
PNN	: Probabilistic Neural Network
TDN	: Time Delay Networks
TN	: True Negative
TP	: True Positive
SSE	: Sum of Squares Error
SVM	: Support Vector Machine
WHO	: World Health Organization
YSA	: Yapay Sinir Ağları

ŞEKİLLER LİSTESİ

Şekil 1.1. Ükelere göre farklı yaş aralıklarında demir eksikliği anemisi hastalığına bağlı yaşam süresi oranı (100,000 kişide bir) [10].....	2
Şekil 2.1. Demir eksikliği anemisine sahip bir hastanın periferik kan resmi [31]...	7
Şekil 2.2. MCV değerine göre anemilerin şematik bir gösterimi	9
Şekil 2.3. Demir eksikliği anemisine bağlı olarak göz solukluk, incelme ve kırılma oluşumları [38]	14
Şekil 3.1. Bir sinir hücresinin yapısı [14]	22
Şekil 3.2. İleri beslemeli geri yayımlı bir yapay sinir ağı modeli [46]	24
Şekil 3.3. İki kümeli bir k-ortalama algoritmasının örnek çalışma şeması	29
Şekil 3.4. DVM yapısı [50]	30
Şekil 3.5. Destek vektör şeması [51].....	31
Şekil 3.6. x_1 ve x_2 adında 2 sınıfa sahip Naive Bayes modeli. Kırmızı çizgi karar sınırını, diğer çizgilerde sınıfların kendi yoğunluklarını gösteren gauss modelleridir [56]	34
Şekil 3.7. Çalışmadan alınmış örnek bir regresyon ağacı şeması	36
Şekil 4.1. Çalışmada kullandığımız algoritma şeması	38
Şekil 4.2. Çalışmada kullandığımız ileri beslemeli geri yayımlı yapay sinir ağının algoritma şeması.....	40
Şekil 4.3. İleri beslemeli yapay sinir ağı test sonuçlarının grafiği	42
Şekil 4.4. DVM test sonuçlarının grafiği	44
Şekil 4.5. Regresyon ağacı test sonuçlarının grafiği	45
Şekil 4.6. Naive Bayes algoritması test sonuçlarının grafiği	47
Şekil 4.7. K-ortalama algoritması test sonuçlarının grafiği	49
Şekil 5.1. Tüm algoritmaların başarımlar oranları	51
Şekil 5.2. Tüm algoritmaların hassaslık oranları	51
Şekil 5.3. Tüm algoritmaların özgüllük oranları	52
Şekil 5.4. Tüm algoritmaların ortalamalarının karşılaştırılması	53

TABLolar LİSTESİ

Tablo 1.1. Yaş'a göre en düşük hemoglobin oranları [3].....	1
Tablo 1.2. Çeşitli kaynaklardan alınmış cinsiyete baęlı en düşük hemoglobin oranları	2
Tablo 1.3. Tam kan sayımı elemanlarının tam isimleri ve deęer aralıkları [24].....	3
Tablo 1.4. Literatürdeki sonuçların karşılaştırılması	5
Tablo 3.1. ROC parametreleri ve açıklamaları	37
Tablo 4.1. Veri setinden alınmış örnek veriler.....	39
Tablo 4.2. İleri beslemeli yapay sinir aęı tüm test sonuçları	41
Tablo 4.3. İleri beslemeli yapay sinir aęı ROC analizi tablosu	41
Tablo 4.4. İleri beslemeli yapay sinir aęı testlerinin ortalama deęerleri.....	42
Tablo 4.5. DVM tüm test sonuçları.....	43
Tablo 4.6. DVM ROC analizi tablosu.....	43
Tablo 4.7. DVM testlerinin ortalama deęerleri	44
Tablo 4.8. Regresyon aęacı tüm test sonuçları	44
Tablo 4.9. Regresyon aęacı ROC analizi tablosu	45
Tablo 4.10. Regresyon aęacı testlerinin ortalama deęerleri.....	46
Tablo 4.11. Naive Bayes algoritması tüm test sonuçları.....	46
Tablo 4.12. Naive Bayes algoritması ROC analizi tablosu.....	46
Tablo 4.13. Naive Bayes algoritması testlerinin ortalama deęerleri	47
Tablo 4.14. K-ortalama algoritması tüm test sonuçları.....	48
Tablo 4.15. K-ortalama algoritması ROC analizi tablosu.....	48
Tablo 4.16. K-ortalama algoritması testlerinin ortalama deęerleri	49
Tablo 5.1. Tüm algoritmaların ortalama deęerleri	50

ÖZET

Anahtar Kelimeler: Demir Eksiliği Anemisi Tespiti, Sınıflandırma Algoritmaları Karşılaştırması

Demir eksikliği anemisi dünyada en sık görülen anemi türüdür. Ortaya çıkma sebebi olarak vücudun ihtiyaç duyduğu demir oranının, yiyeceklerden alımının ve bağırsaklardan emiliminin az olması sonucu ortaya çıkar. Dünya nüfusu içerisinde orta düzeyde demir eksikliği anemisine sahip 610 milyona yakın kişinin olduğu bilinmektedir. Bu oran dünya nüfusunun yaklaşık %8.8'dir. Kadınlarda görülme oranı (%9.9), erkeklerde görülme oranından (%7.8) daha fazladır. Hafif derecede demir eksikliği anemisine sahip ise 375 milyon kişi olduğu bilinmektedir. Bu oran yaklaşık olarak %5.41'dir.

Anemi hastalığının tespiti için kan testlerinin bir doktorun incelemesi gereklidir. Bu çalışmada amaç, alınan kan testlerinin bir doktorun kontrolü yerine, bir sınıflandırma algoritması ile hızlıca yapılması sağlamaktır.

Bu çalışmada, MATLAB üzerinde RBC, HGB, HCT, MCV, MCH, MCHC gibi kan değerleri ile hastalık teşhisi yapılmıştır. 1800 hasta verisinden 359 u test aşamasında, 1441 i ise eğitim aşamasında kullanılmıştır. Problemin çözümünde ileri beslemeli yapay sinir ağı, destek vektör makinesi, Naive Bayes algoritması, k-ortalama algoritması ve regresyon ağaçları yöntemleri kullanılmıştır. Sınıflandırma algoritmaları MATLAB ortamında kullanılmıştır. Son aşamada, kullandığımız algoritmaların sonuçları arasında ROC eğrisi yöntemi kullanılarak başarımların analizi yapılmıştır.

Gerçekleştirilen deneyde elde edilen sonuçlara göre, en yüksek başarımların oranı %99.53 olarak ileri beslemeli yapay sinir ağı tarafından elde edilmiştir.

THE COMPARISON OF CLASSIFICATION ALGORITHMS FOR DIAGNOSIS OF IRON-DEFICIENCY ANEMIA

SUMMARY

Keywords: The Diagnosis Of Iron-Deficiency Anemia, Comparison of Classification Algorithms

Iron-deficiency anemia is the most common anemia disease in the world. It is caused by insufficient dietary intake and absorption of iron or iron loss from bleeding. A moderate degree of iron-deficiency anemia affects about 610 million people. This rate is %8.8 of world population. Iron-deficiency anemia is more common in female (%9.9) than males (%7.8). A mild of iron-deficiency anemia affects another 375 million people. This rate is about %5.41.

In diagnosis of iron-deficiency anemia, a physician analyzes subject's complete blood test. In the implemented network model was automated for the detection of iron-deficiency anemia disease by using feedforward backpropagation neural network.

In this study, subject's disease condition was decided by using RBC, HGB, HCT, MCV, MCH, MCHC values of subjects' complete blood test on MATLAB. 359 subjects' data were used for test and 1441 subjects' data were used for training within 1800 iron-deficiency anemia samples. In this study, feedforward neural network, support vector machine, Naive Bayes algorithm, k-means algorithm and regression trees methods were used for problem solving. Classification algorithms were used on MATLAB. In the last step, ROC curve method was used for performance analysis of used algorithms' results.

The best success rate of performed tests have been achieved %99.53 by feedforward neural network.

BÖLÜM 1. GİRİŞ

Demir eksikliği anemisi dünyada en sık görülen anemi tipidir. Ortaya çıkma sebebi olarak vücudun ihtiyaç duyduğu demir oranının, yiyeceklerden alımının ve bağırsaklardan emiliminin az olması sonucu ortaya çıkar [1]. Dünya nüfusu içerisinde orta düzeyde demir eksikliği anemisine sahip 609 milyona yakın kişinin olduğu bilinmektedir. Bu oran dünya nüfusunun yaklaşık %8.8'dir. Kadınlarda görülme oranı (%9.9), erkeklerde görülme oranından (%7.8) daha fazladır. Hafif derecede demir eksikliği anemisine sahip ise 375 milyon kişi olduğu bilinmektedir. Bu oran yaklaşık %5.41'dir [2].

Tablo 1.1. Yaşa göre en düşük hemoglobin oranları [3]

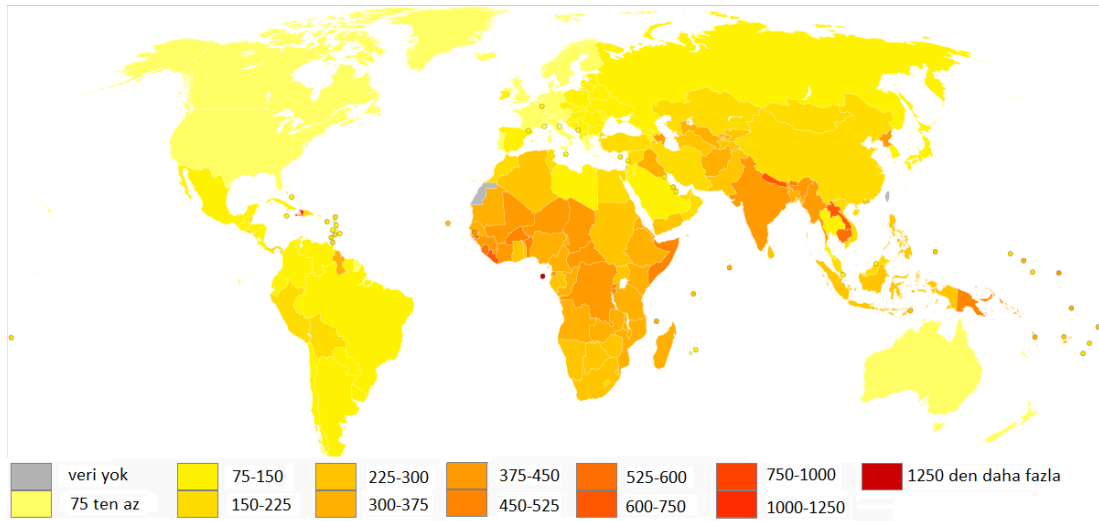
Açıklama	Değer
6 aylık – 6 yaş arası çocuk	11
6 – 14 yaş arası çocuk	12
Yetişkin erkekler	13
Hamile olmayan yetişkin erkekler	12
Hamile yetişkin kadınlar	11

Dünya Sağlık Örgütüne göre anemi hastalığının tespit edilebilmesi için hastanın hemoglobin konsantrasyonu ve hematokrit oranı bilinmelidir. Hemoglobin konsantrasyonunun cinsiyete bağlı olarak belirli oranların altında olmaması gerekir. Bu oranla birlikte hematokrit oranı da belirli bir değer altında olmamalıdır. Hematokrit oranı kadınlarda %39, erkeklerde %36, hemoglobin konsantrasyonu ise erkeklerde %13, kadınlarda ise %12 nin altına düşmemelidir. Bunları bu aralıkların altına düşmesi anemi hastalığının göstergesidir [3].

Tablo 1.2. Çeşitli kaynaklardan alınmış cinsiyete bağlı en düşük hemoglobin oranları

Kaynak	Kadın (g/dL)	Erkek (g/dL)
Blanc [3]	12	13
Rapaport [4]	12	14
Tietz [5]	11.7	13.2
Jandl [6]	12.2	14.2
Goyette [7]	11.7	13.2
Lee [8]	11.6	13.2
Beutler [1]	12.3	14.0
Hoffman [9]	12.0	13.5

Anemi hastalığının tespiti için kan testlerinin bir doktorun incelemesi gereklidir. Bu çalışmada alınan kan testlerinin bir doktorun kontrolü yerine, ileri beslemeli yapay sinir ağı, destek vektör makinesi, Naive Bayes algoritması, regresyon ağaçları, k-ortalama algoritması gibi yöntemler kullanılmıştır.



Şekil 1.1. Ükelere göre farklı yaş aralıklarında demir eksikliği anemisi hastalığına bağlı yaşam süresi oranı (100,000 kişide bir) [10]

Farklı sınıflandırma yöntemleri ile çok çeşitli problemler kolaylıkla ve etkili bir şekilde çözülebilmektedir. İstatistiksel yöntemlerin bazıları, yapay sinir ağları, bulanık mantık, destek vektör makinesi (DVM), k-ortalama algoritması, genetik algoritmalarıdır. Bu algoritmalar, geleneksel yöntemlerden daha hızlı ve daha az maliyetlerle aynı problemlere çözüm üretebilmektedirler. Yapay sinir ağları (YSA),

örüntü tanıma alanında en yaygın kullanılan yöntemlerden birisidir. Yapay sinir ağlarının kullanım alanları neredeyse sınırsızdır. Bu alanlar genel olarak 6 ana başlığa ayrılabilir. Sayısal işaret işleme, lineer transformasyonlar, örüntü sınıflandırma, tanıma, zaman serisi modelleme, sistem tanımlama [11]. İleri beslemeli yapay sinir ağları ise en çok bilinen ve kullanılan yapay sinir ağı modelidir. Geri yayılım ise en çok kullanılan eğitim algoritmasıdır. DVM, k-ortalama algoritması, Naive Bayes algoritması, regresyon ağaçları gibi yöntemlerinde kullanım alanları çok geniştir. Bu yöntemlerin bazılarının kullanım alanlarına örnek vermek gerekirse; kanser hastalığı tespitinde YSA kullanımı [12], YSA ile örüntü tanıma [13][14], YSA ile optik karakter tanıma [15], YSA ile güneş termal sisteminin performans tahmini [16], DVM ile örüntü tanıma [17][18], Naive Bayes ile Çince yazı sınıflandırması [19], regresyon ağacı ile ekolojik verinin modellenmesi [20] ve benzeri örnekleri çoktur. Bazı araştırmacıların kıyaslamaları sonucu klasik istatistiksel metotlar arasında yapay sinir ağlarının probleme yönelik daha iyi genelleştirme yapabildiği ortaya koyulmuştur [21][22][23].

Tablo 1.3. Tam kan sayımı elemanlarının tam isimleri ve değer aralıkları [24]

Açıklama	Değerler En Düşük – En Yüksek
Red Blood Cell (RBC)	4.5 - 6
Hemoglobin (HGB)	12 – 16
Hematocrit (HCT)	36 – 48
Mean Corpuscular Volume (MCV)	80 – 100
Mean Corpuscular Hemoglobin (MCH)	27 - 34
Mean Corpuscular Hemoglobin Concentration (MCHC)	31 - 37

1.1. Literatür Özeti

Literatür de anemi hastalığının tespitine yönelik kullanılan istatistiksel yöntemler arasında ileri beslemeli geri yayımlı yapay sinir ağı (FFN), olasılıksal sinir ağı (PNN), vektör parçalama yöntemi (LVQ), zaman gecikmeli sinir ağı (TDN), ileri beslemeli kademeli korelasyon ağı (CFN) ve dağıtık sistem temelli gecikmeli sinir ağı (DDN), bulanık mantık, yapay bağışıklık sistemi, adaptif ağ tabanlı bulanık mantık çıkarım sistemi (ANFIS) ve lojistik regresyon analizi kullanılmıştır.

Wongseree ve arkadaşları yapay sinir ağıları yardımıyla talasemi (akdeniz anemisi) hastalığının tespiti için yaptıkları testler sonucunda 10 ayrı talasemi hastalığının sınıflandırılması için yapılan testlerde ortalama olarak %90 başarımla elde edildiğini bildirmişlerdir [25].

Birndorf ve arkadaşları üç çeşit mikrositik anemi hastalığı üzerinde durmuşlardır. Bunlar: demir eksikliği anemisi, hemoglobinopati (orak hücre anemisi, talasemi gibi), kronik hastalık anemisi. Bu anemiler üzerinde yapay sinir ağıları ile yapılan testler sonucunda %96.5 oranında başarımla elde edildiğini bildirmişlerdir [26].

Azarkhish ve arkadaşları adaptif ağı tabanlı bulanık mantık çıkarım sistemi, lojistik regresyon ve yapay sinir ağı kullanarak demir eksikliği anemisi üzerinde yaptıkları testlerde lojistik regresyon analizi ile %62.96 başarımla, yapay sinir ağıları ile %96.29 başarımla ve adaptif ağı tabanlı bulanık mantık çıkarım sistemi ile ise %90.74 oranında başarımla elde edildiğini bildirmişlerdir [27].

Yılmaz ve arkadaşları FFN, PNN, LVQ, TDN, CFN ve DDN ağılarını kullanarak demir eksikliği anemisi üzerinde yaptıkları testler sonucunda, LVQ başarımla sonucu %99.33, CFN başarımla sonucu %98.95, DDN başarımla sonucu %99.16, TDN başarımla sonucu %99.16, PNN başarımla sonucu %98.52, FFN başarımla sonucu %99.16 olarak elde edildiğini bildirmişlerdir [24].

Yılmaz ve arkadaşları bulanık mantık kullanarak demir eksikliği anemisi üzerinde yaptıkları testler sonucunda yaklaşık %89.5 oranında başarımla elde edildiğini bildirmişlerdir [28].

Yurtay ve arkadaşları, yapay bağışıklık sistemi kullanarak demir eksikliği anemisi üzerinde yaptıkları testler sonucunda %98 oranında başarımla elde edildiğini bildirmişlerdir [29].

Tablo 1.4. Literatürdeki sonuçların karşılaştırılması

Kaynak	Yöntem	Başarım Oranları (%)
Wongseree ve arkadaşları [25]	YSA	90
Birndorf ve arkadaşları [26]	YSA	96.5
Azarkhish ve arkadaşları [27]	YSA	96.29
Azarkhish ve arkadaşları [27]	ANFIS	90.74
Azarkhish ve arkadaşları [27]	Lojistik Regresyon	62.96
Yılmaz ve arkadaşları [24]	FFN	99.16
Yılmaz ve arkadaşları [24]	CFN	98.95
Yılmaz ve arkadaşları [24]	DDN	99.16
Yılmaz ve arkadaşları [24]	TDN	99.16
Yılmaz ve arkadaşları [24]	LVQ	99.33
Yılmaz ve arkadaşları [24]	PNN	98.52
Yılmaz ve arkadaşları [28]	Bulanık Mantık	89.5
Yurtay ve arkadaşları [29]	Yapay Bağışıklık Sistemi	98,00

1.2. Tez Çalışmasının Amacı

Bu çalışmada, demir eksikliği anemisi tespiti için çeşitli istatistiksel metotlar ile bu sistemin tasarımı gerçekleştirilmiştir. Hazırlanan sistemimizde hastaların RBC, HGB, HCT, MCV, MCH, MCHC kan değerleri kullanılarak bu değerler MATLAB üzerinde işlenmiştir. Bu işlemlerin sonucunda kişinin hasta ya da hasta olmadığı tespiti yapılmıştır. Gerçekleştirilen deneylerde 359 hasta test verisi için, 1441 hasta eğitim verisi için kullanılmıştır. Gerçekleştirilen deneyde elde edilen sonuçlara göre, en yüksek başarımlı oranı %99.53 olarak ileri beslemeli yapay sinir ağı tarafından elde edilmiştir.

1.3. Tez Çalışmasının Katkısı

Bu tez çalışmasının literatüre katkıları şunlardır:

1. Demir eksikliği anemisinin tespiti için birçok yöntem kullanılarak medikal alanında bu yöntemlerin başarımlı oranının yüksek olduğu belirlenmiştir.

2. İstatistiksel yöntemler arasından ileri beslemeli yapay sinir ağlarının benzer çalışmalarda kullanılabilceđi ve diđer yöntemlere oranla daha iyi sonuçlar elde edilebileceđi tespit edilmiştir.
3. Son olarak da bu ve benzeri hastalıkların tespitinde istatistiksel yöntem ile yüksek başarımlar elde edilebileceđi belirlenmiştir.

1.4. Tez Organizasyonu

Tezin diđer bölümleri şu şekilde organize edilmiştir. Bölüm 2 de anemi hastalığı hakkında detaylı bilgilere yer verilmiştir. Bölüm 3 de bu hastalığın tespitinde ve diđer işlemlerde kullanacağımız ileri beslemeli geri yayımlı yapay sinir ađı, destek vektör makinesi (DVM), k-ortalama algoritması, Naive Bayes algoritması, regresyon ağaçları, Holdout yöntemi ve ROC analizi hakkında bilgiler verilmiştir. Bölüm 4 de hazırlanan sistemlere ait bilgiler ve deneysel sonuçlar sunulmuştur. Bölüm 5 de ise çalışmadan elde edilen sonuçlar, karşılaştırmaları açıklanmıştır.

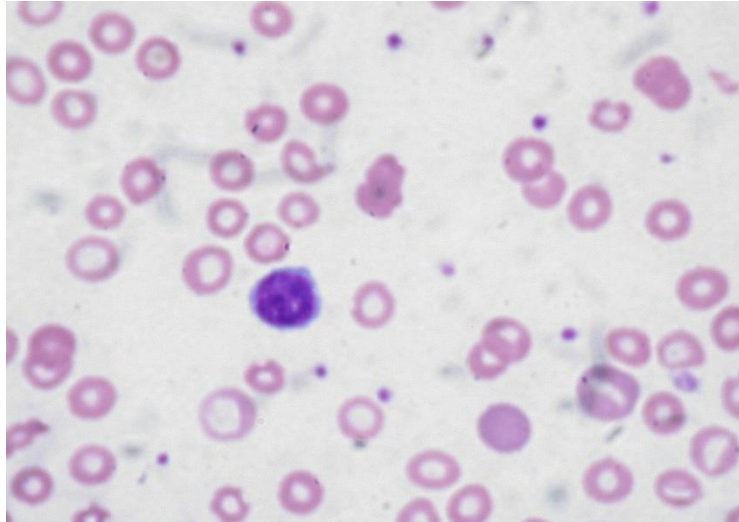
BÖLÜM 2. ANEMİ HASTALIĞI

Anemi, toplam kırmızı kan hücre sayısının (alyuvar) normalden daha düşük olması veya eritrositlerin içerisindeki hemoglobin oranının az olmasından kaynaklanan bir hastalıktır [30].

Hemoglobin kana kırmızı rengini veren demir zengini bir proteindir. Bu protein kırmızı kan hücrelerinin akciğerlerden tüm vücuda oksijen taşımaya yardımcı olur.

Anemi hastası bir kişide, vücut yeterince oksijen içeren kan alamaz. Bunun sonucu olarak da kişi yorgun ya da zayıf hisseder. Bununla birlikte başka belirtiler de olabilir. Örnek olarak; nefes darlığı, baş dönmesi veya baş ağrıları gibi.

Ağır veya uzun süreli anemiler kalbe, beyne ve vücuttaki diğer organlara zarar verir. Çok şiddetli anemiler ölümlere yol açabilir.



Şekil 2.1. Demir eksikliği anemisine sahip bir hastanın periferik kan resmi [31]

2.1. Anemiye Genel Bakış

Kan, kırmızı kan hücreleri (alyuvar), beyaz kan hücreleri (akyuvar), trombositler ve plazmadan (kanın akışkan kısmı) oluşur.

Alyuvarlar ortasında delikler olmayan çörelere veya disk şekline benzerler. Alyuvarlar vücuda oksijen taşıırken aynı zamanda da bir atık ürün olan karbodioksitten vücuda arındırırılar. Bu hücreler ve trombositler süngerimsi doku denilen kısımdaki kırmızı kemik iliklerinde üretilirler. Trombositler yaralanmalarda, kesiklerde kanın pıhtılaşmasına ve kanamayı durdurmaya yararlar.

Akyuvarlar ise kemik içerisindeki sarı kemik iliklerinde üretilirler. Bu hücreler enfeksiyonlara karşı savaşırılar. Bazı tip anemiler bu üç kan hücresi çeşidinin az bulunmasından kaynaklanmaktadır.

Anemilerin temel olarak üç nedeni vardır; kan kaybı, alyuvar hücrelerinin az üretilmesi veya yüksek oranda kırmızı kan hücresi yıkımıdır.

2.2. Anemi Hakkında

Birçok çeşit anemi, hafif seviyeli, orta seviyelidir ve bu anemiler kolaylıkla tedavi edilebilir. Hatta sağlıklı beslenme ile bunların bazıları engellenebilir. Diğer çeşit hafif ve orta seviyeli anemiler ise beslenme takviyeleri ile çözülebilir.

Bunlara rağmen, bazı çeşit anemiler ağır, uzun ömürlü ve hatta tespit edilmeyip, tedavi edilmezse hayati tehlikesi olan anemilerdir. Anemi belirtileri bir kişide ortaya çıkmışsa mutlaka bir doktora görülmelidir. Tedavisi şiddetine ve aneminin sebebine bağlı olarak değişecektir.

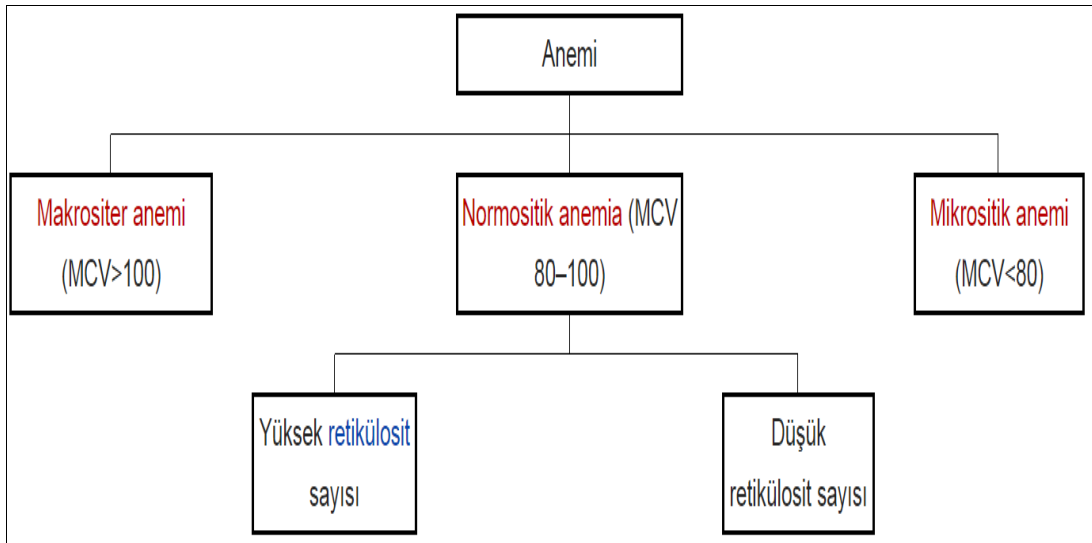
Aneminin diğer adları olarak, yorgun kan, kan eksikliği ve kandaki demir eksikliği de denilmektedir [32].

Birçok çeşit anemi türü bulunmaktadır. Bu türlerin kendine özel nedenleri ve özellikleri bulunmaktadır. Bunlardan bazıları;

- Aplastik anemi
- Cooley anemisi
- Diamond-blackfan anemisi
- Fanconi anemisi
- Folik asit eksikliğine bağlı anemi
- Hemolitik anemi
- Demir eksikliği anemisi
- Penisyöz anemisi
- Orak hücreli anemi
- Akdeniz anemisi [33]

2.2.1. Anemilerin morfolojik olarak sınıflandırılması

Temel olarak anemilerin morfolojik sınıflandırması MCV değerine göre yapılır. Anemileri morfolojik olarak sınıflandırma şematik gösterimi Şekil 2.2 deki gibidir.



Şekil 2.2. MCV değerine göre anemilerin şematik bir gösterimi

Makrositer anemiler arasında megaloblastik anemiler, normokrom normositik anemiler arasında akut kan kaybı, hipokrom mikrositer anemiler içinde demir eksikliği anemisi sayılabilir [34].

2.3. Aneminin Nedenleri

Anemiye üç ana sebep neden olmaktadır.

1. Kan kaybı,
2. Alyuvar hücrelerinin az üretilmesi,
3. Yüksek oranda kırmızı kan hücresi yıkımı

2.3.1. Kan kaybı

Kan kaybı, anemi hastalığının özellikle demir eksikliği anemisinin en çok karşılaşılan nedenlerinden birisidir. Kan kaybı kısa süreli ya da sürekliliği olan bir durumdan kaynaklanabilir.

Aylık adet dönemleri veya kanamaları, sindirim veya idrar yollarındaki iç kanamalar bu kan kayıplarına neden olabilir. Ayrıca ameliyat, travma veya kanser hastalığı da kan kaybına neden olabilir.

Eğer çok fazla kan kaybı yaşanmışsa, vücut ihtiyacı olandan daha az alyuvara sahip olduğundan dolayı anemi hastalığı ortaya çıkabilir.

2.3.2. Alyuvar hücrelerinin az üretilmesi

Kalıtımdan gelen ya da sonradan oluşan durumlar vücudun ihtiyaç duyduğu alyuvarlar sayısının üretilmesini engelleyebilir. “Sonradan oluşan” ifadesi ile anlatılan bu durumun doğuştan gelmediği yani sonradan oluştuğudur. “Kalıtımdan” ifadesi ile anlatılmak istenen ise bu özelliğin ebeveynlerden kalıtım yolu (genler) ile kişiye geçtiğidir.

Sonradan oluşan durumlarda, yetersiz beslenme, anormal hormon seviyeleri ve bazı kronik hastalıklar ile hamilelik anemi hastalığının oluşmasında etkilidir.

Kalıtımdan gelen ya da sonradan oluşan aplastik anemi de vücudun ihtiyaç duyduğu alyuvarlar sayısının üretilmesini engelleyebilir.

2.3.2.1. Beslenme

Demir minerali, folik asit veya B12 vitamini eksik olan bir beslenme, vücudun ihtiyaç duyduğu alyuvarlar sayısının üretilmesini engelleyebilir. Alyuvar üretimi için vücut düşük miktarda da olsa C vitamini, B2 vitamini ve bakıra ihtiyaç duymaktadır.

Vücutta besinlerin yeterince emilmemesi ile de vücudun ihtiyaç duyduğu alyuvarlar sayısının üretilmesini engellenebilir.

2.3.2.2. Hormonlar

Vücut alyuvar üretimi için eritropoetin hormonuna ihtiyaç duymaktadır. Bu hormon kemiklerdeki kırmızı kan iliklerini alyuvar üretmesi için uyarır. Bu hormonun normalden düşük seviyelerde olması anemiye neden olur.

2.3.2.3. Hastalıklar ve tedavileri

Böbrek hastalıkları ve kanser gibi kronik hastalıklar, vücudun alyuvar üretimini zorlaştırmaktadırlar.

Bazı kanser tedavileri kemik iliklerine veya kırmızı kan hücrelerinin oksijen taşıma kabiliyetlerine zarar verir. Eğer ki kemik ilikleri zar görürse, yıkılan alyuvarı sayısı kadar alyuvar üretimi hızlı olmaz.

HIV/AIDS [35] hastaları, enfeksiyonlar veya kullandıkları ilaçlar nedeniyle anemi hastalığına yakalanabilirler.

2.3.2.4. Hamilelik

Hamilelik süresince, kandaki vücudun ihtiyaç duyduğu orandan daha düşük demir ve folik asit oranlarından dolayı anemi hastalığı ortaya çıkabilir.

Hamileliğin ilk 6 ayı boyunca, kadınların kanındaki sıvı kısım (plazma), alyuvar üretiminden çok daha hızlı bir şekilde artar. Bu kandaki sulanma nedeniyle anemi ortaya çıkabilir.

2.3.2.5. Aplastik anemi

Bazı bebekler yeterince alyuvar üretme kabiliyetleri olmadan dünyaya gelirler. Bu durum aplastik olarak adlandırılır. Aplastik anemiye sahip olan bebekler ve çocuklar kanlarındaki alyuvar sayılarının artırılması için bazen kan nakline ihtiyaç duyarlar.

Bazı tedaviler, toksinler ve bulaşıcı hastalıklar da doğum dışında da aplastik anemiye neden olabilir.

2.3.3. Yüksek oranda kırmızı kan hücresi yıkımı

Kalıtımdan gelen ya da sonradan oluşan durumlar vücudun normalden daha fazla alyuvar yıkımı yapmasına neden olabilir. Sonradan oluşan durumlara bir örnek olarak dalak genişlemesi ya da hastalığıdır.

Dalak yıpranmış alyuvarları vücuttan temizleyen bir organdır. Dalak genişlemiş ya da hastalanmışsa, normalden daha fazla alyuvar hücresinin vücuttan temizlenmesine neden olabilir. Bu da anemi hastalığına neden olabilir.

Kalıtımdan gelen durumlara örnek verecek olursak, orak hücreli anemi, akdeniz anemisi ve bazı enzimlerin üretimindeki eksiklikler normalden daha fazla alyuvarın yıkımına neden olabilir. Bu durumlarda üretilen sağlıklı alyuvar sayısından çok daha hızlı bir şekilde alyuvar ölümlerinin gerçekleşmesine neden olur.

Bir diđer örnek ise hemolitik anemidir. Bu anemide vücutta normalden daha fazla alyuvar yıkımına neden olur. Kalıttımdan gelen ya da sonradan oluşan bazı durumlar hemolitik anemiye neden olabilir. Örnek olarak; bağıřıklık sistemindeki düzensizlikler, enfeksiyonlar, bazı tedaviler veya kan nakline karşı oluşan reaksiyonlar anemiye neden olabilecek durumlar arasındadırlar [36].

2.4. Kimler Anemi Tarafından Risk Altındadır?

Anemi çok yaygın bir hastalıktır. Tüm yaşlarda, ırklarda ve etnik gruplarda görülebilir. Erkeklerde ve kadınlarda anemi hastası olabilir. Buna rağmen, kadınlar adet dönemleri başladığı andan itibaren yüksek risk altındadırlar.

Bazı bebekler ilk yıllarında demir eksikliğine bağılı anemi için risk altındadırlar. Bu risk, erken doğan, sadece sütle beslenen veya demir ile takviye edilmemiş besinlerle beslenen bebekler içindir. Bu bebeklerde ilk 6 ay içerisinde demir eksikliği anemisi ortaya çıkması söz konusudur.

1 ve 2 yaş arasındaki bebeklerde anemi konusunda risk altındadırlar. Aşırı inek sütü içilmesi sonucu oluşan, yeterince demir minerali içermeyen beslenme şekli ile anemi ortaya çıkması söz konusudur. Çünkü inek sütü büyüme için yeterli olan demir mineralini içermez. Bunun yanında demir minerali içeren besinler alınması ile bu durum çözülebilir.

Yaşlılarda anemi riski altındadırlar. Arařtırmacılar anemiye yaşlılardaki hangi durumların etkilediğini arařtırmaktadırlar [37].

2.5. Aneminin Belirtileri

Aneminin en fazla karşılaşılan belirtisi tükenmişliktir (yorgun hissetme ve zayıflık hissi). Anemiye sahip olan bir kişi normal aktivitelerini yapacak enerjiyi bile bulmakta zorlanır.



Şekil 2.3. Demir eksikliği anemisine bağlı olarak göz solukluk, incelme ve kırılma oluşumları [38]

Aneminin diğer belirtileri ise şunlardır:

- Nefes darlığı
- Baş dönmesi
- Baş ağrısı
- Ayaklarda ve ellerde üşüme
- Deride solgunluk (Şekil 2.3)
- Boyun ağrısı

Kalbin vücuda yeterince oksijen taşıyabilmek için daha çok çalışması sonucu bu işaretler ve belirtiler görülebilir. Hafif ya da orta şiddette bir anemide çok küçük belirtiler olur ya da hiç olmayabilir.

2.5.1. Anemide görülen komplikasyonları

Bazı anemi hastalarında kalp atış ritminde bozukluk görülebilir. Bu bozukluklar kalbin aşırı çalışmasına neden olduğundan dolayı kalbin zarar görmesine ve buna bağlı olarak kalp krizi geçirmesine neden olabilir.

Anemi aynı zamanda diğer organları da etkileyebilir. Çünkü anemiden dolayı organlara yeterince oksijen taşınmamaktadır.

Anemi, HIV / AIDS hastaları ya da kanser hastalarının zayıf düşmesine neden olur. Bu da tedavilerinin iyi bir şekilde ilerlememesine neden olur.

Anemi bazı başka hastalıklara da neden olmaktadır. Böbrek hastalığı ve anemi hastalığı olan hastalar kalp problemleri oluşumlarına daha yatkındırlar. Bazı tür anemiler az sıvı alımı veya çok sıvı kaybına bağlı olarak ortaya çıkabilir. Bu sıvı kayıpları bazen hastanın yaşamını tehdit edebilir [39].

2.6. Anemi Hastalığı Nasıl Tespit Edilir?

Anemi hastalığı, bazı fiziksel testler, aile geçmişi, tıbbi geçmiş ve bazı kan testleri yardımı ile tespit edilebilir.

2.6.1. Tıbbi ve ailesel geçmiş

Anemi belirtilerinin görülüp görülmediği, anemiye sebep olabilecek bir hastalık geçirilip geçirilmediği gibi sorular sorularak anlaşılmaya çalışılır. Bunun dışında anemiye sebep olabilecek bir ilaç alınıp alınmadığı araştırılır. Aynı zamanda ailenizde anemi hastalığının olup olmadığı araştırılır.

2.6.2. Fiziksel test

Aneminin tespiti için yapılacak fiziksel testler şunlardır:

- Kalp atım hızlarının düzensiz mi veya hızlı mı olduğunu dinleyebilir.
- Akciğerlerdeki nefes alımının düzensiz mi veya hızlı mı olduğunu dinleyebilir.
- Karaciğer ve dalak boyutları kontrolü yapılabilir.
- Aynı zamanda kanama olabilme ihtimali yüksek bölgeler olan rektal ve jinekolojik bölgeler kontrol edilebilir.

2.6.3. Tanı testleri ve işlemleri

Aneminin türünü ve şiddetini ölçtürmek için bazı kan testleri ve diğer testler yapılması gerekebilir.

2.6.3.1. Tam kan sayımı

Anemi tespiti için genellikle ilk yapılan test tam kan sayımıdır. Kanımızdaki birçok değerlerin ölçülmesi işlemidir.

Bu test ile hemoglobin ve hematokrit seviyelerini ölçülür. Hemoglobin kandaki oksijeni damarlar yardımıyla vücudun tüm bölgelerine taşıyan ve demir yoğunluğu yüksek proteindir. Hematokrit ise kandaki alyuvarların oran ölçüsüdür. Bu değerlerin düşük çıkması anemi göstergesidir.

Tam kan sayımı ile aynı zamanda akyuvar ve trombosit sayılarına da bakılır. Anormal sonuçlar anemi veya başka bir hastalığın göstergesidir.

Son olarak, tam kan sayımında MCV ye bakılır. MCV, kırmızı kan hücrelerinin ortalama hacmini temsil eder. Örnek olarak demir eksikliği anemisinde alyuvarlar normalden daha küçük olurlar.

2.6.3.2. Diğer test ve işlemler

Eğer ki tam kan sayımı sonucunda anemi hastalığına sahip sonucu çıkarsa diğer bazı testler yapılarak bu sonuç kesinleştirilir. Bu testler örnek olarak şunlardır:

- Hemoglobin elektroforezi. Bu testin amacı kandaki farklı hemoglobin türlerine bakmaktır. Bu testin sonucunda hangi tür anemiye sahip olduğunuz tespit edilir.
- Retikülot sayımı. Bu testin amacı kandaki genç akyuvar sayısını tespit etmektir. Bu test sonucunda kemikteki kırmızı ilik hücrelerinin doğru sayıda alyuvar üretip üretmediği tespit edilir.

- Vücuttaki ve kandaki demir oranı testleri. Bu test serum demiri ve serum ferritin testlerini de içermektedir. Toplam demir bağlama kapasitesi testi ve transferin seviyesi de demir seviyesi ölçümlerinde kullanılmaktadırlar.

Aneminin birçok sebebi olduğundan dolayı, böbrek yetmezliği, kurşun zehirlenmesi (çocuklarda) ve vitamin eksiklikleri testleri de yapılması gerekebilir.

Bunlar dışında bir neden olarak iç kanamadan dolayı hastalığın oluştuğundan şüpheleniliyorsa bunun için dışkı testi yaptırılabilir. Eğer ki, bu test sonucunda kan tespit edilirse endoskopi gibi yöntemler ile bu kanamanın kaynağı tespit edilebilir.

Bunların dışında kan iliği testi ile iliklerin sağlıklı mı ve yeterince alyuvar hücresi ürettiği mu durumları kontrol edilebilir [40].

2.7. Anemi Hastalığı Nasıl Tedavi Edilir?

Anemi hastalığının tedavisi türüne, sebebine ve şiddetine göre değişir. Hastalığın tedavileri arasında beslenme değişiklikleri ya da takviyeleri, ilaçlar, çeşitli tıbbi işlemler veya ameliyat bulunmaktadır.

2.7.1. Tedavinin hedefleri

Tedavilerde hedeflenen kanın taşıyabildiği oksijen oranını artırmaktır. Bu işlem alyuvar sayıları artırılarak ya da hemoglobin oranları artırılarak yapılır. Diğer bir hedefi ise bu hastalığı ortaya çıkaran diğer sebepleri ortadan kaldırmaktır.

2.7.2. Beslenme değişiklikleri ve takviyeleri

Düşük vitamin ve demir alımları anemiye sebep olabilmektedir. Bu düşük mineral alımının sebebi de kötü beslenme veya bazı hastalık ya da durumların sonucunda oluşabilmektedir.

Bu mineral seviyelerinin artırılması için beslenme deęişiklięi veya vitamin ve dięer mineral takviyeleri yapılması gerekmektedir. Genel olarak takviye vitaminler B12 ve folik asittir. Bazen C vitamini verilerek vücudun demir emilimi kolaylaştırılır.

2.7.2.1. Demir

Hemoglobin üretimi için vücut demire ihtiyaç duyar. İnsan vücudu demiri etlerden, sebze ve dięer yiyeceklerden olduęundan daha kolay emer. Anemi hastalığının tedavisi amacıyla özellikle kırmızı et yenilmesi tavsiye edilir.

Et dışında demir içeren çeşitli besinler şunlardır:

- Ispanak ve dięer koyu yeşil renkli lifli sebzeler
- Soya peyniri
- Bezelye, mercimek (kırmızı, beyaz), soya fasulyesi ve nohut
- Kurutulmuş meyveler (kuru erik, kuru üzüm ve kayısı)
- Erik suyu
- Demir bakımından zenginleştirilmiş tahıllar ve ekmekler

Yenilen yiyeceklerin besin deęeri tablosuna bakılarak içerisinde demir bakımından zengin olan yiyeceklerin seçilmesi aneminin tedavisi için tavsiye edilenler arasındadır.

Demir ayrıca takviye olarak alınabilmektedir. Demirle birlikte dięer mineral ve çoklu vitaminler alınarak emilimi artırmak amaçlanmaktadır.

Özellikle erken doğum, doğum ve çok fazla inek sütü içen çocuklarda ve sadece anne sütüyle beslenen çocuklarda demir takviyesi önerilmektedir.

Çok fazla demir alımı zararlı olabilir. Bu nedenle demir alımı ancak bir doktor kontrolünde yapılması gereklidir.

2.7.2.2. B12 vitamini

Düşük seviyedeki B12 vitamin oranı ölümcül bir anemi türüne neden olabilmektedir. Bu anemi türü genellikle B12 vitamin takviyesi ile tedavi edilebilmektedir

B12 vitamini içeren çeşitli besinler şunlardır:

- B12 vitamini bulunduran kahvaltı gevreği
- Ciğer eti, sığır eti, kümes hayvanı etleri ve balıklar
- Yumurtalar ve süt ürünleri (yoğurt, süt ve peynir gibi)
- B12 ile takviye edilmiş soya bazlı içecekler ve vejeteryan burgerleri

2.7.2.3. Folik asit

Folik asit, B vitamininin yiyeceklerde bulunan bir çeşitidir. Folik asit vücutta yeni hücre yapımı ve koruması için gereklidir. Folik asit aynı zamanda hamile kadınlar içinde çok önemlidir. Çünkü bu vitamin ceninin sağlıklı büyümesine yardım eder.

Folik asit içeren çeşitli besinler şunlardır:

- Ekmek, hamur işleri ve folik asit eklenmiş pirinçi
- Ispanak ve diğer koyu yeşil renkli lifli sebzeler
- Börülce ve kuru fasulye
- Sığır ciğeri
- Yumurta
- Muz, portakal, portakal suyu ve diğer meyve ve sebzeler

2.7.2.4. C vitamini

C Vitamini vücudun demir emilimine yardımcı olur. C vitamini bakımından zengin olan yiyecekler arasında meyveler ve sebzeler önde gelir. Özellikle turunçgillerde yüksek oranlarda bulunur. Turunçgiller; portakal, greyfurt, mandalina ve benzeri

meyveleri içerir. Taze ve donmuş meyveler, sebzeler ve suları C vitamini bakımından konserve olanlardan daha zengindirler.

C vitaminini ilaç olarak almak yerine greyfurt yenmesi çoğu kez daha etkilidir.

C vitamini bakımından zengin diğer meyveler ise kivi, çilek, domates ve kavundur.

Sebzeler arasında C vitamini bakımından zengin sebzeler ise brokoli, biberler, brüksel lahanası, lahana, patates ve yeşil yapraklı sebzelerdir(şalgam ve ıspanak gibi).

2.7.2.5. İlaçlar

Anemi hastalığının tedavisi ya da anemi hastalığına sebep olan hastalığın tedavisi için ilaç kullanılması gerekebilir. Bu ilaçlardan bazıları şunlardır:

- Enfeksiyonları tedavi için antibiyotikler
- Genç ve yetişkin kadınlardaki ağır adet kanamalarını engellemek için hormonlar
- İnsan yapımı eritropoetin (Bu protein ile kırmızı kan hücrelerinin yapımı tetiklenir. Bu tedavinin tehlikeleri olmakla birlikte doktor gözetiminde yapılması gerekmektedir.)
- Vücudun kendi kırmızı kan hücrelerini yok etmesini engellemek için bağışıklık sistemi engelleyici ilaçlar
- Kurşun zehirlenmeleri için kelasyon terapisi (Bu tedavi daha çok çocuklarda uygulanır. Çünkü çocuklarda oluşan demir eksikliği anemileri kurşun zehirlenmesi riskini artırır.)

2.7.3. Çeşitli tıbbi işlemler

Ağır olan anemi hastalıkları için çeşitli tıbbi işlemler uygulanabilir.

2.7.3.1. Kan nakli

Kan nakli sıkça uygulanan güvenli bir yöntemdir. Bu yöntem ile hastanın damarlarına kan grubuyla uyuşan kan nakli yapılmaktadır [41].

2.7.3.2. Kan ve kök hücre nakli

Kan ve kök hücre naklinin amacı kusurlu olan kök hücreyi sağlıklı bir donörden alınan kök hücreyle değiştirmektir. Kök hücreler kemik iliklerinde üretilir. Bu hücreler alyuvar, akyuvar ve trombositlere dönüşürler.

Kan nakline benzemektedir. Göğüs üzerindeki bir damara yerleştirilmiş tüp yardımıyla kök hücreler hastaya verilir. Bu hücreler kemik iliklerine ulaşarak yeni kan hücrelerini üretmeye başlar.

2.7.3.3. Ameliyat

Hayati tehlikesi olan veya ağır bir anemi hastalığına sahip hastalara ameliyat yapılır. Örnek olarak kolon kanserine ya da mide ülserine bağlı kontrol dışı kanaması olan kişilere ameliyat gerekir.

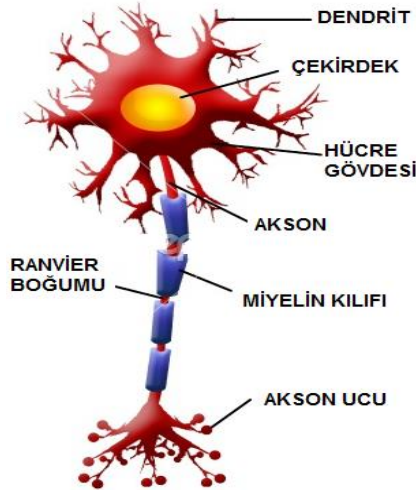
Eğer ki dalak yüksek oranda alyuvar yıkımı gerçekleştiriyorsa, dalağın ameliyat ile alınması gerekebilir. Dalak vücuttaki yıpranmış alyuvarları uzaklaştıran bir organdır. Bu organın genişlemesi veya hastalanması ile normalden daha fazla alyuvar yıkımına neden olabilir. Bu sonuca bağlı olarak anemi hastalığı ortaya çıkabilir [42].

BÖLÜM 3. ANEMİ HASTALIĞININ TESPİTİ İÇİN KULLANILACAK ALGORİTMALAR

Anemi hastalığının tespiti için ileri beslemeli yapay sinir ağı, destek vektör makinesi, k-ortalama algoritması, Naive Bayes algoritması ve regresyon ağaçları kullanılacaktır. Verilerin seçilmesi aşamasında Holdout yöntemi kullanılmıştır. Değerlendirme aşamasında ise ROC analizi kullanılmıştır.

3.1. Yapay Sinir Ağları ve İleri Beslemeli Yapay Sinir Ağı

Yapay sinir ağları, birbirine nöronlarla bağlı sistemler olarak tanımlanan, insan sinir sisteminden esinlenerek sayısal olarak modellenen yapılardır [43]. Elektronik ortam ile belirtmek istenen donanım ve yazılımdır. Bir yapay sinir ağı hem yazılımsal hem de donanımsal olarak oluşturulabilir [44].



Şekil 3.1. Bir sinir hücresinin yapısı [14]

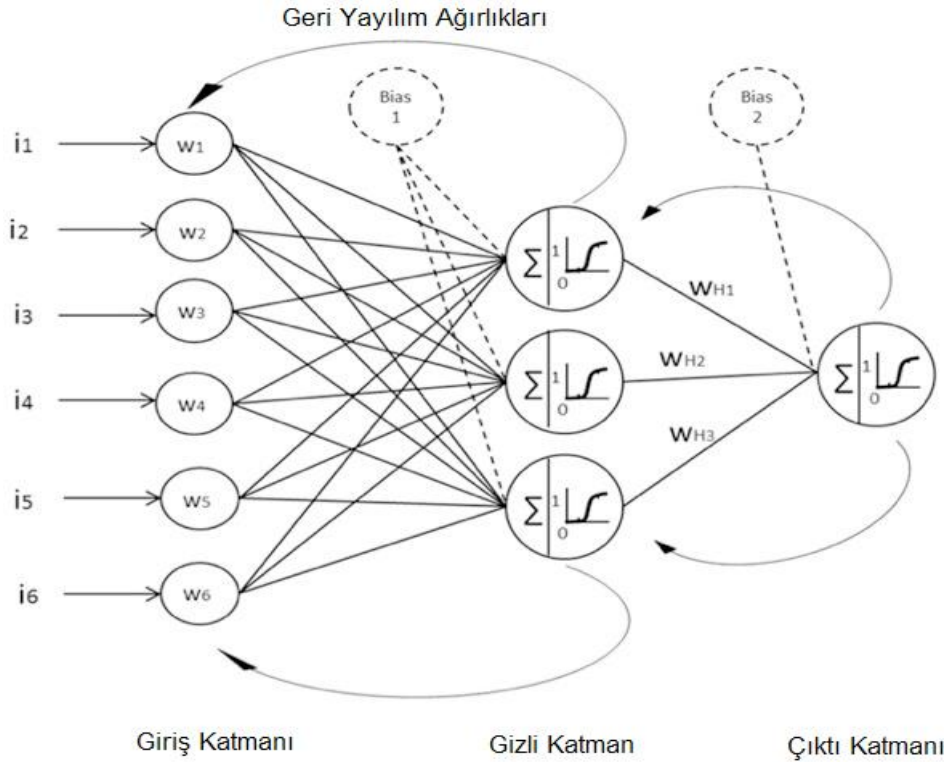
Yapay sinir ağlarının eğitim metotları öğreticili, öğreticisiz ve destekleyici öğrenme modelleri olarak üç ana başlık altında incelenebilir [45].

Her bir giriş değerine karşılık (vektör), bir çıkış değerine sahip öğrenme modeli öğreticili öğrenme modeli olarak ele alınır. Giriş değerlerine karşılık, çıkış değeri bulunmadan sezgisel olarak bunların sınıflandırılması olarak tanımlanan öğrenme modeline öğreticisiz öğrenme denir. Destekleyici öğrenme modeli ise, davranış psikolojisinden esinlenerek kendi içerisinde elde ettiği sonucu, test ve hata sonuçlarına bakarak iyi ya da kötü sonuç diye tanımlaması ile oluşan eğitim modelidir.

Çok katmanlı yapay sinir ağları çalışma adımları şunları içerir:

- Örneklerin toplanması
- Ağın topolojik yapısının belirlenmesi
- Öğrenme parametrelerinin belirlenmesi
- Ağırlıkların başlangıç değerlerinin verilmesi
- Öğrenme setinden örneklerin seçilmesi ve ağın eğitilmesi
- Öğrenme esnasında ileri hesaplama yapılması
- Çıktının, beklenen sonuç ile karşılaştırılması
- Ağırlıkların güncellenmesi

İleri katmanlı yapay sinir ağlarında, ağ ileri doğru hesaplama yapar. Giriş değerleri hiçbir işlem yapılmadan ara katmana belirli işlemler yapılarak gönderilir. Ara katmandan (gizli katman) alınan veriler çeşitli işlemler sonucunda çıkış katmanına yollanır. Bu katmanda oluşan çıktı değeri bizim ağımızın probleme ürettiği sonuçtur. Bu sonuçtaki hata oranının, beklenen sonuç ile arasındaki fark sonucu geri besleme gibi algoritmalarla bu hata minimize edilmeye çalışılır.



Şekil 3.2. İleri beslemeli geri yayımlı bir yapay sinir ağı modeli [46]

Yapay sinir ağlarında aktivasyon öğrenmenin ilk adımı olarak nitelendirilir. Eğer toplam sinyal hücreyi tetikleyebilecek, eşik değerinde ise yani 1 ise o hücre aktif, 0 ise pasiftir.

İleri beslemeli yapay sinir ağları çok katmanlı yapıda ağlardır. İleri beslemeli yapay sinir ağlarının işlem adımlarında iki aşaması bulunmaktadır.

- İleri hesaplama
- Geri hesaplama

İleri doğru hesaplama safhasında bir örnek alınarak girdi katmanından (i_1, i_2, \dots) ağı gösterilmesi ile başlar. Bu katmanda herhangi bir bilgi işleme olmaz. Bu nedenle çıktı ifadeleri (O), girdi ifadeleri (I) ile aynı değeri almaktadır. Girdi katmanının görevi gelen bilgiyi gizli katmana yollamaktır. Girdi katmanındaki k . elemanın çıktısı (3.1) nolu formülde görüldüğü gibi olur.

$$O_k^i = I_k \quad (3.1)$$

Gizli katmana gelen her bir giriş değeri ara katman tarafından bağlantı ağırlığına (W) göre alınır. Bir ara katmandaki proses elemanlarına gelen net girdi (NET) (3.2) nolu formüldeki gibi hesaplanır.

$$NET_j^a = \sum_{k=1}^n W_{kj} O_k^i \quad (3.2)$$

W_{kj} k. girdi elemanını j. ara katman değeriyle olan bağlantısının ağırlık değerini temsil etmektedir. j. ara katman elemanının çıktısı, bu girdinin aktivasyon fonksiyonundan geçirilmesiyle hesaplanır. Birden çok aktivasyon fonksiyonu çeşidi vardır. Genel olarak sigmoid fonksiyonu aktivasyon fonksiyonu olarak kullanılır. Sigmoid fonksiyonu kullanılırsa (3.3) nolu formüldeki gibi bir çıktı oluşur.

$$O_j^a = \frac{1}{1 + e^{-(NET_j^a + \beta_j^a)}} \quad (3.3)$$

Buradaki β_j ifadesi ara katmandaki j. elemana bağlanan eşik değerini temsil etmektedir. Genel olarak eşik değeri 1 olarak alınmaktadır. Sigmoid fonksiyonunun oryantasyonunu belirlemek üzere konulmuştur. Çıktı katmanına gelen değerlerinde NET girdilerinin toplanmasıyla çıktı değerleri oluşturulur. Bu adımdan itibaren geri doğru hesaplama adımları işletilir.

$$E_m = T_m - O_m \quad (3.4)$$

Geri doğru hesaplama safhasında O_m elemanına karşılık gelen T_m değeri arasındaki fark hata (E) oranıdır. Bu hata oranı (3.4) nolu formül ile hesaplanır. Bu hata oranı beklenen hata oranından yüksek ise oranın düşürülmesi için bu hata ağırlık değerlerine dağıtılır.

$$TE = \frac{1}{2} \sum_m E_m^2 \quad (3.5)$$

Ağın toplam hata oranı (TE) ise (3.5) nolu formül ile hesaplanır. Bu hata oranının azaltılması için iki durum söz konusudur.

- Ara katman ile çıktı katmanı arasındaki ağırlıkların güncellenmesi
- Ara katmanlar arası veya ara katman ile girdi katmanı arasındaki ağırlıkların güncellenmesi

$$\Delta W_{jm}^a(t) = \lambda \delta_m O_j^a + \alpha \Delta W_{jm}^a(t-1) \quad (3.6)$$

Ara katmandaki j. elemanın çıktı katmanındaki m. Elemanla bağlayan ağırlığın değişim miktarını ΔW^a ile ifade edecek olursak; herhangi bir t anında ağırlığın değişim miktarı (3.6) nolu denklemdeki gibidir. (3.6) nolu denklemdeki λ öğrenme katsayısını, α momentum katsayısını ifade etmek için kullanılır. Bir ağırlık değişim değerinin diğer bir değeri etkilemesini sağlamak için momentum katsayısı kullanılır. Bu sayede yerel optimum noktaya takılıp kalmadan ağırlıklar güncellenebilir. δ_m ise m. çıktı ünitesinin hatasını ifade etmek için kullanılır.

$$\delta_m = f'(NET).E_m \quad (3.7)$$

(3.7) nolu formülde δ_m değerinin nasıl hesaplandığı gösterilmektedir. Sigmoid fonksiyonunu kullandığımızda (3.7) nolu formülün açılımı (3.8) nolu formüldeki gibi olacaktır.

$$\delta_m = O_m(1-O_m).E_m \quad (3.8)$$

$$W_{jm}^a(t) = W_{jm}^a(t-1) + \Delta W_{jm}^a(t) \quad (3.9)$$

Ağırlıkların yeni değerlerinin hesaplanacağı fonksiyon ise (3.9) nolu formülde ifade edilmiştir.

$$\beta_m^c(t) = \beta_m^c(t-1) + \Delta\beta_m^c(t) \quad (3.10)$$

(3.10) nolu formülde ise eşik değerinin yeni değerinin hesaplanmasını sağlayan formül bulunmaktadır.

$$\Delta W_{kj}^i(t) = \alpha \Delta W_{kj}^i(t-1) + \lambda \delta_j^a O_k^i \quad (3.11)$$

$$\delta_j^a = O_j^a (1 - O_j^a) \sum_m \delta_m W_{jm}^a \quad (3.12)$$

Gizli ve ara katman arasındaki değişimler hesaplanırken sadece bir hata oranı dikkate alınırken, giriş ve ara katman arasında veya ara katmanlar arasındaki ağırlık güncellemeleri için çıktı katmanındaki her elemanın hatasından payını alması gerekir. Bu ağırlıkların değişimi yukarıdaki (3.11) nolu formülde gösterilmiştir. (3.11) nolu formüldeki δ^a ifadesini oluşturan formül ise (3.12) nolu formüldür. (3.12) nolu formülün içeriğinde $f'(NET)$ olarak sigmoid aktivasyon fonksiyonu kullanılmıştır.

$$W_{kj}^i(t) = W_{kj}^i(t-1) + \Delta W_{kj}^i(t) \quad (3.13)$$

Ağırlıkların son şekli (3.13) nolu formüldeki gibi hesaplanmaktadır.

$$\beta_j^a(t) = \beta_j^a(t-1) + \Delta\beta_j^a(t) \quad (3.14)$$

Eşik değerlerindeki değişim ise (3.14) nolu formüldeki gibi hesaplanır. Bu işlemler eğitim boyunca devam eder. Eğitim tamamlandıktan sonra test aşamasında ağırlık güncellemesi yapılmaz. Bu aşamada yapılacak olan, ilk aşamalardaki adımların aynı olup, giriş verilerine karşılık çıkan sonuç son kısımda en yakın değere yuvarlanır. Bu sonuç ile verinin hangi sınıfa ait olduğu tespit edilir.

3.2. K-Ortalama Algoritması

K-ortalama algoritması öğreticisiz bir algoritmadır [47]. Bu nedenle eğitim aşaması yoktur. Algoritmanın temel işleyişinde n nesneden oluşan bir veri kümesini (D) k ($k \leq n$) demete ayırmak yatmaktadır. Bilinen bir k değeri için k-ortalama algoritması aşağıdaki gibidir:

- Rastgele k tane nokta seç (Bu noktalar oluşturulacak öbeklerin merkezi varsayılır)
- Tüm noktaların k merkezlerine olan uzaklıklarını hesapla
- Her noktayı kendisine yakın olan merkezin öbeğine ata
- Oluşturulan öbeklerin merkezini hesapla
- 2.Adımdan itibaren dikkate değer bir değişiklik gözlenmeyip kümeler değişmedikçe iterasyonlara devam et

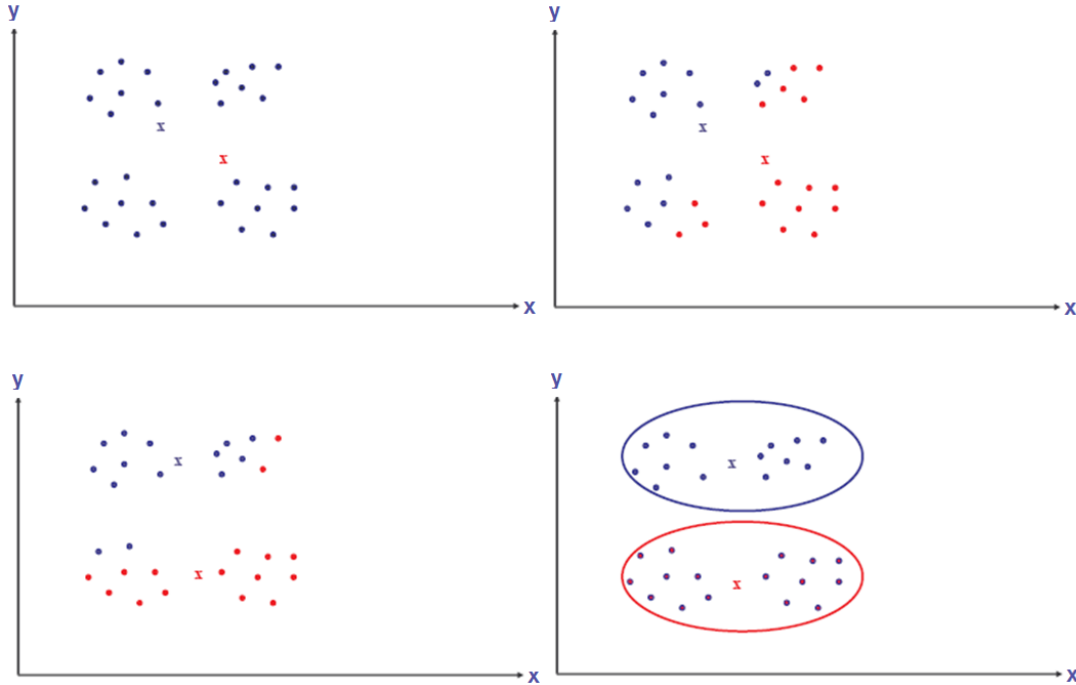
K-ortalama algoritmasının değerlendirmesinde genellikle karesel hata kriteri (SSE) kullanılmaktadır. Bulunan en düşük SSE değeri en iyi sonucu verir [48]. (3.15) nolu formül SSE hesaplanmasında kullanılacak formüldür. Formülde kullanılan C_i kümeyi temsil ederken, bu kümeye ait merkezi c_i temsil etmektedir. x ise kümedeki herhangi bir elemanı göstermektedir. C kümesinin merkezi ise (3.16) nolu formül ile hesaplanmaktadır. Tüm kümelerdeki elemanların tamamı için elemanın küme merkezine uzaklığı karesel olarak hesaplanmakta ve bu uzaklıklar toplanmaktadır.

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} Uzaklık(c_i, x)^2 \quad (3.15)$$

$$c_i = \frac{1}{m_i} \sum_{x \in C_i} x \quad (3.16)$$

K-ortalama algoritması benzeri algoritmaların en önemli parametresi küme sayısını temsil eden k parametresidir. K kullanıcı tanımlı olarak, sonuç için en iyi olacak şekilde belirlenmesi zor bir parametredir. Bu nedenle k farklı değerler verilerek

çalıştırılır ve oluşan sonuçlar doğruluk analizleri ile en iyi sonucu veren k değerinin seçilmesi sağlanır.

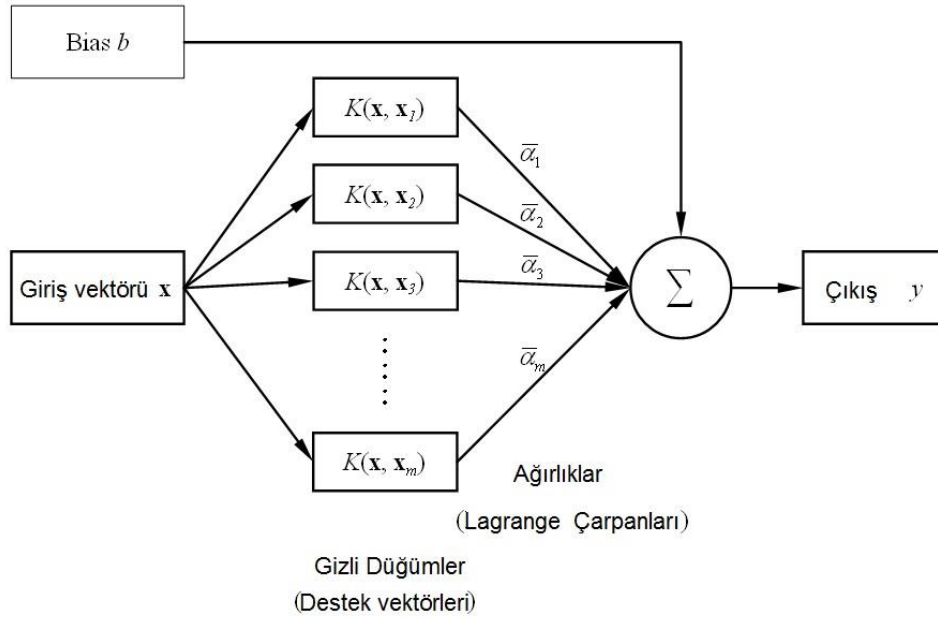


Şekil 3.3. İki kümeli bir k -ortalama algoritmasının örnek çalışma şeması

3.3. Destek Vektör Makinesi (DVM)

Destek vektör makinesi, orjinal adıyla support vector machine (SVM) yöntemi, lineer ve lineer olmayan problemleri çözebilen, giriş verilerini analiz ederek örüntülerini belirleyen yaygın bir algoritmadır. DVM örüntü tanıma problemlerinde, değişkenler arası bağları tespit eden ve bunları sınıflandıran bir algoritmadır. Tek sınıflı DVM, girdi veri kümesini içerisindeki belirli sistemlere göre bu girdilere karşılık gelen çıktıları iki sınıfa ait olma şartı ile sınıflandırmaktadır [15] [49].

Bu çalışmada tek sınıflı bir DVM ye ihtiyaç duyulduğundan dolayı, özellikle bu yöntem üzerinde durulacaktır.



Şekil 3.4. DVM yapısı [50]

Tek sınıf DVM, tek sınıfa ihtiyaç duyan problemler için özelleştirilmiş bir DVM algoritmasıdır.

$$\{x_1, \dots, x_l\} \subseteq X, \quad (3.17)$$

$$y_i \in \{1, -1\} \quad (3.18)$$

(3.17) nolu formüldeki $l \in \mathbb{N}$ gözlem sayısı olmak üzere, X bir \mathbb{R}^N alt kümesi, (3.18) örnek bir hedef vektörü, $\Phi: X \rightarrow F$ öznitelik haritamız olsun. Bu haritamız F uzayındaki Φ basit bir kernel fonksiyonu yardımıyla kolayca hesaplanabilir.

$$k(x, z) = (\Phi(x) \cdot \Phi(z)) \quad (3.19)$$

Bu algoritmanın amacı (3.17) nolu formülle ifade edilen bir örnek uzayı içerisindeki örneklerin büyük bir kısmını içeren dar bir alanda +1 değeri alan ve örnek uzay içerisinde kalan diğer örnekler için ise -1 değeri alacak bir f fonksiyonu tanımlamaktır. Bir kernel fonksiyonu yardımıyla veriler daha yüksek boyutlu öznitelik uzayına

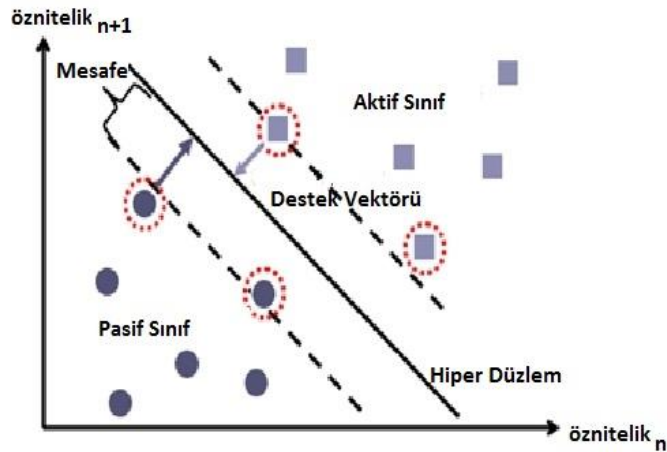
taşınarak, bu verileri merkez noktasından en iyi şekilde iki ayrı sınıfa ayırabilecek bir hiperdüzlem çıkartılır. Bu ifadelere göre (3.20) ve (3.21) nolu formül oluşur.

$$\min_{w \in F, \xi \in R^l, \rho \in R} \frac{1}{2} \|w\|^2 + \frac{1}{\nu l} \sum_i \xi_i - \rho \quad (3.20)$$

$$(w \cdot \Phi(x_i)) \geq \rho - \xi_i, \xi_i \geq 0 \quad (3.21)$$

$\nu \in (0, 1)$ ile ifade edilen ν cezalandırılmış terim, ξ_i ler ise aylak değişkenler olarak alınmıştır. w ve ρ değerlerini optimizasyon problemini çözerek elde ettiğimizde karar fonksiyonumuz (3.22) nolu formüldeki gibi olacaktır:

$$f(x) = \text{sgn}((w \cdot \Phi(x)) - \rho) \quad (3.22)$$



Şekil 3.5. Destek vektör şeması [51]

(3.21) nolu formülün çoğunluğu +1 e eşit olmakla birlikte, kalan kısımdakiler -1 e eşit olacaktır. $\alpha_i, \beta_i \geq 0$ olmak üzere lagrange çarpanları ile (3.23) deki formül elde edilir.

$$L(w, \xi, \rho, \alpha, \beta) = \frac{1}{2} \|w\|^2 + \frac{1}{\nu l} \sum_i \xi_i - \sum_i \beta_i \xi_i - \rho - \sum_i \alpha_i ((w \cdot \Phi(x_i)) - \rho + \xi_i) \quad (3.23)$$

(3.23) nolu fomüldeki w, ζ, ρ değişkenlerine göre kısmı türevlerini alarak 0 a eşitlersek (3.24) ve (3.25) nolu sonuçlar elde edilir.

$$w = \sum_i \alpha_i \Phi(x_i) \quad (3.24)$$

$$\alpha_i = \frac{1}{vl} - \beta_i \leq \frac{1}{vl}, \sum_i \alpha_i = 1 \quad (3.25)$$

Bu işlemlerden sonra (3.22) deki karar fonksiyonumuz (3.26) nolu formüldeki gibi ifade edilir. Aynı zamanda (3.24) nolu ve (3.25) nolu formülleri yerlerine yazarsak (3.27) nolu formül elde edilir.

$$f(x) = \text{sgn}\left(\sum_i \alpha_i k(x_i, x) - \rho\right) \quad (3.26)$$

$$\min_{\alpha} \frac{1}{2} \sum_{ij} \alpha_i \alpha_j k(x_i, x_j) \quad (3.27)$$

$$0 \geq \alpha_i \geq \frac{1}{vl}, \sum_i \alpha_i = 1 \quad (3.28)$$

$$\rho = (w \cdot \Phi(x_i)) = \sum_j \alpha_j k(x_i, x_j) \quad (3.29)$$

α_i değeri ise (3.27) ve (3.28) nolu formüller ile bulunabilir. ρ değişkeni ise (3.29) nolu formülde de görüldüğü üzere herhangi bir i . α_i değeri ve ona karşılık gelen x_i değeri ile elde edilebilir [52].

3.4. Naive Bayes Algoritması

Naive Bayes sınıflandırıcısı, Bayes teoreminin bağımsızlık önermesiyle başitleştirilmiş şeklidir [53]. Bayes teorisi (3.30) nolu formül ile ifade edilir.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (3.30)$$

$P(A|B)$; B olayı gerçekleştiği durumda A olayının meydana gelme olasılığıdır. $P(B|A)$; A olayı gerçekleştiği durumda B olayının meydana gelme olasılığıdır. $P(A)$ ve $P(B)$; A ve B olaylarının önsel olasılıklarıdır. Önsel olasılık Bayes teoremine öznellik kazandırır. Diğer bir ifadeyle $P(A)$ olayı hakkında henüz elde veri toplanmadan bu olay hakkındaki sahip olunan bilgidir. $P(B|A)$ olayı da ardıl olasılıktır. Ardıl olasılıkla anlatılmak istenen A olayının gerçekleşmiş olduğu durumlarda B olayının gerçekleşme ihtimali hakkında bilgi vermesidir [54].

$$x(i), i = 1, 2, \dots, L \quad (3.31)$$

$$x(i) = [x(1), x(2), \dots, x(L)]^T \in R^L \quad (3.32)$$

Naive Bayes algoritması için nicelik kümesi (3.31) nolu formülde görüldüğü gibi kabul edilirse (3.32) nolu formül L boyutlu nicelik vektörünü oluşturur. $X \in R^L$ verildiğine göre ve S ayrıştırılacak sınıflar kümesi olarak kabul edilirse, Bayes teoremine göre (3.33) nolu ve (3.34) nolu formüller yazılır.

$$P(S_i | x) \times p(x) = p(x | S_i) \times P(S_i) \quad (3.33)$$

$$P(x) = \sum_{i=1}^L p(x | S_i) P(S_i) \quad (3.34)$$

(3.33) nolu ve (3.34) nolu formüllerdeki $P(S_i)$; S_i nin öncel olasılığı ($i = 1, 2, \dots, L$), $P(S_i|x)$; S_i nin ardıl olasılığı, $p(x)$; x 'in olasılık yoğunluk fonksiyonu, $p(x|S_i)$ ($i = 1, 2, \dots, L$) x 'in koşullu olasılık yoğunluk fonksiyonudur.

Sonuç olarak (3.32) nolu formüldeki gibi verilecek bir nicelik vektörünün hangi sınıfa ait olup olmadığını anlamak için Bayes karar teorisine göre (3.35) nolu formüldeki koşul doğru ise S_i ye aittir. Diğer bir ifadeyle (3.36) nolu formüldeki koşul sağlanıyorsa S_i ye aittir.

$$P(S_i | x) > P(S_j | x), \forall j \neq i \quad (3.35)$$

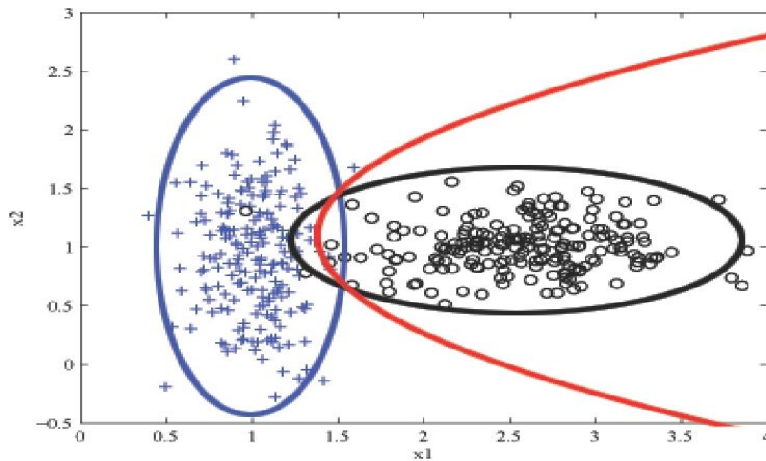
$$P(x | S_i)P(S_i) > P(x | S_j)P(S_j), \forall j \neq i \quad (3.36)$$

Verilen x sınıfının S_i 'ye ait olup olmadığına karar vermek için yukarıda (3.36) ve (3.37) ile ifade ettiğimiz Bayes karar teoreminde istatistik olarak bağımsızlık önermesi kullanılırsa bu sınıflandırmaya Naive Bayes sınıflandırması denir. (3.36) nolu formüldeki $P(x|S_i)$ ifadesini (3.37) deki formüle göre düzenlersek sonuç olarak (3.38) nolu formül oluşacaktır.

$$P(x | S_i) \approx \prod_{k=1}^L P(x_k | S_i) \quad (3.37)$$

$$P(S_i) \prod_{k=1}^L P(x_k | S_i) > P(S_j) \prod_{k=1}^L P(x_k | S_j) \quad (3.38)$$

(3.38) nolu formüldeki $P(S_i)$ ve $P(S_j)$ i ve j sınıflarının öncel olasılıklarıdır. Elde edilen veri kümesinden değerler kolaylıkla bulunabilir. Bu sınıflayıcının kullanım alanı kısıtlı gözükse de yüksek boyutlu uzayda, yeterli sayıdaki veriyle nicelik kümesinin (x) bileşenlerinin istatistiksel olarak bağımsız olma koşulu esnetilerek başarılı sonuçlar elde edilebilir [39] [55].



Şekil 3.6. x_1 ve x_2 adında 2 sınıfa sahip Naive Bayes modeli. Kırmızı çizgi karar sınırını, diğer çizgilerde sınıfların kendi yoğunluklarını gösteren gauss modelleridir [56]

3.5. Regresyon Ağaçları

Regresyon ağaçları, sınıflandırma ağaçlarından farklı yapılardır. Regresyon ağaçlarında sınıflar yoktur bu nedenle de sınıf ayırımı için Gini indeksi kullanılamaz. Regresyon ağaçlarındaki ayırımlar artıkların karelerini azaltma algoritmasıyla gerçekleşir. Bu algoritmanın amacı, iki düğüm için yapılan hesaplamalarda elde edilen toplam varyansın en küçük değerde olmasıdır. Regresyon ağacı yönteminde her düğümdeki minimizasyon problemi (3.39) nolu formüldeki gibi çözümlür.

$$\arg \min_{x_j \leq x_j^R, j=1, \dots, M} [P_l \text{Var}(Y_l) + P_r \text{Var}(Y_r)] \quad (3.39)$$

(3.39) nolu formüldeki P_l ve P_r sırası ile sol ve sağ düğümlerin olasılıklarıdır. M eğitim setindeki değişkenlerin sayısını temsil ederken, değişken j “ x_j ” olarak ifade edilmekte ve x_j^R ise değişken x_j nin en iyi ayırım değerini ifade etmektedir. $\text{Var}(y_l)$ ve $\text{Var}(y_r)$ ise karşılıklı sağ ve sol alt düğümler için sorumlu vektörleri ifade etmektedir. $x_j \leq x_j^R, j=1, \dots, M$ ifadesi ile ifade edilen optimal ayırım sorgulamasıdır.

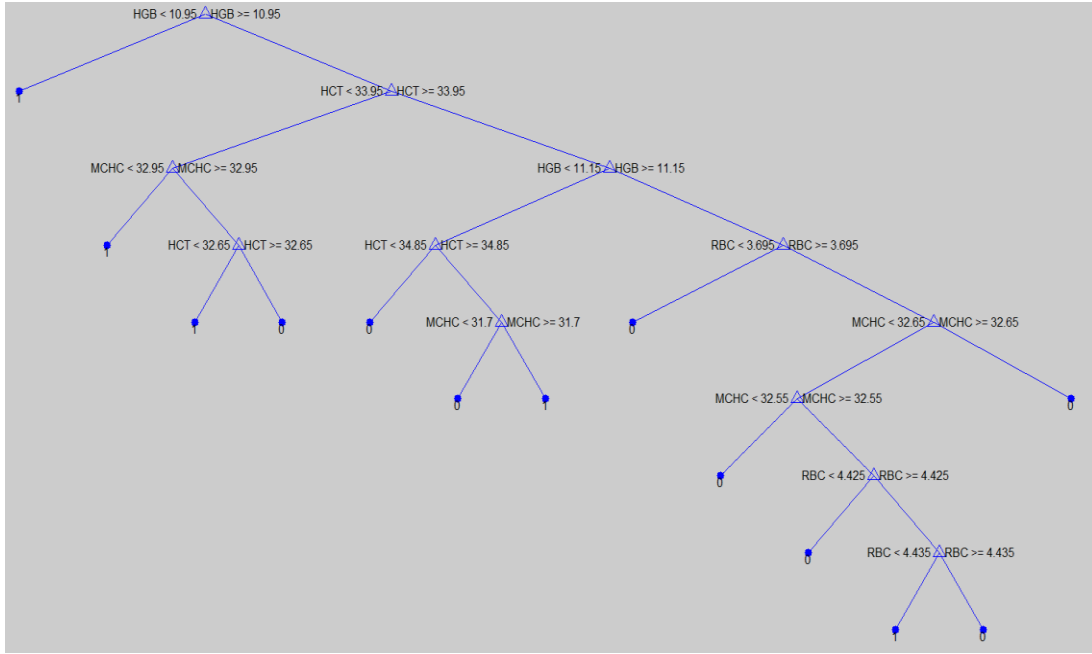
Artıkların karelerini azaltma yöntemi ile sınıf k 'nın nesnelere değer “1”, diğer sınıfların nesnelere de değer “0” ataması yapılırsa bu değerlerin örnek varyansı (3.40) nolu formüle eşit olur.

$$p(k | t)[1 - p(k | t)] \quad (3.40)$$

Katışıklık ölçümü $i(t)$ ise (3.41) nolu formül ile elde edilir.

$$i(t) = 1 - \sum_{k=1}^K p^2(k | t) \quad (3.41)$$

(3.41) nolu formüldeki $p(k|t)$ düğüm t içinde sınıf k 'nın koşullara bağlı özelliklerini, K sınıf sayısını, k sınıf indeksini ve t düğüm indeksini ifade etmektedir [20].



Şekil 3.7. Çalışmadan alınmış örnek bir regresyon ağacı şeması

3.6. Holdout Yöntemi

Bu algoritmanın amacı verilerin rastgele bir şekilde seçilerek daha objektif sonuçlar elde etmektir. Bu model ile veri seti, eğitim seti ve test seti olmak üzere iki ana sınıfa ayrılır [57][58].

3.7. ROC Analizi

ROC analizi, çeşitli istatistiksel yöntemler ile elde edilen sonuçlara dayalı olarak bunları doğruluk, duyarlılık ve özgülük olarak değerlendirilebilmesini sağlayan bir yöntemdir. Son yıllarda makine öğrenmesi ve veri madenciliği yöntemlerinde çok fazla kullanılmaktadır. ROC eğrisi grafiği ise doğru pozitiflerin (true positive), yanlış pozitiflere (false positive) olan oranını ifade edebilir [59].

Bu yöntem içerisinde true positive (TP), false positive (FP), true negative (TN), false negative (FN) gibi kavramlar bulunmaktadır. Bu ifadeler Tablo 3.1. üzerinde açıklanmıştır.

Tablo 3.1. ROC parametreleri ve açıklamaları

	Açıklama
TP: True Positive	Doktorun kesin test sonuçlarına göre hastalıklı olarak tespit ettiği kişiyi yapılan test sonuçlarında hastalıklı olarak tespit etmesidir
TN: True Negative	Doktorun kesin test sonuçlarına göre sağlıklı olarak tespit ettiği kişiyi yapılan test sonuçlarında sağlıklı olarak tespit etmesidir
FP: False Positive	Doktorun kesin test sonuçlarına göre sağlıklı olarak tespit ettiği kişiyi yapılan test sonuçlarının yanlış olarak yani hastalıklı olarak tespit etmesidir
FN: False Negative	Doktorun kesin test sonuçlarına göre hastalıklı olarak tespit ettiği kişiyi yapılan test sonuçlarının yanlış olarak yani sağlıklı olarak tespit etmesidir

$$\text{Doğruluk} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3.42)$$

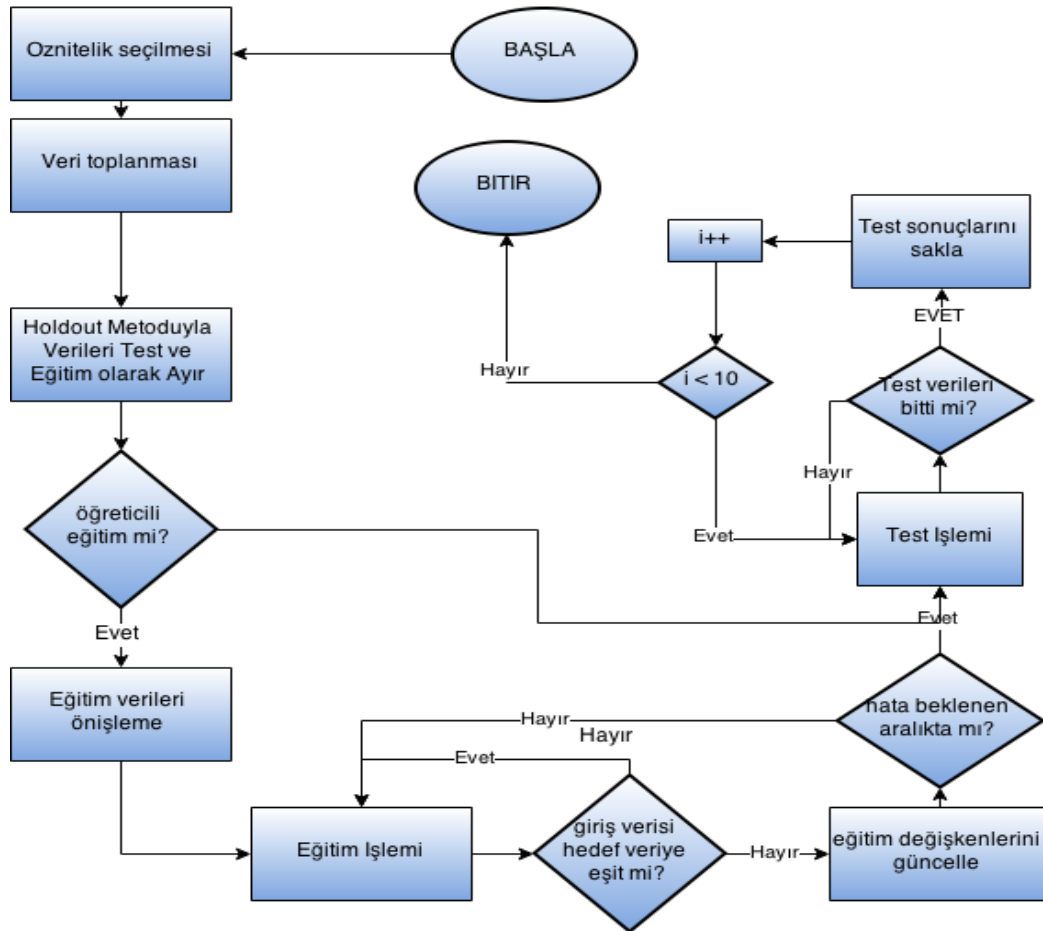
$$\text{Hassasiyet} = \frac{(TP)}{(TP + FN)} \quad (3.43)$$

$$\text{Özgüllük} = \frac{(TN)}{(FP + TN)} \quad (3.44)$$

ROC değerlerimiz içerisindeki doğruluk (3.42) nolu formül ile, duyarlılık (3.43) nolu formül ile, özgüllük değeri ise (3.44) nolu formül ile hesaplanmaktadır. Bu hesaplamalar sonucunda elde edilen değerler ile yöntemlerimiz analiz edilmiştir. Tüm test sonuçları için bu değerler elde edilmiştir.

BÖLÜM 4. İSTATİSTİKSEL YÖNTEMLER YARDIMI İLE ANEMİ HASTALIĞININ TESPİTİ VE BU YÖNTEMLERİN ANALİZLERİ

Tezin bu bölümünde problemi çözmeye yönelik Yapılan işlemler ve kullandığımız yöntemlerin detayları yer almaktadır. İşlem basamakları Şekil 4.1 de gösterilmektedir. Test aşaması 10 kez tekrarlanarak Holdout yöntemiyle veriler rastgele olarak alınmıştır. Bu işlem ile test aşamasının güvenilirliğinin artırılması amaçlanmıştır.



Şekil 4.1. Çalışmada kullandığımız algoritma şeması

4.1. Önişleme

Verilerimizin yapısı örnek olarak Tablo 4.1. de gösterilmiştir. “1” demir eksikliği anemisi olan kişileri, “0” ise sağlıklı kişileri temsil etmektedir.

Tablo 4.1. Veri setinden alınmış örnek veriler

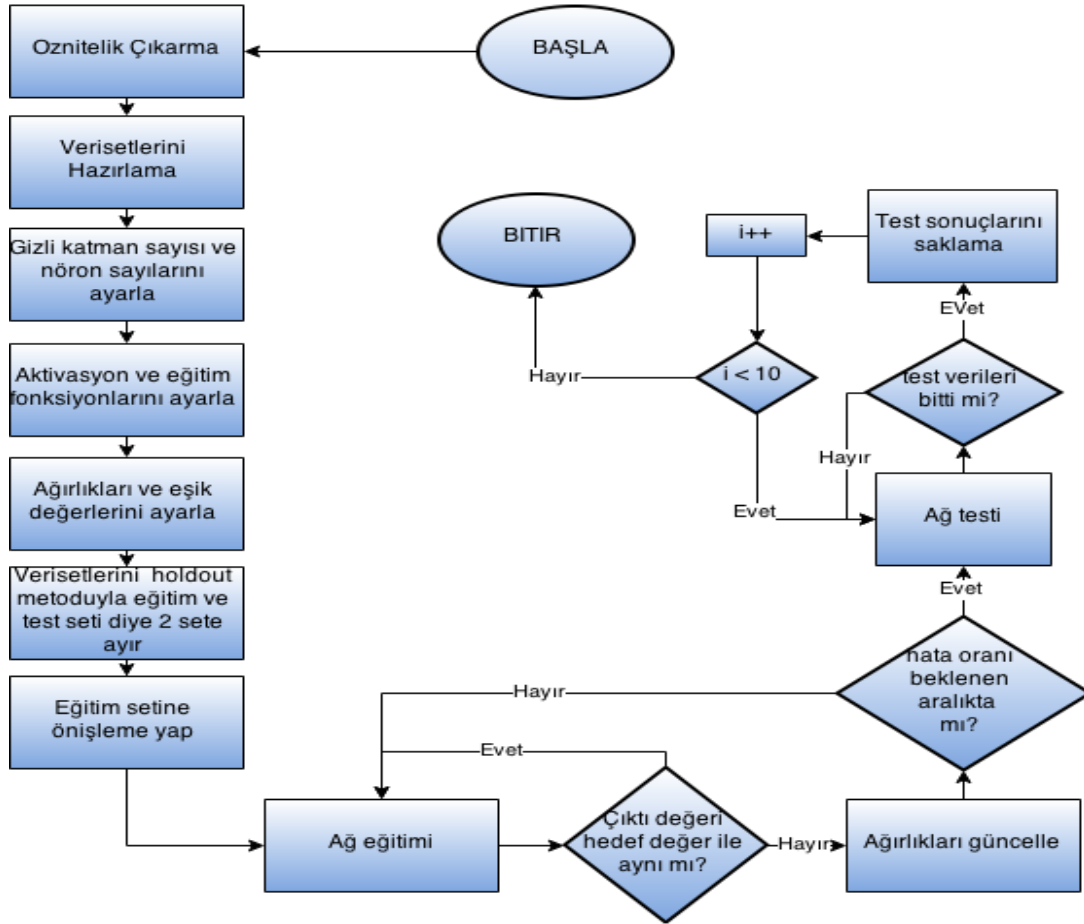
RBC	HGB	HCT	MCV	MCH	MCHC	SONUÇ
3,66	6,42	23,62	64,52	17,52	27,12	1
3,9	7,02	24,82	63,62	17,92	28,22	1
4,65	7,32	28,3	60,92	15,72	25,82	1
4,4	7,42	26,6	60,52	16,82	27,82	1
4,49	7,52	27,9	62,12	16,72	26,92	1
4,63	7,62	28,3	61,12	16,42	26,92	1
4,62	7,72	27,5	59,52	16,72	28,02	1
3,95	11,6	35,2	89,1	29,4	33	0
4,54	11,6	36,5	80,42	25,62	31,8	0
4,23	11,6	34,7	82,02	27,4	33,4	0
4,7	11,6	36,5	77,7	24,7	31,8	0
4,5	11,6	35,1	78,5	26	33	0
4,2	11,6	34,3	80,9	27,4	33,8	0
3,9	11,6	35,5	91	29,7	32,7	0

$$y = \frac{(y_{\max} - y_{\min}) * (x - x_{\min})}{(x_{\max} - x_{\min})} + y_{\min} \quad (4.1)$$

Veri seti (4.1) nolu formül ile önişlemeye tabi tutularak sınıflandırma işlemleri yapılmıştır. (4.1) nolu formül ile değerlerimiz -1 ve 1 aralığına çekilmektedir. Bu işlemimizde önişlemeyi tersine çevirerek orijinal verilerimizi alabilmek için y_{\max} , y_{\min} , x_{\min} , x_{\max} , aralık ve bazı istatistikler saklanmaktadır.

4.2. İleri Beslemeli Yapay Sinir Ağı Testleri

Veri seti içerisinde 1441 veri ile eğitilen ileri beslemeli geri yayımlı yapay sinir ağı, 359 veri ile test edilmiştir.



Şekil 4.2. Çalışmada kullandığımız ileri beslemeli geri yayımlı yapay sinir ağı algoritma şeması

Kullanılan ağı algoritma akış diyagramı Şekil 4.2. de verilmiştir. Kullanılan ileri beslemeli yapay sinir ağında, ara katmanda (gizli katman) 13 nöron, giriş katmanında veri giriş sayımızla aynı olarak 6 nöron, çıkış katmanında ise bir nöron bulunmaktadır. Ağı öğrenme modeli momentum kullanan gradyan azaltımı öğrenme yöntemidir. 6000 iterasyon ile network eğitimi yapılmıştır. Eğitim oranı katsayısı 0.4, momentum katsayısı 0.8 olarak seçilmiştir.

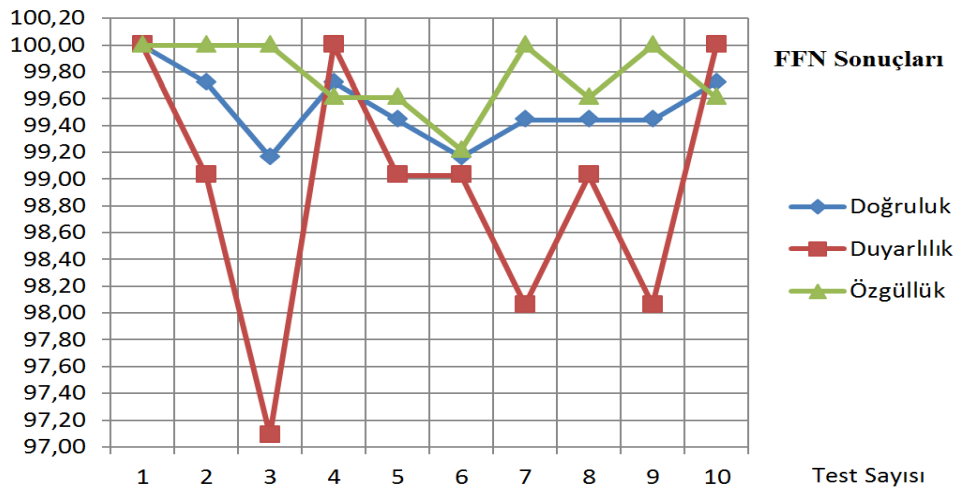
Tablo 4.2. İleri beslemeli yapay sinir ağı tüm test sonuçları

Test Sayısı	TP	TN	FP	FN
1	103	256	0	0
2	102	256	0	1
3	100	256	0	3
4	103	255	1	0
5	102	255	1	1
6	102	254	2	1
7	101	256	0	2
8	102	255	1	1
9	101	256	0	2
10	103	255	1	0

Yapılan testin sayısal sonuçları Tablo 4.2. de gösterilmektedir.

Tablo 4.3. İleri beslemeli yapay sinir ağı ROC analizi tablosu

Test Sayısı	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
1	100,00	100,00	100,00
2	99,72	99,03	100,00
3	99,16	97,09	100,00
4	99,72	100,00	99,61
5	99,44	99,03	99,61
6	99,16	99,03	99,22
7	99,44	98,06	100,00
8	99,44	99,03	99,61
9	99,44	98,06	100,00
10	99,72	100,00	99,61



Şekil 4.3. İleri beslemeli yapay sinir ağı test sonuçlarının grafiği

Yapılan testlerin sayısal ROC analizi sonuçları Tablo 4.3. de gösterilmektedir.

Testlerin duyarlılık, özgüllük ve doğruluk oranlarının grafiği Şekil 4.3. de gösterilmektedir. Testlerin sonucunda elde edilen en iyi sonuç %100 doğruluk, %100 duyarlılık, %100 özgüllük değerleri olmuştur.

Tablo 4.4. İleri beslemeli yapay sinir ağı testlerinin ortalama değerleri

Formül	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
Değer / 10	99,53	98,93	99,77

10 testin tamamının ortalama değerleri Tablo 4.4. de gösterilmektedir.

4.3. Destek Vektör Makinesi Testleri

Veri seti içerisinde 1441 veri ile eğitilen destek vektör makinesi, 359 veri ile test edilmiştir.

Kullanılan DVM de, maksimum iterasyon sayısı 20000, kernel fonksiyonu rbf, rbf sigma değeri 1, metot olarak kuadratik programlama kullanılmıştır.

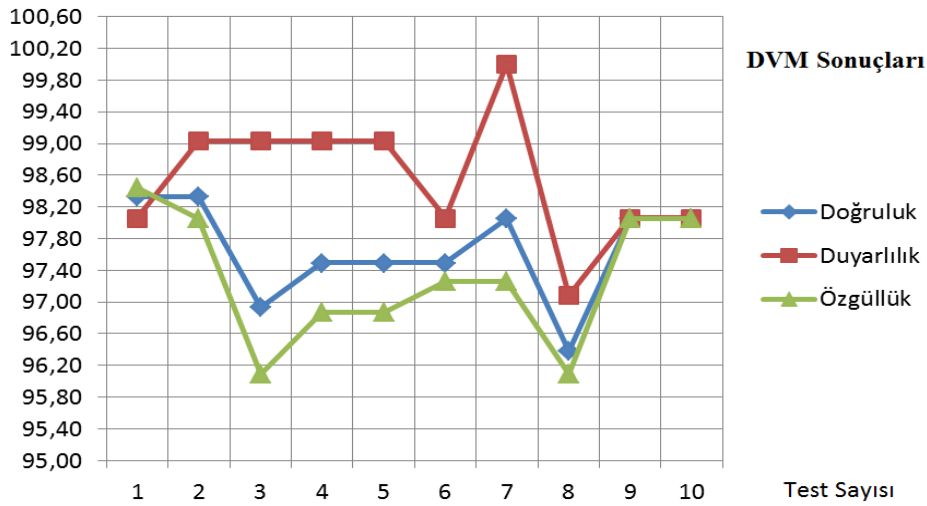
Tablo 4.5. DVM tüm test sonuçları

Test Sayısı	TP	TN	FP	FN
1	101	252	4	2
2	102	251	5	1
3	102	246	10	1
4	102	248	8	1
5	102	248	8	1
6	101	249	7	2
7	103	249	7	0
8	100	246	10	3
9	101	251	5	2
10	101	251	5	2

Yapılan testin sayısal sonuçları Tablo 4.5. de gösterilmektedir.

Tablo 4.6. DVM ROC analizi tablosu

Test Sayısı	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
1	98,33	98,06	98,44
2	98,33	99,03	98,05
3	96,94	99,03	96,09
4	97,49	99,03	96,88
5	97,49	99,03	96,88
6	97,49	98,06	97,27
7	98,05	100,00	97,27
8	96,38	97,09	96,09
9	98,05	98,06	98,05
10	98,05	98,06	98,05



Şekil 4.4. DVM test sonuçlarının grafiği

Yapılan testlerin sayısal ROC analizi sonuçları Tablo 4.6. da gösterilmektedir. Testlerin duyarlılık, özgüllük ve doğruluk oranlarının grafiği Şekil 4.4. de gösterilmektedir.

Tablo 4.7. DVM testlerinin ortalama değerleri

Formül	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
Değer / 10	97,66	98,54	97,30

10 testin tamamının ortalama değerleri Tablo 4.7. de gösterilmektedir.

4.4. Regresyon Ağacı Testleri

Veri seti içerisinde 1441 veri ile eğitilen ileri regresyon ağacı, 359 veri ile test edilmiştir.

Tablo 4.8. Regresyon ağacı tüm test sonuçları

Test Sayısı	TP	TN	FP	FN
1	101	254	2	2
2	103	254	2	0
3	100	254	2	3
4	99	256	0	4
5	101	254	2	2

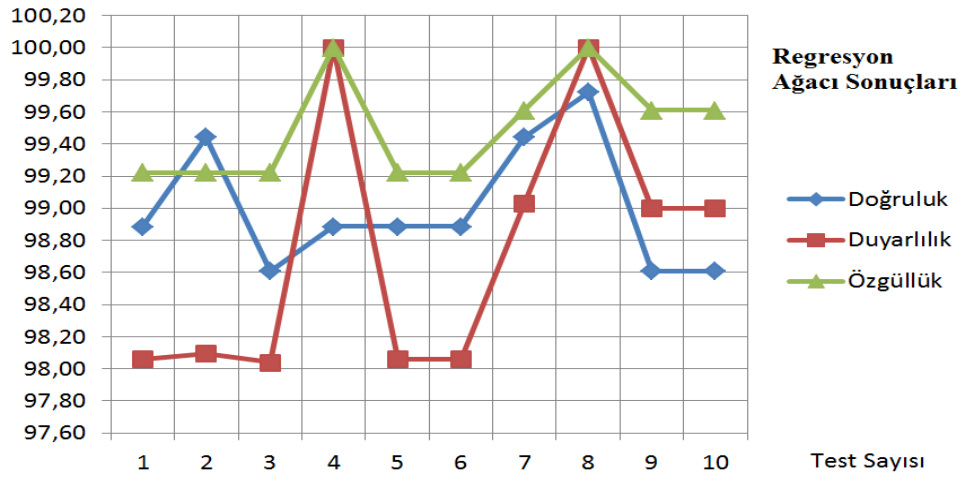
Tablo 4.8. Regresyon ağacı tüm test sonuçları (devamı)

6	101	254	2	2
7	102	255	1	1
8	102	256	0	1
9	99	255	1	4
10	99	255	1	4

Yapılan testin sayısal sonuçları Tablo 4.8. de gösterilmektedir.

Tablo 4.9. Regresyon ağacı ROC analizi tablosu

Test Sayısı	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
1	98,89	98,06	99,22
2	99,44	98,10	99,22
3	98,61	98,04	99,22
4	98,89	100,00	100,00
5	98,89	98,06	99,22
6	98,89	98,06	99,22
7	99,44	99,03	99,61
8	99,72	100,00	100,00
9	98,61	99,00	99,61
10	98,61	99,00	99,61



Şekil 4.5. Regresyon ağacı test sonuçlarının grafiği

Yapılan testlerin sayısal ROC analizi sonuçları Tablo 4.9. da gösterilmektedir.

Testlerin duyarlılık, özgüllük ve doğruluk oranlarının grafiği Şekil 4.5. de gösterilmektedir.

Tablo 4.10. Regresyon ağacı testlerinin ortalama değerleri

Formül	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
Değer / 10	99,00	98,73	99,49

10 testin tamamının ortalama değerleri Tablo 4.10. da gösterilmektedir.

4.5. Naive Bayes Algoritması Testleri

Veri seti içerisinde 1441 veri ile eğitilen Naive Bayes sistemi, 359 veri ile test edilmiştir.

Tablo 4.11. Naive Bayes algoritması tüm test sonuçları

Test Sayısı	TP	TN	FP	FN
1	101	247	9	2
2	100	245	11	3
3	100	244	12	3
4	101	250	6	2
5	102	247	9	1
6	100	249	7	3
7	100	243	13	3
8	101	248	8	2
9	101	245	11	2
10	101	240	16	2

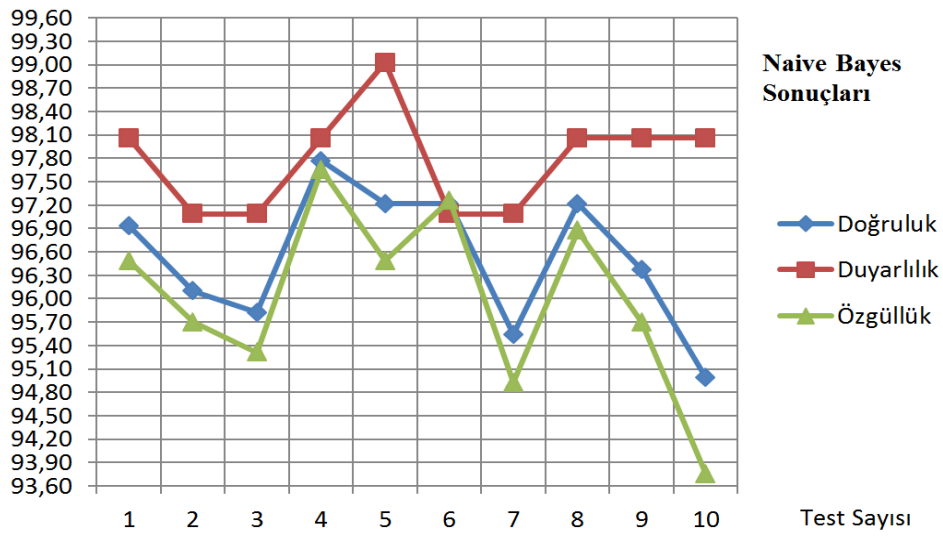
Yapılan testin sayısal sonuçları Tablo 4.11. de gösterilmektedir.

Tablo 4.12. Naive Bayes algoritması ROC analizi tablosu

Test Sayısı	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
1	96,94	98,06	96,48
2	96,10	97,09	95,70
3	95,82	97,09	95,31
4	97,77	98,06	97,66

Tablo 4.12. Naive Bayes algoritması ROC analizi tablosu (devamı)

5	97,21	99,03	96,48
6	97,21	97,09	97,27
7	95,54	97,09	94,92
8	97,21	98,06	96,88
9	96,38	98,06	95,70
10	94,99	98,06	93,75



Şekil 4.6. Naive Bayes algoritması test sonuçlarının grafiği

Yapılan testlerin sayısal ROC analizi sonuçları Tablo 4.12. de gösterilmektedir. Testlerin duyarlılık, özgüllük ve doğruluk oranlarının grafiği Şekil 4.6. da gösterilmektedir.

Tablo 4.13. Naive Bayes algoritması testlerinin ortalama değerleri

Formül	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
Değer / 10	96,52	97,77	96,02

10 testin tamamının ortalama değerleri Tablo 4.13. de gösterilmektedir.

4.6. K-ortalama Algoritması Testleri

Veri seti içerisinde 359 veri ile test edilmiştir.

Kullanılan k-ortalama algoritmasının da, uzaklık ölçüm metodu olarak manhattan mesafesi ölçüm tekniği kullanılmıştır.

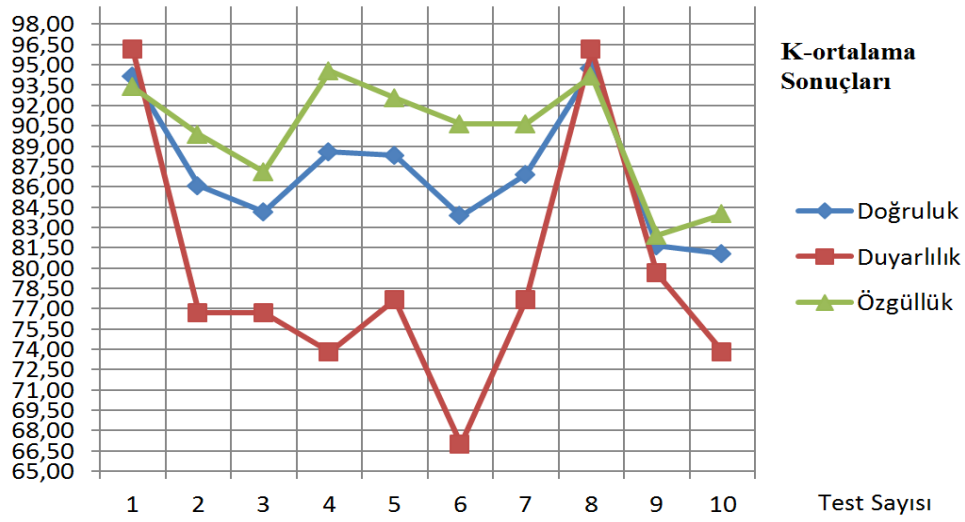
Tablo 4.14. K-ortalama algoritması tüm test sonuçları

Test Sayısı	TP	TN	FP	FN
1	99	239	17	4
2	79	230	26	24
3	79	223	33	24
4	76	242	14	27
5	80	237	19	23
6	69	232	24	34
7	80	232	24	23
8	99	241	15	4
9	82	211	45	21
10	76	215	41	27

Yapılan testin sayısal sonuçları Tablo 4.14. de gösterilmektedir.

Tablo 4.15. K-ortalama algoritması ROC analizi tablosu

Test Sayısı	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
1	86,94	79,51	89,92
2	86,07	76,70	89,84
3	84,12	76,70	87,11
4	88,58	73,79	94,53
5	88,30	77,67	92,58
6	83,84	66,99	90,63
7	86,91	77,67	90,63
8	94,71	96,12	94,14
9	81,62	79,61	82,42
10	81,06	73,79	83,98



Şekil 4.7. K-ortalama algoritması test sonuçlarının grafiği

Yapılan testlerin sayısal ROC analizi sonuçları Tablo 4.15. de gösterilmektedir. Testlerin duyarlılık, özgüllük ve doğruluk oranlarının grafiği Şekil 4.7. de gösterilmektedir.

Tablo 4.16. K-ortalama algoritması testlerinin ortalama değerleri

Formül	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
Değer / 10	86,94	79,51	89,92

10 testin tamamının ortalama değerleri Tablo 4.16. de gösterilmektedir.

BÖLÜM 5. SONUÇ VE DEĞERLENDİRMELER

Bu çalışmada, demir eksikliği anemisinin hastalığının tespiti için çeşitli yöntemler denenmiştir. Hastalık tespiti için, ileri beslemeli geri yayımlı yapay sinir ağı, destek vektör makinesi, Naive Bayes algoritması, regresyon ağaçları ve k-ortalama algoritmaları kullanılmıştır. Bu yöntemlerden elde edilen sonuçlar Bölüm 4 da detaylı bir şekilde gösterilmiştir.

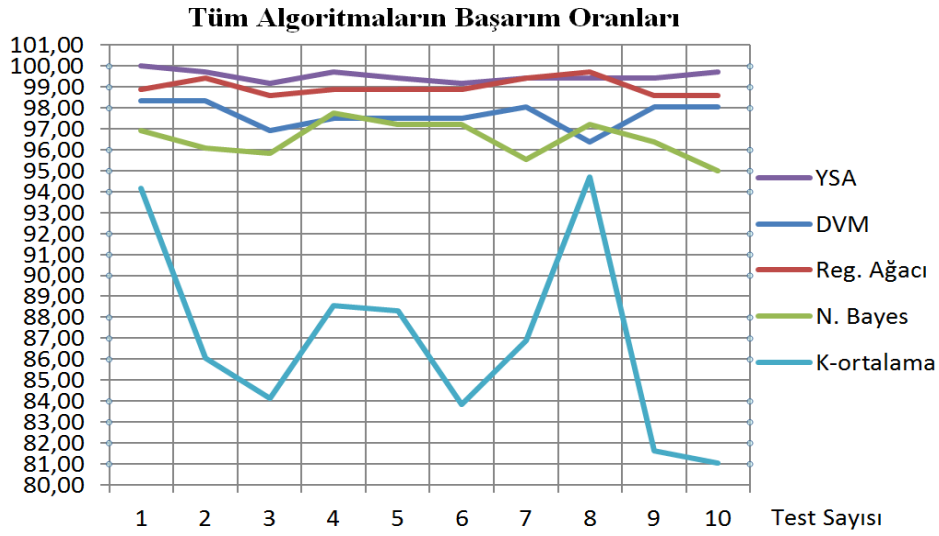
Hastalık tespiti için öncelikle verilerde optimizasyon yapıldı. Daha sonra öğreticili ya da öğreticisiz yöntem kullanılmasına bağlı olarak eğitim aşaması ve test aşaması yapıldı ya da sadece test aşaması yapıldı. Bu işlemler Holdout metodu ile giriş verileri rastgele olacak şekilde 10 kez tekrarlandı. Buradan alınan sonuçlar Bölüm 4 da detaylı bir şekilde gösterildi.

Bölüm 4 te elde edilen sonuçlar arasında kıyaslama ise bu bölümde yapılmıştır.

Tablo 5.1. Tüm algoritmaların ortalama değerleri

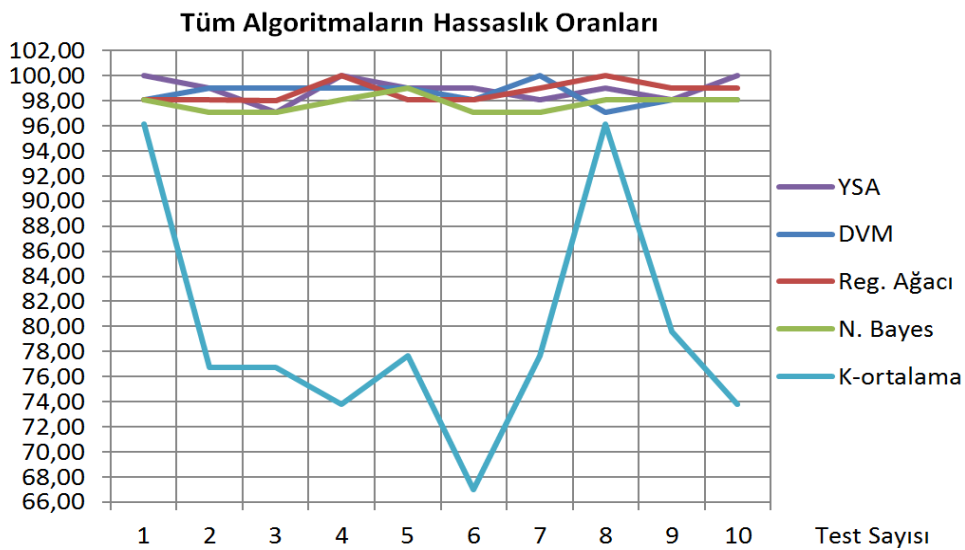
Yöntemler	Doğruluk(%)	Duyarlılık(%)	Özgüllük(%)
YSA	99,53	98,93	99,77
DVM	97,66	98,54	97,30
Regresyon Ağacı	99,00	98,73	99,49
Naive Bayes	96,52	97,77	96,02
K-Ortalama	86,94	79,51	89,92

Tablo 5.1 de görüldüğü üzere kullanılan yöntemlerin ortalama değerleri gösterilmektedir. Bu ortalama değerler alınan tüm test sonuçlarının, test sayısına bölünmesi sonucu elde edilmiştir.



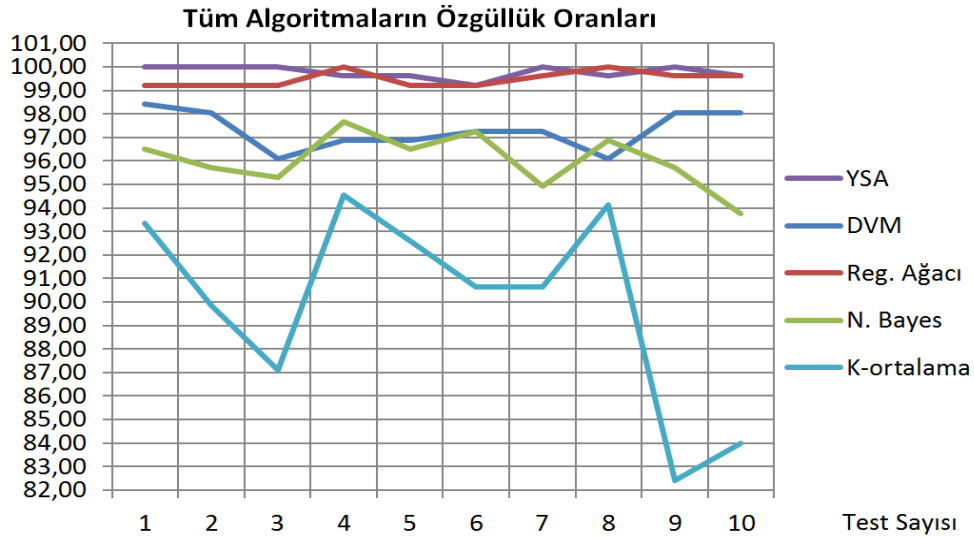
Şekil 5.1. Tüm algoritmaların başarımları

Görsel şekli Şekil 5.1 de görüldüğü üzere yapılan 10 testin sonucunda kullanılan algoritmaların başarımları karşılaştırılmıştır. İleri beslemeli yapay sinir ağıları görüldüğü üzere yedinci ve sekizinci testler dışında tüm testlerde en iyi başarımlarını elde etmiştir. Onu takip eden en iyi yöntem ise regresyon ağaçlarıdır. Sırasıyla başarımlarına göre en iyi sonuçlar öncelik yapay sinir ağıları olmak üzere sırasıyla regresyon ağaçları, destek vektör makinesi, Naive Bayes algoritması ve k-ortalama algoritmasıdır.



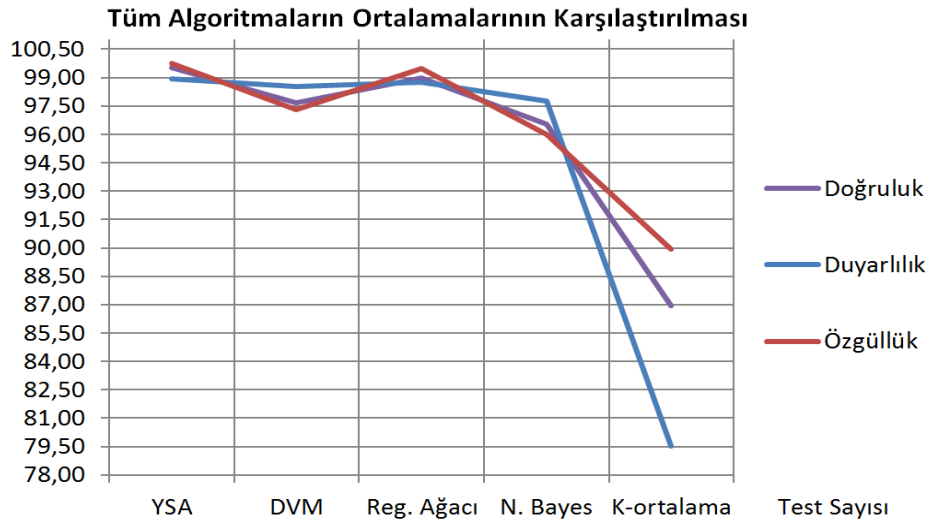
Şekil 5.2. Tüm algoritmaların hassaslık oranları

Görsel şekli Şekil 5.2 de görüldüğü üzere yapılan 10 testin sonucunda kullanılan algoritmaların hassaslık oranları karşılaştırılmıştır. Yapılan testlerin sonucunda k-ortalama dışındaki yöntemlerin birbirine yakın hassaslık oranları elde ettiği görülmektedir.



Şekil 5.3. Tüm algoritmaların özgüllük oranları

Şekil 5.3 de görüldüğü üzere yapılan 10 testin sonucunda kullanılan algoritmaların özgüllük oranları karşılaştırılmıştır. İleri beslemeli yapay sinir ağı ve regresyon ağaçları birbirine yakın sonuçlar elde etmiştir. Sıralama yapılacak olursa, özgüllük oranlarına göre ilk olarak yapay sinir ağları, daha sonra diğerlerinde ise regresyon ağaçları, destek vektör makinesi, Naive Bayes algoritması ve k-ortalama algoritması şeklinde sıralanacaktır.



Şekil 5.4. Tüm algoritmaların ortalamalarının karşılaştırılması

Şekil 5.4 de görüldüğü üzere yapılan testlerin sonucunda doğruluk oranlarının, duyarlılık oranlarının, özgüllük oranlarının ortalamaları alınarak karşılaştırılmıştır. Alınan bu ortalamalara sonucunda en iyi doğruluk oranı olarak %99.53, en iyi duyarlılık oranı olarak %98.93, en iyi özgüllük oranı olarak ise %99.77 elde edilmiştir. Bu elde edilen en iyi sonuçlar ise ileri beslemeli geri yayımlı yapay sinir ağından elde edilmiştir. Ağlar kıyaslanacak olursa en iyi sonuçlar öncelikle ileri beslemeli yapay sinir ağından, daha sonra sırasıyla regresyon ağaçlarından, destek vektör makinesinden, Naive Bayes algoritmasından ve k-ortalama algoritmasından elde edilmiştir.

Bu sonuçlar göz önünde bulundurulduğunda çalışmadaki anemi hastalığının tespiti için en uygun metot ileri beslemeli geri yayımlı yapay sinir ağları olarak tespit edilmiştir.

Ayrıca Tablo 1.4 de gösterilen literatür çalışmalarıyla kıyaslanınca elde edilen sonuçların literatürdeki başarı oranlarından daha yüksek başarı oranlarına sahip olduğu tespit edilmiştir.

5.1. Çalışmanın Faydaları

Yapılan tez çalışmasının faydaları şu şekilde sıralanabilir:

1. Tez çalışmasının 4. bölümünde kullanılan farklı yöntemler ile detaylı bir şekilde demir eksikliği anemisi tanısı için hangi yöntemin daha etkili olabileceği verilmiştir.
2. Anemi hastalığı gibi yöntemlerin bilgisayar destekli ortamlarda teşhis edilebileceği görüşü, test sonuçlarındaki başarı oranları ile desteklenmiştir.
3. Gerçekleştirilen uygulama ile hasta sayısının fazla ama doktor sayısının az olduğu ülkelerde teşhis aşaması için harcanan zamanın en aza indirilmesi sağlanabilecek bir uygulama tasarlanmıştır.

KAYNAKLAR

- [1] BEUTLER, E., LICHTMAN, M. A., COLLER, B. S., KIPPS, T. J. ve SELIGSOHN, U. Williams Hematology, NY: McGraw-Hill, 289-294, 2001.
- [2] VOS, T., FLAXMAN, A. D., NAGHAVI, M. ve d., Years lived with disability (ylds) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010, Lancet, 380, 216-396, 2012.
- [3] BLANC, B., FINCH, C. A., HALLBERG, L. ve d., Nutritional anaemias : Report of a who scientific group (Technical Report 405), World Health Organization, Geneva, Switzerland, 405, 1-401, 1968.
- [4] RAPAPORT, S.I., Diagnosis of anemia, Introduction to Hematology, PA: JB Lippincott Company, 10-38, 1987.
- [5] TIETZ, N. W., Clinical Guide to Laboratory Tests, PA: WB Saunders Co, 1995.
- [6] JANDL, J. H., Heinz body hemolytic anemias, Blood: Textbook of Hematology, MA: Little, Brown ve Company, 503-518, 1996.
- [7] GOYETTE, R. E., Hematology. A Comprehensive Guide to the Diagnosis and Treatment of Blood Disorders, CA: Practice Management Information Corporation (PMIC), 1997.
- [8] LEE, G. R., FOERSTER, J., LUKENS, J., PARASKEVAS, F., GREER, J. P. ve RODGERS, G. M., Anemia: a diagnostic strategy, Wintrobe's Clinical Hematology, MD: Williams and Wilkins, 908-940, 1998.
- [9] HOFFMAN, R., BENZ, E. J. Jr, SHATTIL, S. J. ve FURIE, B., Hematology: Basic Principles and Practice, NY: Churchill-Livingstone, 2004.
- [10] MATHERS, C. D., BERNARD, C., MOESGAARD, I.K. ve d., The global burden of disease in 2002: data sources, methods and results (Technical Report 54), World Health Organization, Geneva, Switzerland, 2003.

- [11] HU, Y. H. ve HWANG, J. N., Neural Network Solutions to Signal Processing Problems in Handbook of Neural Network Signal Processing, CRC Press, ISBN 0-8493-2359-2, 2002.
- [12] GANESAN, N., VENKATESH, K. ve RAMA, M. A, Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data, International Journal of Computer Applications, 1, 76-85, 2010.
- [13] EBRAHIMZADEH, A. ve RANAEE, V., Control chart pattern recognition using an optimized neural network and efficient features, The International Society of Automation, 49, 387-393, 2010.
- [14] OU, G. ve MURPHEY, Y. L., Multi-class pattern classification using neural networks, Pattern Recognition, 40, 4-18, 2007.
- [15] NAMANE, A., GUESSOUM, A., SOUBARI, E. H. ve MEYRUEIS, P., CSM neural network for degraded printed character optical recognition, Journal of Visual Communication and Image Representation, 25, 1171-1186, 2014.
- [16] YAICI, W. ve ENTCHEV, E., Performance prediction of a solar thermal energy system using artificial neural networks, Applied Thermal Engineering, 73, 1346-1357, 2014.
- [17] XANTHOPOULOS, P. ve RAZZAGHI, T., A weighted support vector machine method for control chart pattern recognition, Computers & Industrial Engineering, 70, 134 – 149, 2014.
- [18] BURGESS, J. C., A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery, 2, 121 – 167, 1998.
- [19] LU, S. H., CHIANG, D. A., KEH, H. C. ve HUANG, H. H., Chinese text classification by the Naïve Bayes Classifier and the associative classifier with multiple confidence threshold values, Knowledge-Based Systems, 23, 598-604, 2010.
- [20] ÖZKAN K., Sınıflandırma ve regresyon ağacı tekniği (SRAT) ile ekolojik verinin modellenmesi, SDU Faculty of Forestry Journal, 13, 1-4, 2012.
- [21] CHEN, H., ZHANG, J., XU, Y., CHEN, B. ve ZHANG, K., Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans, Expert Systems with Applications, 39, 11503-11509, 2012.

- [22] KIANG, M.Y., A comparative assessment of classification methods, *Decision Support Systems*, 35, 441–454, 2003.
- [23] NEDIC, V., DESPOTOVIC, D., CVETANOVIC, S., DESPOTOVIC, M. ve BABIC S., Comparison of classical statistical methods and artificial neural network in traffic noise prediction, *Environmental Impact Assessment Review*, 49, 24-30, 2014.
- [24] YILMAZ, Z. ve BOZKURT, M. R., Determination of women iron deficiency anemia using neural networks. *Journal of Medical Systems*, 36, 2941-2945, 2012.
- [25] WONGSEREE, W., CHAIYARATANA, N., VICHITTUMAROS, K., WINICHAGOON, P. ve FUCHAROEN, S., Thalassaemia classification by neural networks and genetic programming, *Information Sciences*, 177, 771-786, 2007.
- [26] BIRNDORF, N. I., PENTECOST, J. O., COAKLEY, J. R. ve SPACKMAN, K. A., An expert system to diagnose anemia and report results directly on hematology forms, *Computers and Biomedical Research*, 29, 16-26, 1996.
- [27] AZARKHISH, I., RAOUFY, M. R. ve GHARIBZADEH, S., Artificial Intelligence Models for Predicting Iron Deficiency Anemia and Iron Serum Level Based on Accessible Laboratory Data, *Journal of Medical Systems*, 36, 2057-2061, 2012.
- [28] YILMAZ, Z. ve OCAK, Ş., *Bulanık Mantık ile Aneminin Belirlenmesi*, 1. Mühendislik ve Teknoloji Sempozyumu, 2008.
- [29] http://www.nyurtay.sakarya.edu.tr/sites/nyurtay.sakarya.edu.tr/file/Kans%EF%BF%BDz1%EF%BF%BDkmakale_v5_.doc, Erişim Tarihi: 06. 11. 2014.
- [30] <http://www.medicinenet.com/script/main/art.asp?articlekey=15491>, Erişim Tarihi: 11. 11. 2014.
- [31] <https://www.flickr.com/photos/euthman/2274260085/in/photostream/>, Erişim Tarihi: 11. 11. 2014.
- [32] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia>, Erişim Tarihi: 12. 11. 2014.
- [33] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia>, Erişim Tarihi: 12. 11. 2014.

- [34] <http://tr.wikipedia.org/wiki/Kans%C4%B1z%C4%B1k>, Eriřim Tarihi: 12. 11. 2014.
- [35] <http://en.wikipedia.org/wiki/HIV/AIDS>, Eriřim Tarihi: 07. 11. 2014.
- [36] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia/causes>, Eriřim Tarihi: 12. 11. 2014.
- [37] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia/atrisk>, Eriřim Tarihi: 12. 11. 2014.
- [38] <http://doktoru.us/genel-dahiliye/demir-eksikligi-anemisi-belirtileri.html>, Eriřim Tarihi: 11. 11. 2014.
- [39] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia/signs>, Eriřim Tarihi: 12. 11. 2014.
- [40] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia/diagnosis>, Eriřim Tarihi: 12. 11. 2014.
- [41] <http://www.nhlbi.nih.gov/health/health-topics/topics/bt>, Eriřim Tarihi: 07. 11. 2014.
- [42] <http://www.nhlbi.nih.gov/health/health-topics/topics/anemia/treatment>, Eriřim Tarihi: 12. 11. 2014.
- [43] ZUPAN, J., Basics of artificial neural networks, Data Handling in Science and Technology, 23, 199-229, 2003.
- [44] ÖZTEMEL, E., Yapay Sinir Ağları, Papatya Yayıncılık, ISBN: 978-975-6797-39-6, 2003.
- [45] SATHYA, R. ve ABRAHAM, A., Comparison of supervised and unsupervised learning algorithms for pattern classification, International Journal of Advanced Research in Artificial Intelligence (IJARAI), 2, 34-36, 2013.
- [46] KAFETZOPOULOU, L. E., BOOCOOCK, D. J., DHONDALAY, G. K. R. ve d., Biomarker Identification in Breast Cancer: Beta-Adrenergic Receptor Signaling and Pathways to Therapeutic Response, Research Network of Computational and Structural Biotechnology, 6, 1-7, 2013.

- [47] KOONSANIT, K., JARUSKULCHAI, C. ve EIUMNOH, A., Determination of the initialization number of clusters in K-means clustering application using Co-occurrence statistics techniques for multispectral satellite imagery, *International Journal of Information and Electronics Engineering*, 2, 785–789, 2012.
- [48] TAN, P. N., STENINBACH, M. ve KUMAR, V. , *Introduction to Data Mining*, Addison Wesley, 2006.
- [49] http://en.wikipedia.org/wiki/Support_vector_machine, Erişim Tarihi: 11. 11. 2014.
- [50] CHEN, S., YU, P., Pruning of support vector networks on flood forecasting, *Journal of Hydrology*, 347,67-78, 2007.
- [51] SERRANO, J., TOMECKOVA ve M., ZVAROVA, J., *Machine Learning Methods for Knowledge Discovery in Medical Data on Atherosclerosis*, *European Journal for Biomedical Informatics*, 2, 6-33, 2006.
- [52] YANG, J., ZHONG, N., LIANG, P., WANG, J. ve d., Brain activation detection by neighborhood one-class SVM, *Cognitive Systems Research*, 11, 16 – 24, 2010.
- [53] QIN, B., XIA, Y., WANG, S. ve DU, X., A novel Bayesian classification for uncertain data, *Knowledge-Based Systems*, 24, 1151-1158, 2011.
- [54] PAWLAK, Z., A Rough Set View on Bayes' Theorem, *International Journal Of Intelligent Systems*, 18, 487-498, 2003.
- [55] http://tr.wikipedia.org/wiki/Naive_Bayes_s%C4%B1n%C4%B1fland%C4%B1r%C4%B1c%C4%B1, Erişim Tarihi: 08. 11. 2014.
- [56] [https://alliance.seas.upenn.edu/~cis520/wiki/index.php?n=Lectures.Nai veBayes](https://alliance.seas.upenn.edu/~cis520/wiki/index.php?n=Lectures.Nai%20veBayes), Erişim Tarihi: 17. 11. 2014.
- [57] http://en.wikipedia.org/wiki/Receiver_Operating_Characteristic_Curve_Explorer_and_Tester, Erişim Tarihi: 18. 11. 2014.
- [58] <http://www.cs.cmu.edu/~schneide/tut5/node42.html>, Erişim Tarihi: 19. 11. 2014.
- [59] FAWCETT, T., An Introduction to ROC Analysis, *Pattern Recognition Letters*, 27, 861–874, 2006.

ÖZGEÇMİŞ

Volkan SEYMEN 1990 yılında İstanbul ili Kadıköy ilçesinde doğdu. İlköğretimini Küçükyalı Merkez İlkokulu, ortaöğretimini Mustafa Saffet Anadolu Lisesi'nde tamamladı. Lisans eğitimine 2009 yılında Sakarya Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği bölümünde başladı. 2013 yılında lisans eğitimini tamamladı. Aynı yıl içerisinde Sakarya Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği bölümünde Araştırma Görevlisi olarak çalışmaya başladı. Aynı zamanda Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar ve Bilişim Mühendisliği'nde ve Mekatronik Mühendisliği'nde yüksek lisans eğitimine devam etmektedir.