T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

TIP BİLİŞİMİ VE VERİ MADENCİLİĞİ UYGULAMALARI: EEG SİNYALLERİNDEKİ EPİLEPTİFORM AKTİVİTEYE VERİ MADENCİLİĞİ YÖNTEMLERİNİN UYGULANMASI

Ümit Can KUMDERELİ

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

Danışman : Doç. Dr. Erdem UÇAR

EDİRNE, 2012

T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

TIP BİLİŞİMİ VE VERİ MADENCİLİĞİ UYGULAMALARI: EEG SİNYALLERİNDEKİ EPİLEPTİFORM AKTİVİTEYE VERİ MADENCİLİĞİ YÖNTEMLERİNİN UYGULANMASI

Ümit Can KUMDERELİ

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

Tez Danışmanı : Doç. Dr. Erdem UÇAR

EDİRNE, 2012

T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

TIP BİLİŞİMİ VE VERİ MADENCİLİĞİ UYGULAMALARI: EEG SİNYALLERİNDEKİ EPİLEPTİFORM AKTİVİTEYE VERİ MADENCİLİĞİ YÖNTEMLERİNİN UYGULANMASI

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

Ümit Can KUMDERELİ

Bu tez 27/ 07 / 2012 tarihinde Aşağıdaki Jüri Tarafından Kabul Edilmiştir.

Doç. Dr. Erdem UÇAR	Doç. Dr. Necdet SÜT	Yrd. Doç. Dr. İlhan UMUT
(Danışman)	(Üye)	(Üye)

Yüksek Lisans Tezi
Tıp Bilişimi ve Veri Madenciliği Uygulamaları: EEG Sinyallerindeki Epileptiform Aktiviteye Veri Madenciliği Yöntemlerinin Uygulanması
T.Ü. Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

ÖZET

İnsanoğlunun ilk var olduğu günlerden bu yana tıp bilimi çok hızlı bir şekilde gelişimi sürdürmektedir. Her geçen yıl hem teşhis hem de tedavi alanlarında çok başarılı sistemler geliştirildiği görülmektedir. Özellikle günümüzde bilişim sistemlerinin de kullanılmasıyla hastane bilgi istemleri ve klinik bilgi sistemleri, tıbbi görüntüleme ve sinyal analizi, tıbbi karar destek, nöro bilimler gibi alanlarda çok başarılı olunduğu herkesçe bilinmektedir. Tıp bilişimi de yukarıda saydığımız alanlarda tıp ve bilişim alanlarını harmanlayarak her iki alanda da çalışmak isteyen araştırmacı ve akademisyenleri yepyeni bir çalışma sahası ortaya koymuştur.

Bu çalışmada, tıp bilişiminin bir alt kolu olan tıbbi görüntü ve sinyal analizi alanı ele alınmıştır. Elektroansefalografi (EEG) cihazlarından almış olduğumuz veriler kendi geliştirdiğimiz yazılım sayesinde analiz edilip, ayrık fourier ve dalgacık dönüşümüne tabi tutulmuştur. Diğer yapılan çalışmalardan ayrı olarak 19 kanallı 30 dakikalık rutin çekimler sonucu elde edilen veriler kullanılmış olup bu veri setine uygun öznitelik parametreleri seçilmiştir. Verilerin veri madenciliği sürecinde öznitelik parametrelerine karşılık gelen öznitelik değerleri hesaplanmıştır. Ardından veriler üzerinde çeşitli sınıflandırma algoritmaları uygulanarak; bu algoritmaların yüksek doğruluk oranı ile epileptik aktiviteyi teşhis edenleri belirlenmiştir. Böylece nöroloji uzmanlarının ileride epileptik aktiviteyi belirlemede teşhis süresini ve teşhisin doğruluğunu artıracak bir model oluşturulmaya çalışılmıştır.

Yıl **:** 2012

Sayfa Sayısı **:** 130

Anahtar Kelimeler **:** Veri madenciliği, epileptiform aktivite, EEG,WEKA

Master Thesis
Medical Informatics and Data Mining Applications: Application of Data Mining Methods in to Epileptiform Activity in EEG Signals
Trakya University. Institute of Naturel Sciences
Computer Engineering Department

ABSTRACT

Since the first days when human beings have a very fast development of medical science continues. With each passing year, both systems are very successful in the areas of diagnosis and treatment are developed. Prompts the use of information systems, particularly in our hospital and clinical information systems, medical imaging and signal analysis, medical decision support, we were very successful in areas such as neuro sciences known to all. The above-mentioned areas in the fields of medicine, medical informatics and informatics researchers and academics who want to work in both areas by blending a whole new field of study has revealed.

In this study, medical informatics and a lower arm of the medical image signal analysis area is considered. Electroencephalography (EEG) devices have received data will be analyzed through their own developed software, were subjected to discrete Fourier and wavelet transform. Other studies conducted in a separate 19 channels of 30 minutes has been used routinely in this data set-ups of the data from the appropriate attribute parameters chosen. Attribute data, data mining process parameters corresponding to the calculated attribute values. Then the various classification algorithms on the data with the application of these algorithms to diagnose epileptic activity was determined with high accuracy. Thus, the epileptic activity in the neurology experts in determining the future of a model to increase the accuracy of the diagnosis time and attempted to establish the diagnosis.

Year **:** 2012

Number of Pages **:** 130

Keywords **:** Data Mining, Epileptiform Activity, EEG, WEKA

ÖNSÖZ VE TEŞEKKÜRLER

Çalışmalarımı yönlendiren ve araştırmalarımın her aşamasında bilgi, öneri ve yardımlarını esirgemeyen danışmanım Doç. Dr. Erdem UÇAR'a teşekkürlerimi sunarım.

Çalışmanın nöroloji ile ilgili bölümlerinde TÜ Tıp Fakültesi Nöroloji ABD öğretim üyesi Prof. Dr. Yahya ÇELİK ve Yrd. Doç. Dr. Aslan TEKTAŞ'a, EEG sinyallerinin analizini sağlayan programı yazmamda destek olan hocam TÜ Mühendislik -Mimarlık Fakültesi öğretim üyesi Yrd. Doç. Dr. İlhan UMUT'a' ve veri madenciliği ve analiz kısmındaki katkılarından dolayı değerli mesai arkadaşım Arş. Gör. E. Serdar GÜNER'e teşekkürü bir borç bilirim.

Ayrıca tezin oluşturulma sürecinde desteklerini hep yanımda hissettiğim eşime ve aileme sonsuz sevgi ve saygılarımı sunarım.

Ümit Can KUMDERELİ
Edirne, 2012

İÇİNDEKİLER

ÖZET iv
ABSTRACT v
ÖNSÖZ VE TEŞEKKÜRLER vi
İÇİNDEKİLER vii
SİMGELER VE KISALTMALAR xii
TABLOLAR LİSTESİ xiv
ŞEKİLLER LİSTESİ xv
BÖLÜM 1 1
GİRİŞ 1
BÖLÜM 2. 7
EPİLEPSİ (SARA) HASTALIĞI 7
2.1. Epilepsi (Sara) Nedir? 7
2.2. Epilepsi' nin Tarihçesi 9
2.3. Epilepsi Olan Ünlüler 10
2.4. Epilepsi'ye Neden Olan Etkenler 11
2.5. Epilepsiyi Tetikleyen Etkenler 11
2.6. Kalıtımsal (Genetik) faktörler 12
2.7. Nöbet Türleri 13
2.7.1. Kısmi nöbetler 13
2.7.1.1. Basit kısmi nöbetler 13
2.2.1.2. Karmaşık kısmi nöbetler 14
2.7.2. Dalma (Absans) nöbeti 14
2.7.3. Miyoklonik nöbetler 14
2.7.4. Tonik-klonik (grand mal) nöbetler 15
2.7.4.1. Tonik nöbetler 15
2.7.4.2. Atonik nöbetler 15
BÖLÜM 3 17
ELEKTROENSEFALOGRAFİ (EEG) 17
3.1. EEG Nedir? 17

3.2. EEG Aktivitesinin Fizyolojik Temeli 18
3.3. EEG Kaydı 20
3.3.1. Elektrotlar 20
3.3.2. Filtreler 21
3.3.3. Artefaktlar 21
3.3.4. Kayıt 22
3.4. EEG Aktivitesini Tanımlayan Nitelikler 23
3.4.1. Dalga Formu 23
3.4.2. Tekrarlama 25
3.4.3. Frekans 25
3.4.4. Genlik(Amplitüt) 26
3.4.5. Dağılım 27
3.4.6. Faz İlişkisi 28
3.4.7. Zamanlama 28
3.4.8. Israrlılık 29
3.4.9. Reaktivite 29
3.5. EEG'de Kaydedilen Normal Aktivite 29
3.5.1. Alfa Aktivitesi 29
3.5.2. Beta Aktivitesi 30
3.5.3. Teta Aktivitesi 31
3.5.4. Delta Aktivitesi 31
3.5.5. Mü Ritmi 31
3.5.6. Lambda Dalgaları 32
3.5.7. 14 ve 6 Hz Dikenler 32
3.5.8. Küçük Keskin Dikenler 32
3.5.9. 6 Hz Diken- Dalga Aktivitesi 33
3.5.10. Uykuya Geçişin Ritmik Temporal Başalımları 33
3.5.11. Ritmik Teta Boşalımları 33
3.5.12. Yaşla İlgili EEG Değişiklikleri 34
3.6. Anormal EEG'nin Tanımlanması 35
3.7. EEG Desenleri ile Nörolojik Hastalıklar Arasındaki Bağıntı 35
BÖLÜM 4 37
EEG SİNYAL ANALİZİ VE MODELLEME METODLARI 37
4.1. Sinyal Analiz Metotlarının Gelişimi 37

4.2. Fourier Temelli Sinyal Analizi 38
4.2.1. Akan fourier dönüşümü 39
4.2.2. Kısa zaman fourier dönüşümü 39
4.2.3. Zak dönüşümü 40
4.2.4. Ayrık fourier dönüşümü 41
4.2.5. Hızlı fourier dönüşümü 42
4.5. Dalgacık Dönüşümü (Wavelet) ile Sinyal Analizi 42
4.5.1. Sürekli Dalgacık Dönüşümü 43
4.5.2. Ayrık Dalgacık Dönüşümü 44
BÖLÜM 5 46
VERİ MADENCİLİĞİ (DATA MINING) 46
5.1. Veri Madenciliği Nedir? 46
5.2. Veri Madenciliği Tarihçesi 48
5.3. Veri Madenciliğinin Gelişimini Etkileyen Faktörler 49
5.4. Veri Madenciliği Uygulama Alanları 50
5.5. Veri Madenciliği ve Diğer Disiplinler 51
5.6. Veri Madenciliği Gereksinimleri 52
5.7. Veri Madenciliği Uygulanmasında Karşılaşılan Güçlükler 53
5.8. Veri Madenciliği Türleri 54
5.9. Veri Madenciliği Yazılımları 54
5.9.1. Spss 54
5.9.2. Clementine 55
5.9.3. Sas 55
5.9.4. Enterprise miner 55
5.9.5. Kxen 56
5.9.6. Insightful miner 56
5.9.7. Affinium model 56
5.9.8. Statistica Data Miner 57
5.9.9. Inlen 57
5.9.10. DBMiner 57
5.9.11. Darwin 58
5.9.12. WEKA 58
5.10. Veri Madenciliği Süreci 59
5.10.1. İşi ve İş Ortamını Anlamak 60

5.10.2. Veriyi Anlamak....60
5.10.3. Veri Hazırlama....61
5.10.4. Modelleme....62
5.10.5. Değerlendirme....62
5.10.6. Yayma....63
5.11. Veri Madenciliği ve Makine Öğrenmesi Kavramı....64
5.12. Veri Madenciliği Modelleri....66
5.12.1. Tanımlama ve Ayrımlama....68
5.12.2. Birliktelik Analizi....68
5.12.3. Sınıflandırma ve Öngörü Yöntemi....69
5.12.3.1. Karar ağaçları....70
5.12.3.2. Yapay sinir ağları....71
5.12.3.2.1 Radial basis function (RBF) network....73
5.12.3.3. Bellek tabanlı yöntemler....73
5.12.3.4. k-en yakın komşu....74
5.12.3.5. Genetik algoritmalar....75
5.12.3.6. Naive Bayes ile sınıflandırma yöntemi....77
5.12.3.7. Bayes Ağları (Bayes Net) yöntemi....78
5.12.3.8. Vaka tabanlı nedenleşme yöntemi....78
5.12.3.9. Kaba küme teorisi yöntemi....79
5.12.4. Kümeleme analizi....79
5.12.5. İstisna analizi....80
5.12.6. Evrimsel analiz....81
5.13. Tıbbi Verilerle Veri Madenciliği....82
5.13.1. Tıbbi verilerle veri ambarının oluşturulması....82
5.13.2. Tıbbi verilerin özellikleri....83
BÖLÜM 6....85
GERÇEKLEŞTİRİLEN UYGULAMA....85
6.1. Çalışmada Kullanılan Verilerin Elde Edilmesi....85
6.2. Verilerden Öznitelik Parametrelerinin Seçilmesi....89
6.3. Weka'da Model Kurma....92
6.4. Model Değerlendirme....94
BÖLÜM 7....99
SONUÇLAR VE ÖNERİLER....99

KAYNAKLAR 101
EKLER.....106
ÖZGEÇMİŞ.....115

SİMGELER VE KISALTMALAR

ADD : Ayrık Dalgacık Dönüşümü

AFD : Ayrık Fourier Dönüşümü

AR : Otoregresif (Auto Regressive)

ARMA : Otoregresif Hareketli Ortalama (Auto Regressive Moving Average)

BT : Bilgisayarlı Tomografi

C&RT : Classification and Regresion Trees

CHAID : Chi-squared Automatic Interaction Detector

ÇKA : Çok Katmanlı Ağ

DD : Dalgacık Dönüşümü

Deeg : Dijital Elektroensefalogram

Edf :European Data Format

EEG : Elektroensefalogram

EKG : Elektrokardiyogram

EMG : Elektromiyogram

EV : Özvektör (Eigenvector)

FD : Fourier Dönüşümü

GA : Genetik Algoritma

GID3 : Generalized Iterative Dichotomiser 3

GSY : Güç Spektral Yoğunlugu

HFD : Hızlı Fourier Dönüşümü

Hz : Hertz

ID3 : Iterative Dichotomiser 3

KZFD : Kısa Zamanlı Fourier Dönüşümü

MA : Hareketli Ortalama (Moving Average)

MRG : Magnetik Rezonans Görüntüleme

MUSIC : Çoklu Sinyal Sınıflama (Multiple Signal Classification)

p-p : Tepeden Tepeye (Peak to Peak)

RTFA : Radyal Tabanlı Fonksiyon Ağı

WEKA : Waikato Environment for Knowledge Analysis

WNN : Dalgacık Sinir Ağı (Wavelet Neural Network)

YSA : Yapay Sinir Ağı

α : Alfa

β : Beta

γ : Gamma

δ : Delta

ε : Teta

TABLOLAR LİSTESİ

Tablo 1. EEG Dalgalarının Frekans Altbandları 26
Tablo 2. Veri Madenciliği Tarihçesi 48
Tablo 3.Veri Madenciliği Yazılımlarının Kıyaslanması 59
Tablo 4. Hava Durumu Veri Seti.....71
Tablo 5. Düzensizlik Matrisi.....95
Tablo 6. Sınıflara Göre Doğruluk Çizelgesi.....96
Tablo 7. Weka 'daki Sınıflandırma Algoritmalarının Başarı Oranları.....98

ŞEKİLLER LİSTESİ

Şekil 1. Kafatasının saçlı derisine yerleştirilen elektrotlardan alınan sinyaller 18
Şekil 2. Elektroensefalografi (EEG) kaydının alınması 22
Şekil 3. EEG Görülen Dalga Formları 25
Şekil 4. (A)Sinyal ile küçük ölçekli dalgacığın karşılaştırılması. (B)Sinyal ile kaydırılmış küçük ölçekli dalgacığın karşılaştırılması. 43
Şekil 5 (A) Sinyal ile kaydırılmış büyük ölçekli dalgacığın birleşimi. (B) İki boyutlu ölçek- zaman eğrisi. 44
Şekil 6. Dalgacık ayrıştırma ağacı 45
Şekil 7. Veri madenciliği aşamaları 47
Şekil 8. Veri Madenciliğinin Kullanıldığı Alanlara Göre Dağılımı 51
Şekil 9. Veri Madenciliği ve Diğer Disiplinler 52
Şekil 10. CRİSP-DM Veri Madenciliği Süreci 60
Şekil 11. En Çok Kullanılan Veri Madenciliği Teknikleri 67
Şekil 12.Örnek Bir Karar Ağacı Yapısı 71
Şekil 13. Örnek Bir Yapay Sinir Ağı Yapısı 72
Şekil 14.Genetik Algoritma Süreci 76
Şekil 15. Şekil 16. İstisna Analizi ve Küme Oluşumları 81
Şekil 17. Uluslararası 10-20 Sistemi ve Elektrotların Yerleşimi 86
Şekil 18.Örnek EEG Trasesi Üzerinde Montajlar 87
Şekil 19. Polyman Programı Yardımıyla LP ve HP Filtrelerinin Uygulanması 87
Şekil 20. Epilepsi Hastası Olanlar ile Kontrol Gruplarının Karşılaştırılması 88
Şekil 21.Verilerin Analizini Gerçekleştiren Yazılım 89
Şekil 22. Analiz Tablosu 90
Şekil 23. ARFF Dosya Formatındaki Veri Kümesi Parçası 92
Şekil 24. WEKA'da Ön İşleme Paneli 93
Şekil 25. Weka Classifier (Sınıflandırıcı) Paneli 95

BÖLÜM 1

GİRİŞ

İçinde bulunduğumuz bilgi çağının vazgeçilmez parçaları olan Bilişim Teknolojileri ve Bilişim Sistemleri hayatın hemen her alanında etkin bir şekilde kullanılmaktadır. Bu sistemlerin ve teknolojilerin etkin bir şekilde kullanıldığı sektörlerin başında sağlık sektörü gelmektedir. Sağlık sektöründe kullanılan bilişim sistemleri Sağlık Bilişim Sistemleri olarak adlandırılmaktadır. Sağlık Bilişim Sistemleri'nden teşhis ve tedavi sistemleri, hasta bakımı, tıbbi bilgi yönetimi, tıbbi dokümantasyon ve kalite yönetimi gibi alanlarda yararlanılmaktadır. Sağlık Bilişim Sistemleri, sağlık bakım hizmetlerinin kalitesini geliştirmekte, hekimlerin hastalıkları erken teşhis etmesine yardımcı olmaktadır. Sağlık Bilişim Sistemleri hastalıklara erken teşhis ile tedaviye ayrılan süreyi azaltmakla, sağlık bakım maliyetlerini azaltarak, sağlık bakım hizmetlerinden elde edilen gelirlerin yükselmesini sağlamaktadır. (Farboudi,2009)

Son yıllarda Veri Madenciliği uygulamaları da sağlık kurumlarında Sağlık Bilişim Sistemleri bağlamında kullanılan bir uygulama olarak kullanılmaktadır. Veri Madenciliği uygulamaları sağlık kurumlarında hem teşhis ve tedavi hizmetlerinin etkinliğinin artırılması hem de kurumlarda daha etkin bir yöntem sağlamak amacıyla kararları destekleyen bir sistem olarak değerlendirilebilir. Bu bağlamda veri madenciliği; kurumlarda hastalara uygulanan tedavi yöntemlerinin geliştirilmesi ve etkinliğinin artırılması, tedavi sürelerinin kısaltılmasına yönelik araştırmaların yapılması, bölgesel anlamda hastalık gruplarının belirlenerek sağlık hizmetlerinin etkinliğinin artırılması, yönetimle ilgili kararların alınması, kararları destekleyici bilgilerin etkinliğinin artırılması gibi alanlarda sağlık kurumlarına önemli katkılar sağlayabilmektedir (Farboudi,2009)

İnsanın sinir sistemini kontrol eden, tüm bilgilerimizi saklayan, zeka, sevgi ve üzüntü gibi duygusal işlevleri olan, düşünen ve karar veren bir organ olduğu göz önüne alınırsa, beynin insan hayatındaki önemi ancak kavranır. Beyin organı konusundaki çalışmalar yüzyıllardır devam etse de, beyin organının yapısı gereği bu konu, doktorları ve bilim adamlarını zorlayan birçok belirsizlik ve bilinmeyeni barındıran bir araştırma alanı olmuştur.(Demir,2008)

İnsan beyninin çalışırken bir sinyal ürettiği bilinmektedir. Yaklaşık olarak 150 yıldır araştırılan, ispatlanmış ve günümüzde kayıt ve analiz edilebilen sinyallerdir. Bu sinyallere kısaca EEG veya tıp alanındaki ismiyle Elektroansefalogram denilir (Demir,2008). İlk defa 1929 yılında, Hans Berger, kafaya yerleştirilen elektrotlar ve bunlara bağlı bir galvanometre yardımıyla insan beyninde elektriksel aktivitenin varlığını ortaya koydu. 1934' de Adrian ve Matthews, elektrotlarla alınan EEG işaretlerini kuvvetlendirip kaydedilmesini sağladılar. 1939' da P. A. Davis, uyanık bir insanın EEG' sinde, sese karşı uyarılan yanıtların(EP) olduğunu buldu. İkinci dünya savaşından sonra, elektronik ve bilgisayarlar alanlarında meydana gelen büyük gelişmelerin olumlu sonuçları, bu alanda da kendisini gösterdi ve EEG' nin klinik uygulamaları geliştirilerek tanımı üzerinde yapılan araştırmalar olumlu sonuçlar verdi.(Batar,2005)

Bu sinyaller beyin organının aktivitesine göre şekil değiştirmektedir. Beyin organının üretmiş olduğu sinyallerin frekansları da belirlenmiştir, bu frekans değerleri ölçülebilmekte ve analiz edilebilmektedir. Elektronik, bilgisayar ve sinyal işleme yöntemlerindeki gelişmeler tıp elektroniği konusunda son zamanlarda kullanımı oldukça yaygınlaşmaktadır. Bu gelişmeler, doktorlara teşhis ve tedavi de kolaylık sağlamaktadır. (Demir,2008)

EEG işaretlerinin frekans bileşenleri oldukça önemli olduğu gibi, farklı bölgelerden alınan benzer EEG işaretleri arasındaki faz ilişkileri oldukça ilgi çekicidir. Bu tip bilgiler, EEG' nin kaynağının incelenmesinde çok faydalı olup beyin çalışması ile ilgili ilave bilgiler elde edilmesini sağlar. Kafa üzerinden algılanan EEG' lerin genliği tepeden tepeye 1-100μV ve frekans bandı ise 0.5-100 Hz' dir. Ölçümler beyin üzerinden doğrudan alındığı takdirde, genlik 10 kat daha artar. EEG, EKG ve EMG işaretlerinde olduğu gibi şekil bakımından değil, kapsadığı frekanslara göre

değerlendirilmektedir. EEG işaretleri periyodik değildir; genlik, faz ve frekansları sürekli değişir. Bu nedenle anlamlı bir data elde etmek için ölçümlerin oldukça uzun bir sürede yapılması gerekir. EEG işaretlerinin ana frekansı ile beyin aktivitesi yakından ilişkilidir. Aktivite ile frekans birlikte yükselir. (Batar,2005)

EEG işareti geniş bir frekans bandına sahip olmakla birlikte, klinik ve fizyolojik ilgi 0.5 ile 30 Hz arasına yoğunlaşmıştır. Bu frekans aralığı belli frekans bantlarına ayrılmıştır. Bu yaklaşım EEG spektrumunun temelde delta (0.5-4 Hz), teta (4-8 Hz), alfa (8-13 Hz) ve beta (13-30 Hz) şeklinde dört frekans bandı içerdiği gözlemlerine dayanmaktadır. (Albayrak, 2008)

Sinyal işlemecilerin en çok bildiği Fourier analizi bir sinyali zaman alanından frekans alanına dönüştürüp frekans bileşenleri önemli olduğu zaman sıkça başvurulan bir tekniktir. Fourier analizinde, frekans alanına geçildiğinde zaman alanı kaybolur. Durağan sinyallerde fourier analizi iyi bir yöntem olmaktadır, fakat EEG işaretleri durağan olmayıp sürekli değişir. Bu işaretlerde arada sırda kısa süreli ortaya çıkan dik darbeler ve kompleks dalgalar teşhis açısından önemli bilgiler taşır. Bu durumda bu özel spektral bileşenlerin hangi zaman aralığında meydana geldiği önemli olabilir ve fourier analizi yetersiz kalır (Batar,2005).

Dalgacık analizi, Dalgacıklar durağan veya durağan olmayan sinyallerin zaman-ölçek analizi için olanak sağlar. Dalgacıklar sonlu sürelidirler bu yüzden yerel sinyal özelliklerinin analizini mümkün kılarlar. Dalgacık dönüşümleri tüm sinyal frekans-zaman bilgisini korurlar. Bu sebeplerden dolayı durağan olmayan gerçek doğal sinyallerin dalgacık temelli metotlarla işlenmesi geleneksel metotlardan daha iyi sonuçlar sağlarlar. Dalgacığın en önemli avantajlarından birisi, büyük işaretlerin yöresel analizine olanak tanımasıdır. (Batar,2005)

En önemli organımız olan beyinde oluşabilecek hastalıklar doğal olarak insanın hayatını direk ve çok olumsuz etkilemektedir. Dolayısıyla beyin ile ilgili hastalıklar öncelikli hale gelmiş ve iki yüzyıla yakın zamandır bilim dünyasında yoğun araştırmaların konusu olmuştur. (Demir,2008)

Modern tıbbın gelişimi ile özellikle nörolojik kökenli rahatsızlıkların tanı ve tedavisinin izlenmesinde kullanılan EEG teknikleri de benzer bir gelişim göstermiştir. Özellikle

Epilepsi(Sara) hastalığının teşhisi konulmada EEG'nin yararı yadsınamaz. Epilepsi nöbet halinde kendini belirten bir hastalıktır. Nöbetler, beyindeki elektriksel dengenin aniden geçici olarak bozulmasıyla meydana gelmektedir. Nöbetlerin şekli kişiden kişiye değişmektedir. Bazıları yere düşüp kollarıyla ve bacaklarıyla çırpınarak bunu belli ederken, bazıları vücutlarında tuhaf karıncalanmalar hisseder, garip sesler duyar veya kısa bir süre için etraflarına dik dik bakarlar. Bu durum epilepsinin çeşitli nedenlerden ortaya çıkmasından kaynaklanmaktadır.(Albayrak,2008)

Epilepsiden ancak, bir yıl içinde birden fazla epileptik nöbet geçirildikten sonra bahsedilebilmektedir. Sadece bir tek nöbet geçirilmesi halinde, epilepsi var anlamı çıkarılmamalıdır. Tedaviye ise ancak, hastaya kesin 'epilepsi' teşhisi konulduktan sonra geçilebilir.(Albayrak,2008) Burada Epilepsiyi doğru ve zamanında tespit etmenin önemi ortaya çıkıyor.

Çalışmada kullanılacak olan diğer bir yöntem ise , verideki trendleri, ilişkileri ve profilleri belirlemek için veriyi sınıflandıran bir analitik araç ve bilgisayar yazılım paketi olarak tanımlanan veri madenciliğidir. Veri madenciliğinde eldeki veri türüne ve elde edilen sonuçların kullanım amacına göre farklı birçok teknik mevcuttur. Tez çalışmasında sınıflandırma yöntemi kullanılmıştır. Sınıflandırma yönteminde temel prensip, veri içerdiği özelliğe göre önceden belirlenmiş sınıflardan birisine dâhil edilir. Bir sınıflandırma yönteminin kalitesi bu prensibi sağlaması ile doğu orantılıdır. Sınıflandırma yapabilmek için çok sayıda yöntem geliştirilmiştir. Karar ağaçları bunların arasında yer almaktadır. Karar ağacı oluşturulduktan sonra bu ağaçtan karar kuralları türetilir ve test verisi üzerinde denenir. Olumlu sonuç elde edilirse yeni gözlemleri sınıflandırmak için bu kurallar kullanılır. (Farboudi,2009) Özellikle Bayes temelli algoritmalar(Naive Bayes, BayesNet vb.) tıbbi verilerde yüksek başarı oranlarıyla dikkat çekmektedirler.

Türkiye'de yapılan lisansüstü tez çalışmaları Haziran 2012 itibariyle incelendiğinde, veri madenciliği ile ilgili toplam 165 adet çalışma bulunduğu göze çarpmaktadır. Bu çalışmaların büyük bir bölümü işletme, ekonomi, istatistik, bilgisayar bilimleri anabilim dallarında ilgili olduğu görülmektedir. Tıbbi veriler kullanılarak yapılmış veri madenciliği süreci içeren çalışmaların çok az olması sebebiyle tez konusu

olarak eeg sinyallerindeki epileptiform aktiviteye veri madenciliği yöntemlerinin uygulanması seçilmiştir.

Bu çalışmanın amacı; tıp bilişiminin tıbbi sinyaller ve bu sinyallerin analizi ile veri madenciliği alanının sınıflandırma metotlarını içeren böylece iki alanı da kapsayan bir çalışma ortaya koymaktır. Çalışmada 25-30 dakikalık 19 kanallı EEG cihazlarından 25'i Epilepsi(Sara) hastası, 25'i sağlıklı olan 50 kişiden alınan edf formatındaki elektroansefalogram verileri, gerçekleştirilen yazılım sayesinde ayrık dalgacık dönüşümü tabi tutularak her kanal altbandlarına (δ, Θ, α, β) ayrılmıştır. Her kanalın altbandlarının ortalamaları ve hastalık durumunu içeren veriler öznitelik parametreleri olarak seçilmiştir. Bu öznitelik değerlerine veri madenciliği sınıflama algoritmaları uygulanarak; epileptik aktivite tespit edilmeye çalışılmıştır. En yüksek doğruluk oranını verdiği tespit edilen algoritmalar yardımı ile bu alanda çalışan uzmanların epileptik aktivite teşhisi koymalarına kolaylaştıracak ve konulan teşhislerde doğruluk oranını yükseltecek bir model oluşturulması hedeflenmiştir.

Yedi bölümden oluşan bu tez aşağıdaki şekilde düzenlenmiştir:

Giriş bölümünde, epilepsi(sara)hastalığı ve bu hastalığın tespitinde kullanılan elektroansefalografi (EEG) sinyallerinin analiz yöntemlerinden bahsedilmiş ve buradan elde edilecek veriler üzerinde veri madenciliği kullanılarak yapılacak çalışmanın amacı anlatılmıştır.

İkinci bölümde, tespitine çalışılan epilepsi(sara)hastalığının tanımı, nedenleri ve epilepsi nöbet türleri hakkında genel bilgiler verilmiştir.

Üçüncü bölümde, elektroansefalografi (EEG)'nin ne olduğu, EEG kaydının nasıl gerçekleştirildiği ve EEG'yi tanımlayan nitelikler anlatılmıştır. Ayrıca normal ve anormal EEG patternlerinin özellikleri üzerinde durulmuştur.

Dördüncü bölümde, elektroansefalografi (EEG) sinyallerinin analizinde kullanılan klasik sinyal analizi metotları (Fourier analizi ve türleri) ile modern sinyal analizi metotları (Dalgacık dönüşümü ve türleri) farklarına ve özelliklerine değinilmiştir.

Beşinci bölümde, veri madenciliği hakkında geniş bilgiler verilmiş ve veri madenciliği yazılımları, veri madenciliği süreci ile veri madenciliğinin genel kabul görmüş modellerinden bahsedilmiştir. Ayrıca çalışmanın konusu olan tıbbi verilerle veri madenciliği konusu üzerinde durulmuştur.

Altıncı bölümde de gerçekleştirilen çalışmada kullanılan veriler, gerçekleştirilen yazılım ile bu verilerden öznitelik değerlerinin elde edilmesi ve bu veriler üzerine veri madenciliği sınıflandırma algoritmaları kullanılarak oluşan sonuçların değerlendirilmesi yer alıyor.

Tezin sonuç ve öneriler kısmında ise, yapılan katkıları anlatılarak, gelecekteki çalışmalar için önerilerde bulunulmuştur.

BÖLÜM 2.

EPİLEPSİ (SARA) HASTALIĞI

2.1. Epilepsi (Sara) Nedir?

Duyusal ve bilişsel merkezimiz olan beynimiz, bizi diğer canlılardan ayıran en önemli organımızdır. 100 milyar hücreden oluşan beyin herhangi bir bilgisayarla karşılaştırılmayacak kadar karmaşık ve üstün bir sisteme sahiptir.(ETD,2012a)

Bilim ve teknoloji alanında atılan önemli adımlara rağmen insan beyni hala birçok sır barındırmaktadır.(ETD,2012a)Bu sırlardan birisi olan ve halk arasında "sara hastalığı" olarak da bilinen epilepsi, beyindeki hücrelerin kontrol edilemeyen, ani, aşırı ve anormal deşarjlarına bağlı olarak ortaya çıkan bir durumdur. (Epilepsi Rehberi,2012)

"Epilepsi" terimi temel olarak Yunanca "nöbete yakalanmak" anlamına gelen ("epilambenein") kelimesinden türetilmiştir. Diğer bir deyişle "epilepsi", "nöbet" veya "kasılma hastalığı" dır. Bununla beraber farklı türlerde kasılma hastalıkları bulunduğu için epilepsi (sara) olarak adlandırılması daha yerinde olacaktır. (Epilepsi Rehberi,2012)

Basitleştirilmiş şekliyle epilepsi nöbeti kısa süreli beyin fonksiyon bozukluğuna bağlıdır. Beyin, insan vücudunun ana kumanda merkezi gibidir. Beyin hücreleri arasındaki uyumlu çalışma, elektriksel sinyallerle sağlanır. Nöbetin nedeni, bir tür beklenmeyen elektriksel uyarı olarak düşünülebilir. Kısaca; epileptik nöbet beynin kuvvetli ve ani elektriksel boşalımı sonucu oluşan kısa süreli ve geçici bir durumdur.(Incaklinik,2012)

Epilepsi, dünyanın her bölgesinde, erkek ve kadında, her türlü ırkta ve yaklaşık 100 kişide bir oranında görülebilen bir hastalıktır. Epilepsi hastalığının toplumdaki yıllık eklenen hasta oranı 45/100000 civarındadır. Ayrıca yaklaşık 20 kişiden birinde yaşam boyunca bir kez nöbet görülebilir ve bu kişilerde daha sonra nöbet tekrarlamayabilir.(Incaklinik,2012)

Herkese bazı belirli durumlarda örneğin; elektrik şokuna, oksijen yetmezliğine veya ciddi bir baş yaralanmasına bağlı olarak epileptik bir nöbet geçirebilir. Buna karşın, epilepsi nöbetleri belirgin hiçbir nedene bağlı olmaksızın, normalde tamamı ile sağlıklı olan bir kişide görüldüğünde ve tekrarlanmaya meyil ettiğinde epilepsi olarak adlandırılmaktadır. Böyle bir durum epilepsinin kronik, uzun süreli bir hastalık olduğunu gösterir (buna karşın, bu durum epilepsinin yaşam boyu süreceği anlamına gelmez). (Epilepsi Rehberi,2012)

Epilepsi beyinde başlayan bir hastalıktır. Zihinsel bozukluklarla hiçbir ilgisi yoktur ve istisna sayılabilecek vakalar haricinde zeka gerilemesine yol açmaz. Diğer bir çok hastalıkta olduğu gibi epilepside de benign (iyi huylu) ve malign (kötü huylu) türler bulunmaktadır. Epilepsi tedavi edilebilir bir hastalıktır ve bir çok vakada hastalık tamamen iyileşebilir.(Epilepsi Rehberi,2012)

Nöbetin nedeni tümör yada başka bir hastalık değilse, epilepsinin ilerlemesi söz konusu değildir, bazen yaşla birlikte nöbet sıklığı da azalabilir. (Incaklinik,2012)

Epilepsi nöbetleri, çoğu zaman insana çok uzun sürüyor gibi gelse de 1-3 dakika içinde kasılmalar biter ve hastalar belli bir süre sonra nöbet öncesindeki normal aktivitelerini kazanırlar. (Incaklinik,2012)

Epilepsi bulaşıcı bir hastalık değildir. Epilepsili kişinin hastalığının başkaları tarafından bilinmemesi için bir neden yoktur. Yakın arkadaşlarınız, akraba ve komşularınız, öğretmeniniz hastalığınız hakkında bilgi sahibi olmalıdır. Epileptik kişi nöbetleri iyi kontrol edildiğinde, iyi yaşayabilir, eğlenebilir, çalışabilir ve toplumun aktif bireyi olabilir.(Incaklinik,2012)

Epilepsili kişi evlenebilir ve çocuk sahibi olabilir. Epilepsili kadınlara planlı gebelik önerilir, bu nedenle gebe kalmadan önce mutlaka doktoru ile görüşmelidir.

Tedavide kullanılan ilaçların çocuk üzerine değişik etkileri nedeniyle; gebelik öncesi nöbetlerin tipine ve durumuna bakılarak uygun ilaç ve dozu doktor tarafından düzenlenmelidir.(Incaklinik,2012)

2.2. Epilepsi' nin Tarihçesi

Epileptik nöbetleri ve epilepsiyi tedavi etme denemeleri, insanlığın ilk çağlarına kadar uzanır. Tarihin her devrinde epilepsinin sebebi hakkındaki yorumlara dayanılarak tedaviler uygulanırdı.(ETD,2012b)

Hipokrat'tan önceki zamanlarda; "kutsal hastalığın" tanrılar tarafından gönderildiğine inanılır, bağışlar yapılır, kurbanlar kesilir, papazların ve doktorların yönetiminde, ibadethanelerde dini ayinler yapılırdı.(ETD,2012b)

Hipokrat ve O'nun izinden giden hekimler; epilepsinin tıbbi sebeplerine inanıyorlar "Humoralpathologie: Sâfte- Lehre", hastalıkların hepsinin, kan ve vücut ifrazatlarının hatalı bir karışımından ileri geldiğine kabulleniyorlar ve tedaviyi de tabiî bir temele dayandırıyorlardı. Bu tarz bir tedavinin temelinde ise diyet,yani akıllıca düzenli bir yaşam tarzı yatıyordu.(ETD,2012b)

Hıristiyanlığın orta çağlarında, epilepsinin ; tıbbi sebeplerden kaynaklandığı unutulmuş. Hastalığın sebepleri daha ziyade; şeytana, kötü ruhlara, izafe edilmiş (cin hastalığı), tedavide ona göre değişmişti. İbadet, oruç, kurban, kutsal yerlerin (türbe) ziyareti, "Eksorzism" ruh çıkarma gibi usuller, tedavinin ana unsurlarını oluşturuyordu. Azizlerden, doğrudan doğruya yardım istenmiş veya onların aracılığı ile tanrıdan yardım dilenilmiştir. Çok sayıda kutsal sayılan objeler "Devotionalien" epilepsiye karşı kullanılmıştır.(ETD,2012b)

Hıristiyanlık inancından kaynaklanan tedavi gayretleri yanı sıra, batıl itikatlara dayanan tedavi yöntemleri de vardı.Bu yöntemler,bu güne kadar uygulana gelmiştir (Günah çıkarmak, büyü yapmak, "fetişizm"hamaylı taşımak) .(ETD,2012b)

Orta çağın sonlarında, rönesans devrinde; nebatı ilâçların yanı sıra, kimyevi olarak tanınan maddelerde sara tedavisinde kullanılmıştır. Bunlardan en tanınmışları: Bakır (ilk çağlarda da kullanılıyordu), çinko oksit, gümüş nitrat, cıva , bizmut ve kalaydır.(ETD,2012b)

Bütün bu nebati ve madeni maddelerin, günümüzde; epilepsi tedavisinde hiçbir etkisinin olmadığı bilinmektedir. 19. Yüzyılın ortalarından beri, tıbbî ilimler ve tıbbî bilgilerin ışığında, epilepsiye iyi gelen ilâçlar bulundu. Bunlardan ilk iki tanesi, Brom (1857) ve Phenobarbital (1912) dir. Epilepsiye karşı etkili ilâçlar olarak, günümüzde de kullanılmaktadır.(ETD,2012b)

Bugün ki tıpta, epilepsiye karşı yalnız başına ve kombine olarak, başarı ile kullanılan 20 ye yakın kimyevi madde sayesinde, hastaların %60 ı nöbetlerden kurtarılabiliyor,%20 si ise bariz bir iyileşme gösteriyor. Epilepsili hastaların beşte birinde ise modern epilepsi ilaçlarının hiçbir yararı olmuyor. Tedaviye cevap vermeyen hastalara, modern epilepsi cerrahisi fayda sağlayabiliyor.(ETD,2012b)

2.3. Epilepsi Olan Ünlüler

Epilepsinin, akıl hastalıkları veya zekâ geriliği ile hiçbir ilgisi yoktur. Epilepsi nöbetleri geçirmelerine rağmen, epilepsi hastaları dâhiyane işler becermişlerdir. (ETD,2012b)

Aşağıdaki listede isimleri geçen bu değerli insanlar ya bütün hayatları boyunca veya hayatlarının belli bir döneminde kronik epilepsiye yakalanmışlardır. Tanınmış, “prominent epilepsi” hastaları arasında, hanımların erkeklere oranla daha az temsil edilmelerinin nedeni hanımların eski yüzyıllarda yüksek mevkilere daha az gemleri dolayısıyla daha az tanınmalarından kaynaklanmaktadır. .(ETD,2012b)

Bunlar arasında Vincent van Gogh (Hollandalı ressam), G. Julius Caesar(Romalı devlet adamı), Gustave Flaubert (Fransız yazar), F.M. Dostoyevsky (Rus yazar), Heracles, Napoleon Bonaparte (Fransa imparatoru), Alfred Nobel (İsveçli kimyager),

Lord Byron(İngiliz şair) , Büyük İskender (Makedonya Kralı), Hermann von Helmholtz (Alman fizikçi), Moliere (Fransız yazar), Vladimir Ilyich Lenin(Rus devrimci), Socrates (Yunan Filozof) gibi ünlü isimler mevcuttur.(ETD,2012b)

2.4. Epilepsi'ye Neden Olan Etkenler

1) Beyinde oluşan tümörler
2) İskemik lezyon: Beyne giden kan akımı azaldığında (iskemi), beyin dokusundaki besin maddeleri ve oksijen azalması
3) Konjentinal mal formasyon: Doğuştan gelen bozukluklar.
4) Gebelik döneminde annenin ilaç ve alkol alımı, bebeğin gelişimini etkileyecek mikrobik hastalıklar
5) Doğum sırasında oluşabilecek beyin zedelenmesi, kanaması, beynin oksijensiz kalması
6) Doğum sonrası menenjit, beyin iltihabı gibi rahatsızlıklar
7) Febril konvulziyon: Ateşe bağlı istem dışı şiddetli kasılmalar.
8) Enfeksiyon: Tüm vücudu etkileyen ya da şiddetli olan enfeksiyonlar
9) Troid bezi hastalıkları
10) Özellikle B6 vitamini eksikliği gibi nedenlerle epilepsi oluşabildiği saptanmıştır.(Wikipedia,2012a)

2.5. Epilepsiyi Tetikleyen Etkenler

Epilepsi hastası bir kişinin nöbet geçirmesinin nedeni genellikle pek belli değildir. Fakat aşağıda sayacağımız durumlar neticesinde kişilerin epilepsi nöbeti geçirdiği gözlemlenmiştir:

1) Aşırı alkol tüketiminden sonra,
2) Fazla kafein alımı sonucunda uyku düzensizliklerinde,
3) Ağır diyetler sırasında,
4) Yüksek ateşli hastalık geçirilmesi sırasında,
5) Far ışıkları, eğlence mekânlarındaki güçlü ışık etkenleri sebebiyle,

6) Şiddetli kafa travması geçirilmesi sonucunda,
7) Kadınlarda hormonal değişiklikler sırasında,
8) Yetersiz uyku uyunması neticesinde,
9) Çalışılan ortamın stresli olması gibi sebepler sayılabilir.(Wikipedia,2012b)

2.6. Kalıtımsal (Genetik) faktörler

Çok genel sözcüklerle anlatmak gerekirse, bir çok hastalığın gelişmesine yol açan iki temel koşul vardır: Birincisi kişinin kendinde bulunan bir eğilim ve ikincisi ise dış kaynaklı veya sonradan edinilmiş hastalığı ortaya çıkaran bir sebeptir. Kişide bulunan eğilimin özellikle önemli rol oynadığı bazı hastalıklar vardır (örneğin şeker hastalığı veya gut hastalığı). Öte yandan, diğer bazı hastalıkların ortaya çıkışında ise belirli tetikleyici dış (edinilmiş) nedenler asıl sebeptir(örneğin; akciğer kanseri). (Epilepsi Rehberi,2012)

Bir hastalığa olan eğilimin boyutu kalıtımla (genlerle) ilgilidir ve anne babadan çocuklara geçer; Bu yüzden burada genetik veya ailevi bir eğilimin varlığından bahsederiz. Örneğin; bazı ailelerin şeker hastalığına diğerlerinden daha fazla eğilimi vardır fakat yine de şeker hastalığı kalıtsal bir hastalık değildir. Ancak hastalığın kendisi kalıtımın belirli kurallarına göre çocuğa aktarılmışsa burada kalıtsal hastalıklardan bahsederiz. (Epilepsi Rehberi,2012)

Epilepsi kalıtsal bir hastalık değildir, epilepsi hastalığı anne-babadan çocuğa geçemez. Bununla beraber epilepsiye olan eğilimin boyutu aileden aileye değişir. Bu yüzdendir ki, her ne kadar hastalığın ortaya çıkması için eğilimin bir dış etkenle harekete geçirilmesi gerekse de bir ailenin birkaç üyesi epilepsi nöbetleri geçirebilir veya kronik epilepsi geliştirebilir. (Epilepsi Rehberi,2012)

Tabii ki hastalığın ortaya çıkışında bir dış etkenden çok, kalıtımın hayati rol oynadığı epilepsi türleri de vardır. Bu tür epilepsi "birincil olarak genetik bağlantılı epilepsi" olarak adlandırılır. Bunun için kullanılan tıbbi terim ise "idiopatik epilepsi"dir. Çocuklukta bunun örneği, sık boşluk nöbetleri(pignolepsi)ve jacksonian veya rolandik epilepsidir. (Epilepsi Rehberi,2012)

Bununla beraber, bir çok epilepsi türünün gelişmesinde dış (edinilmiş) faktörler genetik eğilimden çok daha önemli bir rol oynar. Bu tür epilepsilere ise "semptomatik epilepsiler " denir. (Epilepsi Rehberi,2012)

2.7. Nöbet Türleri

Nöbetlerin, çok karmaşık olmasına karşın, genellikle iki kategoride sınıflanması uluslararası düzeyde kabul görmektedir: Birincisi kısmi (parsiyal ya da fokal) ikincisi ise, jeneralize (genel, yaygın) nöbet olarak adlandırılmaktadır. Kısmi ya da fokal nöbetler beynin bir bölümünde başlar [beyindeki herhangi bir odakta] ve vücudun, beynin o bölümü tarafında kontrol edilen kısımlarını etkiler. Jeneralize nöbetler tüm beyne yayılır ve bu nedenle tüm vücudu etkiler.(Epilepsy Center,2012)

2.7.1. Kısmi nöbetler

Epilepsiye neden olan bozulma, beynin belli bir kısmında (dilimde) ortaya çıkmaktadır. Bütün nöbet belirtileri çok çeşitli olup, bozulmanın oraya çıktığı beyin kısmına bağlıdır. Bazı nöbetlerde insanın şuuru açık kalmakta, bazılarında ise kaybolmaktadır. Kısmi nöbetler, basit, karmaşık ve sekonder jeneralize nöbetler olarak üç kısma ayrılmaktadır. Sayılanlar arasındaki son nöbet türü, bir parsiyel nöbetin jeneralize nöbete geçebilmesi anlamına gelmektedir. (Nationaal Epilepsie Fonds,2012).

2.7.1.1. Basit kısmi nöbetler

Basit kısmi nöbetler beynin yalnızca bir bölümünü etkileyen lokalize nöbetlerdir. Kişide görülen belirtiler, beynin o bölümünün kontrol ettiği işleve bağlı olacaktır. Nöbet, istem dışı hareketleri ya da kol ve bacaklarda sertliği, daha önce yaşamıştım hissini (deja vu), hoş olmayan kokular duyma ya da tatlar hissetme veya midede titreme, heyecan ya da bulantı gibi belirtileri kapsayabilir. Nöbet sırasında kişinin bilinci tamamen yerindedir. Nöbet genellikle bir dakikadan daha kısa sürer ve kişi daha sonra iyileşir. Kısmi nöbet yayıldığında ve tüm beyni kapsadığında ikincil jeneralize nöbet olarak adlandırılır.(Epilepsy Center,2012)

2.2.1.2. Karmaşık kısmi nöbetler

Bu tür nöbet beynin yalnızca bir bölümünü etkilemektedir. Ancak, kişinin bilinç durumu değişmektedir. Kişide genellikle bilinç bulanıklığı ve şaşkınlık görünebilir ve kişi giysileri ile sinirli sinirli oynama, sakız çiğneme hareketleri ya da tuhaf sesler çıkarma gibi garip hareketler yapabilir. Nöbet genellikle bir iki dakika sürer. Ancak kişide nöbetten sonra birkaç dakika ile birkaç saat arasında bilinç bulanıklığı ve uyuşukluk olabilir.(Epilepsy Center,2012)

2.7.2. Dalma (Absans) nöbeti

Bu, tüm beyne yayılan jeneralize bir nöbet olup, çocuklarda daha yaygın olarak görülür. Bu nöbette kişi çevresinden haberdar değildir, ancak nadiren yere düşer. Gözlerini dikip sabit bakar, gözleriyle aranır ya da göz kapaklarını kırpıştırır.(Epilepsy Center,2012)

Dalma (absans) nöbetleri ile hayale dalma arasındaki farkı anlamak zor olabilir. Ancak, dalma nöbetleri ani başlar, birkaç saniye sürer, aniden sona erer ve kişi nöbet önceki durumuna geri döner. Bu nöbetler birkaç saniye sürmesine karşın, günde birçok defa yinelenebilir ve bu da sonuçta öğrenmeyi olumsuz etkiler.(Epilepsy Center,2012)

2.7.3. Miyoklonik nöbetler

Miyoklonik (sıçrama) nöbetleri kaslarda kısa süreli ani kontrol dışı kasılmalar ile kendini gösterir. Bu nöbetler genellikle uyandıktan kısa bir süre sonra ya da kişi yorgun olduğunda yatarken meydana gelir. Jeneralize bir nöbet olduğundan bilinç kaybı oluşur. Ancak, bu çok kısa süreli ve zor farkedilen bir bilinç kaybıdır.(Epilepsy Center,2012)

2.7.4. Tonik-klonik (grand mal) nöbetler

Tonik-klonik nöbetler tüm beyne yayılan jeneralize nöbetlerdir. Bu nöbet türü, epilepsi denilince çoğu kişinin aklına gelen nöbet türüdür. Bazı kişilerde, nöbet başlamadan hemen önce daha önce yaşamıştım hissi, midede tuhaf bir duygu ya da tuhaf tat ve koku gibi ön belirtiler (aura) görülebilir. Bu ön belirtilerin (auranın) kendisi basit kısmi bir nöbettir.(Epilepsy Center,2012)

Bir tonik-klonik nöbet sırasında kişi kaskatı kesilir ve yere düşer [tonik aşama]. Daha sonra kol ve bacakları güçlü, simetrik ve ritmik hareketler halinde kasılmaya başlar [klonik aşama]. Ağzından köpük gelir, yüzü morarır veya kızarır ya da idrar ve/veya dışkısını kaçırır.(Epilepsy Center,2012)

Bu nöbet türünü izlemek ürkütücü olmasına karşın, nöbetin kendisinin nöbet geçiren kişiye zarar verme olasılığı yoktur. Bununla beraber, kişi kusabilir veya dilini ısırabilir ve bazen yere düşerken veya kasılma sırasında çevredeki nesnelere çarparak yaralanabilir. Nöbet normal olarak birkaç dakika sonra sona erer. Bu sırada kişide genellikle bilinç bulanıklığı ve uyuşukluk görülür. Kişi, baş ağrısından yakınabilir ve uyumak isteyebilir. Bu uyuşukluk hissi birkaç saat sürebilir. .(Epilepsy Center,2012)

2.7.4.1. Tonik nöbetler

Tonik nöbetler kasların kaskatı kesilmesine neden olan jeneralize nöbetlerdir. Kişi eğer ayaktaysa, oldukça şiddetli bir şekilde yere düşecektir. Bu nöbetler uykudayken ardısıra nöbetler (küme şeklinde) halinde oluşabilir. Ancak, kişi uyanıkken oluştuğunda sıklıkla baş yaralanmalarına neden olabilir. Eğer uygunsa, kişinin baş yaralanmasından kaçınmak için koruyucu kask takması tavsiye edilmektedir. Kişinin yaralanması halinde tıbbı yardım isteyiniz.(Epilepsy Center,2012)

2.7.4.2. Atonik nöbetler

Atonik nöbetler, kişinin kaslarını gevşeterek yere yıkılmasına neden olan jeneralize nöbetlerdir. Nöbetler çoğunlukla 'düşme atakları' ya da astatik nöbetler olarak adlandırılır. Bu nöbetler baş veya yüz yaralanmalarına yol açabilir. Sürekli yaralanmadan kaçınmak için başa koruyucu kask takılması tavsiye edilir. Nöbet sonrası iyileşme genellikle oldukça hızlıdır. Kişinin yaralanması halinde tıbbı yardım isteyiniz.(Epilepsy Center,2012)

BÖLÜM 3

ELEKTROENSEFALOGRAFİ (EEG)

3.1. EEG Nedir?

Elektroensefalografi (EEG), saçlı deriye yerleştirilen elektrotlar aracılığıyla kaydedilen serebral biyoelektriksel aktivitedir. İlk kez 1929 yılında Hans Berger tarafından insan beyninden elektriksel aktivite kaydı yapılmasından günümüze kadar, PET, fMR ve magnetoensefalografiye (MEG) rağmen halen serebral biyoelektriksel aktivitenin incelenmesinde geçerliliğini korumaktadır.(Bora vd.,2008)

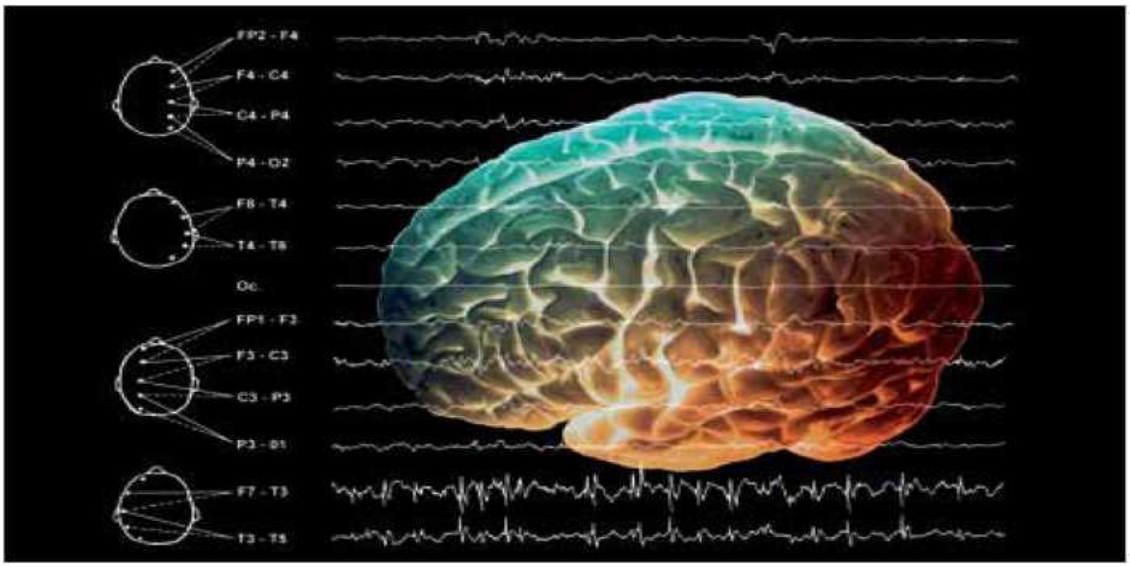
EEG, epilepsi tanısının konulmasında, nöbet sınıflamasında ve hastaların takibinde kullanılan en önemli laboratuvar yöntemidir. Nöbet öyküsü olan hastanın EEG'sinde "epileptiform aktivite" saptanması, benzer aktivitenin hiç nöbet geçirmemiş birçok bireyde de görülebilmesi nedeniyle, tanıyı şüphenin ötesine götürmez ve EEG'nin normal olması da epilepsi tanısını dışlamaz. Kesinleşmiş epilepsi tanısı olan hastalarda EEG bulguları, hastalığı sınıflandırma, fokal veya lateralize bir epileptik odak olduğunu tanımlama, uygun tedaviyi seçme, prognoz için yol gösterici olarak ve hastalığın gidişini izlemek için kullanılır.(Bora vd.,2008)

EEG; diğer nörolojik hastalıkların araştırılmasında da kullanılabilir. Bu hastalıklarda spesifik olmayan, tanı koymaya yardımcı ve tanıyı destekleyen EEG anormallikleri saptanabilir. Herpes simpleks ensefaliti tanısını düşündüren tekrarlayıcı yavaş dalga kompleksleri buna iyi bir örnek olup yavaş virüs enfeksiyonlarında da (subakut sklerozan panensefalit-SSPE, Creutzfeldt-Jacobs hastalığı gibi) tanı konulmasına katkı sağlar. Hastanın klinik durumundaki iyileşme veya kötüleşmeyi, objektif olarak değerlendirebilmek için metabolik bozukluğu olan hastaların

izlenmesinde; men- tal durum veya bilinç değişikliğinin, uyku ve bozuklukları ile beyin ölümünün değerlendirilmesinde de EEG kullanılır.(Bora vd.,2008)

3.2. EEG Aktivitesinin Fizyolojik Temeli

Saçlı deriden kaydedilen EEG aktivitesinin kaynağı, kortikal piramidal hücrelerin postsinaptik potansiyelleridir. Bu postsinaptik potansiyeller, hücrenin içi ve dışı arasındaki elektriksel potansiyel farkından oluşurlar ve kortekste toplanarak beyni saran yapılardan saçlı deriye yayılırlar. Elektriksel potansiyel değişikliklerinin toplanması, korteksin vertikal olarak dizilmiş geniş piramidal nöronlarında gerçekleşir. Sub- kortikal yapılar veya horizontal yerleşimli kortikal hücrelerden kaynaklanan potansiyellerin, skalpten alman EEG'ye katkısı çok azdır veya hiç yoktur. Sulkal derinliklerde ortaya çıkan potansiyellerin skalp üzerinden kaydedilmesi yüzeyde olanlara göre daha güçtür.(Bora vd.,2008)



Şekil 1. Kafatasının saçlı derisine yerleştirilen elektrotlardan alınan sinyaller

Sinaptik aktivitenin eksitatör ya da inhibitör olmasına göre postsinaptik membran depolarize ya da hiperpolarize olur. Eksitatör postsinaptik potansiyeller (EPSP), hücre içine pozitif yüklü iyonların akımını sağlayarak hücrenin depolarizasyonuna neden olur. Sodyum ve klor hücre dışında daha yoğun iken, potasyum hücre içinde yoğunluk gösterir. Bu eşit olmayan dağılım, istirahat elektriksel

potansiyelinde anlamlı bir fark oluşturur. İnhibitör postsinaptik potansiyeller (İPSP) ise hücre dışına net iyon akımını artırarak hücrenin hiperpolarizasyonunu sağlar. Binlerce, hatta milyonlarca nöronun net EPSP ve İPSP'lerinin toplamının oluşturduğu elektriksel akım, yüzeysel EEG aktivitesini oluşturmaktadır.(Bora vd.,2008)

İnsan EEG'si istirahat ve uyanıklık halinde, ritmik bir aktivite gösterir. EEG aktivitesini yaratan jeneratörlerin yerleşimi hakkında farklı görüşler vardır. Piramidal hücrelerin tekrarlayıcı postsinaptik potansiyellerinin kortikal ya da subkortikal jeneratörlerden kaynaklandığı konusunda kesin fikir birliği yoktur. Korteksin spontan ritmler oluşturan nöron ağına sahip olduğu bilinmektedir. Öte yandan talamokortikal projeksiyonlar yoluyla talamik ve beyin sapı yapılarının bu ağı kontrol ettiği de bilinmektedir. Retiküler ağ ve talamik nukleuslardan kaynaklanan ritmik boşalımlar, kortikal potansiyelleri uyarmakta ya da desenkronize edebilmektedir. Deneysel çalışmalarda; talamus kaynaklı uyku iğciklerinin beyin sapının talamusun alt seviyesinden kesilmesi ile çok az etkilendiği gösterilmiştir. Bu bulgular, ritmik aktivitenin genellikle talamik ritm oluşturan hücrelerden kaynaklandığı görüşünü desteklemektedir.(Bora vd.,2008)

Epileptiform aktivite, epileptojenik odak içindeki nöronlar tarafından senkronize edilen bir grup hücrenin membran potansiyellerinin paroksismal depolarizasyon şiftleri sonucu oluşur. Paroksizmal depolarizasyon şifti, bir nöbet öncesi olayıdır. Paroksizmal depolarizasyon şifti çok hızlı ve çok fazla sayıda aksiyon potansiyeli oluşturur. Hücre içi depolarizasyondan kaynaklanan akım apikal dendrite doğru yönlenir, tüm nöronun somadendritik bölgesini kapsar ve hücre dışında yüzeyel diken dalga aktivitesini oluşturur. Paroksizmal depolarizasyon şiftinin ardından oluşan uzamış hiperpolarizasyon ile impulslar, o nörondan diğerlerine yayılır ve hipersenkronizasyon ortaya çıkar. Tüm bu verilere rağmen, EEG'de karşılaşılan anormal ritmlerin fizyolojik temeli ve nöbet aktivitesinin altında yatan hipersenkronizasyondan sorumlu mekanizmalar hala iyi açıklanamamış durumdadır.(Bora vd.,2008)

3.3. EEG Kaydı

3.3.1. Elektrotlar

EEG skalp üzerine ile yerleştirilen metal elektrotlar kaydedilir. Elektrotlar iletken bir madde ile kaplı olarak kafa derisine yerleştirilir ve başlık veya kafa bandı basıncı ile yerinde tutulur. Alternatif olarak iğne elektrotlar doğrudan skalp içine yerleştirilebilir. Her bir elektrotun yeri; nazion, inion, sağ ve sol preauriküler noktalardan yapılan standart ölçümlerle belirlenerek uluslararası 10-20 sistemine göre yerleştirilir. Bu sistemde yetişkindeki komşu elektrotlar, birbirlerinden 5-7 cm uzaklıktadır. Daha fazla yakınlık gerektiğinde (örneğin; epileptojenik bir odağın bölgesini saptarken), elektrotlar orta pozisyonda yerleştirilebilir. Seçilmiş hastalarda, nöbet sırasında değişik alanların aktivitesini ölçebilecek şekilde de elektrotlar yerleştirilebilir. Uluslararası 10-20 sistemindeki 21 elektrota ek olarak %10 aralıklarla elektrot yerleşimleri kullanılabilmektedir. Bu elektrotların yerleştirilmesi ve isimlendirilmesi "American Electroencephalographic Society" tarafından standardize edilmiştir. Bu öneride 10-20 sistemiyle karşılaştırıldığında dört elektrot farklı bir şekilde isimlendirilmiştir: T_7, T_8, P_7 ve P_8 'dir. Bu elektrotlar şekilde siyah üzerine beyaz yazılı olarak çizilmiştir. Elektrotlar arasındaki potansiyel farklılıklar amplifiye edilir ve hareket eden bir kağıt üzerinde birçok yazıcı tarafından kaydedilir. Günümüzde; gelişmiş dijital sistemler sayesinde EEG kayıtları monitörden izlenebilmekte ve gerekli görülen bölümlerin yazdırılması mümkün olmaktadır. (Bora vd.,2008)

Kayıt yaklaşımları, elektrotlar arasındaki potansiyel farklılık ölçümü (bipolar derivasyon) veya tüm elektrotlar ile bilinen bir referans nokta arasındaki ölçüm (monopolar derivasyon) şeklinde değişebilir. Her bir tekniğin avantajları ve dezavantajları vardır, fakat rutin yaklaşımlar için en azından bu montajlardan ikisinin alındığı EEG trasesi elde edilmelidir. Genellikle montajlar eşdeğer elektrotların sıraları arasında longitudinal veya transvers olarak geçecek şekilde yerleştirilir.(Bora vd.,2008)

Beynin ulaşılamayan bölgelerinin aktivitesini kaydetmek için farklı elektrotlar kullanılabilir. Nazofarengeal elektrot, temporal lobun anteromedyal yüzünün aktivitesinin kaydedilmesini sağlar. Temporal lobun anteroinferior kısmının aktivitesi

sfenoidal elektrotlarla kaydedilebilir. Sfenoidal elektrotlar, nazofarengeal elektrotlara göre artefaktlara daha az yol açmaktadır, ancak uygulanmaları daha zordur. Elektrot steril bir iğne veya uç kısmı hariç izole bir çengelden oluşur ve bir doktor tarafından lokal anestezi altında perkutan olarak sfenoid kemiğe komşu, foramen ovalenin hemen önüne yerleştirilir. KPN'leri olan hastalarda elektrofizyolojik değerlendirmede kullanılan diğer aksesuvar elektrotlar; yüzeysel sfenoid elektrotlar, mini sfenoid elektrotlar ve anterior temporal elektrotlardır.(Bora vd.,2008)

3.3.2. Filtreler

Filtreler, EEG'den gelen görece yüksek ve düşük frekanslardaki dalgaları dışlamak için kullanılır. EEG cihazının 3 çeşit filtresi vardır:

a) Yavaş dalgaların amplitüdünü azaltan (geçişini engelleyen) alçak frekans filtresi

b) Hızlı dalgaların amplitüdünü azaltan (geçişini engelleyen) yüksek frekans filtresi

c) En yaygın elektrik artefaktı olan 50-60 Hz. lik dalgaların amplitüdünü azaltan (geçişini engelleyen) 50-60 Hz filtresi (Notch filtre) (Türkiye'de şehir elektrik akımı 50 Hz, ABD'de 60 Hz. dir)

EEG kayıtları sırasında genellikle, alçak frekans filtresi 0. 5- 1Hz yüksek frekans filtresi 70 Hz olarak kullanılır, ancak gerekli görülen durumlarda EEG'nin daha iyi değerlendirilebilmek için filtre ayarları bu frekansların dışında da değiştirilebilir.(Bora vd.,2008)

3.3.3. Artefaktlar

Artefakt, serebral kaynaklı olmayan sinyallerdir. Elektrotlar, kayıt ekipmanı ve kayıt aletlerinden kaynaklanan birçok biyoelektriksel olmayan artefaktlar olabilir. Artefakt, çoğu zaman elektrot ve saçlı deri arasındaki temas bozukluğundan kaynaklanır. Elektrotun yerinden çıkması, amplifikatörlerden kaynaklanan aşırı gürültü, dış kaynaklardan gelen akımlara bağlı artefaktlar, elektrostatik potansiyeller (hastanın çevresindeki sağlık personelinden kaynaklanan) veya intravenöz infüzyonlar (infüzyon

sıvısının damlaması ile ortaya çıkan keskin dalgalar) şeklinde artefaktlar olabilir. Biyoelektrik artefaktlar, serebral kaynaklı olmayan potansiyeller olup oküler, kardiyak, glossokinetik, yutkunma, kas ve hareket artefaktları olarak hastadan kaynaklanırlar. (Bora vd.,2008)

3.3.4. Kayıt

EEG kaydı genellikle sessiz, rahat şartlarda yapılır; şekil 2'de görüldüğü gibi çekim sırasında hasta gözleri kapalı olarak oturur veya uzanır. Konvansiyonel EEG cihazları ile her bir montajdan yaklaşık 2-3 dakika ve toplamda yaklaşık 15-20 dakika süreli, 30mm/s kağıt hızı ile kayıtlar yapılır. Her bir bağlantının aktivitesinin kaydı esnasında, gözlerini tekrar kapatmadan önce hastadan yaklaşık 10 saniye süre ile gözlerini açması istenir, böylece zemin aktivitesinin cevabı değerlendirilebilir.



Şekil 2. Elektroensefalografi (EEG) kaydının alınması

Testin rutin kısmı tamamlandıktan sonra anormallikleri provoke etmek amacı ile aktivasyon yöntemleri ile kayıt devam eder. EEG anormalliklerini provoke etmek veya ortaya çıkarmak için 3-5 dakika süre ile hiperventilasyon yaptırmak etkili ve genellikle iyi tolere edilen bir yöntemdir. Hastadan belirlenen sürede normal hızında ancak derin

nefesler alıp vermesi istenir. Arteryel CO2 basıncındaki düşme, serebral vazokonstrüksiyona ve böylece serebral anoksiye neden olur. Bu genellikle EEG anormalliklerini açığa çıkarmak için yapılır. Hiperventilasyon genellikle çocuklarda yetişkinlere göre daha belirgin değişikliklere neden olmakta, fakat bireysel cevaplar arasında önemli farklılıklar olması nedeniyle normal sınırların tanımlanmasını zorlaştırmaktadır.

Aralıklı fotik stimülasyon, fotosensitiviteyi araştırmak için kullanılır. Verilen herhangi bir flaş frekansında yaklaşık 5 saniye hastanın gözleri açık ve daha sonra 5 saniye kapalı iken EEG kaydedilir. Genellikle saniyede 1-30 Hz arası flaş frekansları kullanılır, fakat daha geniş frekans sınırları da uygulanabilir.

Uyku sırasında veya uyku deprivasyonunun ardından kayıt yapmak, rutin çekimde ortaya çıkmayan EEG anormalliklerini provoke edebilir. Şüpheli epilepsisi ve özellikle uyku ile ilişkili nöbetleri olan hastaların araştırılmasında kullanılmaktadır.

Değişik işitsel uyaranlar epilepsili hastalarda EEG anormalliklerini presipite edebilir, ancak komalı hastaların değerlendirilmesi hariç EEG laboratuvarlarında rutin olarak kullanılmamıştır. Paroksizmal EEG anormalliklerini kolaylaştıran diğer uyaranlar taktil uyaran ve okumaktır. Birçok farmakolojik aktivasyon yöntemi de tanımlanmıştır, fakat bunlar rutin uygulamada olmamaları ve hasta için risk oluşturmaları nedeniyle günümüzde yaygın olarak kullanılmamaktadır.(Bora vd.,2008)

3.4. EEG Aktivitesini Tanımlayan Nitelikler

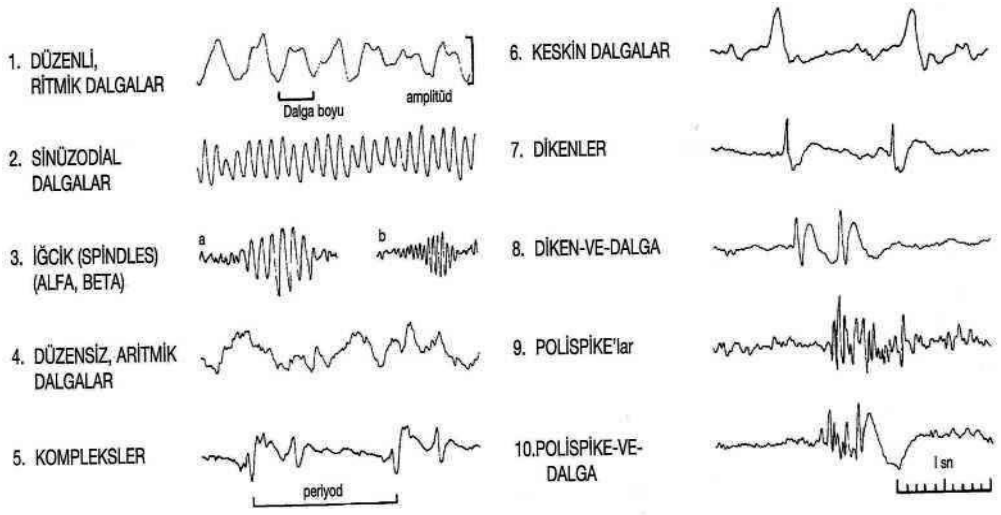
3.4.1. Dalga Formu

Dalga formu veya biçimi, bir dalganın görünüşünü veya morfolojisini tanımlamak için kullanılan terimlerdir. Dalga formuna bakılmaksızın, iki kayıt elektrotu arasındaki herhangi bir elektriksel potansiyel değişikliğine dalga denilmektedir. Her dalga veya dalga dizisi aktivite olarak adlandırılır. Birçok dalga formunun düzenli bir yapıda olduğu bilinmektedir. Yani simetrik yükselme ve alçalma fazı nedeniyle oldukça değişken bir yapıya sahiptir. Diğer düzenli dalga formları yay şeklinde veya testere dişi

şeklinde olabilirler. Düzensiz dalgaların şekilleri ve süreleri eşit olmamaktadır.(Fisch ve Spehlmann,1998)

Dalgalar; monofazik, difazik, trifazik ve polifazik olarak da sınıflandırılabilirler. Monofazik dalgalar; izoelektrik çizgiden yukarı veya aşağı yönde tek bir defleksiyon gösteren dalgalar olarak bilinmektedir. Difazik dalganın zıt yönlerde iki bileşeni, trifazik dalganın ise izoelektrik çizgi etrafında değişen üç bileşeni bulunmaktadır. Polifazik dalganın ise, değişik yönlerde iki veya daha fazla bileşeni olduğu görülmektedir. .(Fisch ve Spehlmann,1998)

Epileptiform paternler genellikle diken veya keskin dalga olarak kabul edilen dalga formlarıdır. Diken (spike), 20-70 msn'lik keskin konturlu dalga formudur. Keskin dalgalar (sharp wave), 70-200 msn süreli olmakta ve diken kadar keskin konturlu olmayabilmektedir. Diken dalgalar bazen yavaş bir dalga tarafından izlenebilmektedir. Bunun sonucu olarak, düzenli aralıklarla yineleyebilen diken ve dalga kompleksi oluşmaktadır. 3 Hz'den düşük hızlarda yineleyen diken-dalga kompleksleri, yavaş diken ve dalga kompleksleri olarak isimlendirilmektedir. Bir keskin dalgayı yavaş bir dalga izlerse, bu dalgalar keskin ve yavaş dalga kompleksini oluşturmaktadır. Bazı durumlarda iki ya da çok diken ardı sıra ortaya çıkıp, polispike kompleksi olarak da isimlendirilen, multiple diken kompleksini (multiple spike complexes) oluşturmaktadır. Bu kompleksler yavaş bir dalga tarafından izlenebilmekte ve böylece çoklu diken-yavaş dalga kompleksini ya da polispike ve yavaş dalga kompleksinin bir parçasını oluşturmaktadırlar. .(Fisch ve Spehlmann,1998)



Şekil 3. EEG Görülen Dalga Formları

3.4.2. Tekrarlama

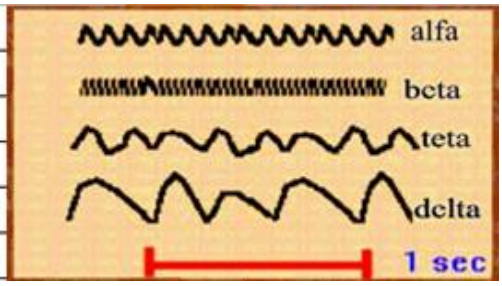
Dalgaların tekrarlaması ritmik veya aritmik olabilmektedir. Ritmik tekrarlı dalgalarda, tek tek dalgalar arasında benzer aralıklar bulunmaktadır. Genellikle düzenli ve sıklıkla da sinüzoidal şeklinde olmaktadır. Uyku iğcikleri (spindles) amplitüd olarak yavaş yavaş yükselen ve daha sonra alçalan ritmik tekrarlı dalga gruplarıdır. Aritmik tekrarlı dalgalar, tek tek dalgalar arasında değişken, düzensiz aralıklarla tanımlanmaktadır. Bunlar değişik frekanslardaki dalga dizileri olarak kabul edilebilir. Genelde düzensiz bir biçime sahiptirler .(Fisch ve Spehlmann,1998)

3.4.3. Frekans

Frekans; tekrarlı bir dalganın bir saniye içinde kaç kez yinelendiğini göstermektedir. Bir saniyede 3 çevrimi tamamlayan bir dalgaya 3 Hz'lik ya da saniyede 3 kez yineleyen dalga adı verilmektedir. Bir dalganın ya da yinelemeli dalganın frekansı, tek bir dalganın süresi ve dalga boyu ölçülerek ve evrik değeri hesaplanarak belirlenmektedir. Örneğin, tekrarlansın ya da tekrarlanmasın, 250 msn ya da 1/4 saniye süren bir dalganın frekansı, 4 Hz olarak hesaplanır. Tek dalgalar ve kompleksler, dalga boylarından daha uzun aralıklarla yineleyebilir. Bu durumda periyodik diye isimlendirilirler ve dalgalar arasındaki zaman aralığına da periyot denilmektedir .(Fisch

ve Spehlmann,1998) Tablo1'de olduğu gibi EEG dalgalarının frekansı genellikle beş alt frekans bandına ya da gruba ayrılmaktadır. Bunlar;

	Altbandın Adı	Simgesi	Aralığı
1	Delta	δ	1-4 Hz
2	Teta	Θ	4-8 Hz
3	Alfa	α	8-12 Hz
4	Beta	β	12-22 Hz
5	Gama	γ	22-30 Hz



Tablo 1. EEG Dalgalarının Frekans Altbandları

3.4.4. Genlik(Amplitüt)

EEG dalgalarının genliği, mikrovolt (µV) düzeyinde ölçülmektedir. Bir dalganın toplam dikey (vertikal) uzunluğunun, aynı kazanç (gain) ve filtre ayarlarında kaydedilmiş kalibrasyon sinyali yüksekliğiyle kıyaslanması ve ölçülmesiyle belirlenmektedir. Örneğin; bir EEG dalgasının yüksekliği 14mm ve 50 µV'lik kalibrasyon sinyali 7 mm olarak ölçülmüşse, dalganın genliği 100 µV'dir.(Fisch ve Spehlmann,1998)

Genlik; hiçbir zaman kalem sapmasının boyuna dayanarak belirtilmemelidir. Çünkü sapma, cihazın ayarlanmasına bağlı olarak değişmektedir. Klinik EEG'de genlik çoğu zaman mikrovolt olarak değil, kabaca düşük (20 µV altı), orta veya ılımlı (2050 µV arası) ya da yüksek (50 µV üstü) olarak adlandırılmaktadır. Ne var ki bu terimler sıklıkla, aynı kayıt içindeki diğer dalgaların genliğine göre belli dalgaların genliğini tanımlamak için kullanılmaktadır. Örneğin 40-50 µV'lik bir zeminde ortaya çıkan 60 µV'lik bir dalganın yüksek genliği olduğu söylenememektedir. (Fisch ve Spehlmann,1998)

Önemli bir anormallik, başın karşılıklı iki yanında eşzamanlı olarak kaydedilen aktivitenin genliğindeki asimetri olarak bilinmektedir. Eğer asimetri sürekli ise, genlikteki en ufak değişiklik bile klinik önem taşımaktadır. Özellikle erişkin EEG'si açısından -alfa ritmi dışında- geçerli olmaktadır.(Fisch ve Spehlmann,1998)

Genlikteki değişimler bazen beyin dışı etmenler, özellikle empedans farkı olan ve eşit mesafelerde yerleştirilmeyen elektrotlar nedeniyle oluşabilmektedir. Bu nedenle anormal genliğin gerçek olduğunu kabul etmeden önce teknisyen, elektrotların doğru yerleştirildiğini ve empedansı denetlemelidir.(Fisch ve Spehlmann,1998)

3.4.5. Dağılım

Dağılım başın farklı bölgeleri üzerine yerleştirilen elektrotlarla kaydedilen elektriksel aktivitenin oluşumunu anlatmaktadır. EEG paternleri başın her iki tarafında geniş alanlarda olabileceği gibi tek hemisferde de bulunabilmekte ya da küçük bir alanda sınırlı kalabilmektedir. Yaygın, difüz veya jeneralize dağılım, başın tümünde veya tümüne yakın kısmında aynı zamanda görülen aktiviteyi ifade etmektedir. Jeneralize aktivitenin, dağılım alanı içinde açıkça maksimum bir amplitüdü olabilmekte ve bulunduğu alanda yapılan bipolar kayıtlarda faz karşılaşması ve referans montajda çok yüksek amplitüdle tanınmaktadır.(Fisch ve Spehlmann,1998)

Lateralize dağılım, sadece ya da çoğunlukla başın tek yanında görülen aktiviteyi ifade etmektedir. Lateralize aktivite anormaldir; ya anormal aktivitenin bulunduğu tarafta ya da normal aktivitenin bulunmadığı tarafta serebral bir anormalliği düşündürmektedir. Bazı normal paternler başın bir tarafında bir süre görüldükten birkaç saniye ya da dakika sonra diğer alanda oluşabilmektedir.(Fisch ve Spehlmann,1998)

Fokal aktivite, bir alan üzerindeki bir veya birkaç elektrotla sınırlı olarak saptanabilen aktivitedir. Bazı komşu elektrotlar, bu aktiviteyi daha düşük amplitüdlü olarak alabilmektedirler. Bu sınırlı dağılım, bir bölgede maksimum amplitüdü olabilecek, geniş ya da jeneralize bir dağılımdan ayırt edilmesi gerekmektedir. Bu ayırım özellikle anormal yavaş ve keskin dalga söz konusu olduğunda önemli taşımaktadır.(Fisch ve Spehlmann,1998)

3.4.6. Faz İlişkisi

Faz; bir veya birkaç kanaldaki dalga bileşenlerinin zamanlama ve polaritesini ifade etmektedir. Değişik frekanstaki dalgalar değişik kanallarda ortaya çıkabilmektedir. Böylece aynı anda pikler ve çukurlar oluşmaktadır. Bu dalgaların faz birlikteliği (inphase) içinde olduğu söylenebilmektedir. Eğer frekanstaki dalgalar bu tarzda bir araya gelmemişlerse o zaman faz ayrılığı (out of phase) vardır denilmektedir. Faz farklılıkları, faz açılarına dayanarak ifade edilebilmektedir. Örneğin, zıt yöndeki pik noktaları 180° faz ayrılığı göstermektedir. Böylesi bir "faz karşılaşması", bipolar kayıtlarda EEG potansiyellerinin kaynağının en büyük göstergesi durumundadır. Faz, tek bir kanalda, bir ritmin değişik bileşenleri arasındaki zaman ilişkisini göstermektedir. Örneğin, sinüsoidal bir dalganın pik noktası ile sıfır noktası arasında 90°, bir sonraki pik ile 360° bulunmaktadır.(Fisch ve Spehlmann,1998)

3.4.7. Zamanlama

Başın değişik alanlarındaki dalgaların zamanlaması aynı veya farklı olabilmektedir. Simültane (eşzamanlı) ve senkron terimleri iki olayın aynı zamanda oluştuğunu ifade etmektedir. Bu terimler genellikle aynı anlamda kullanılmaktadır. Ancak "senkron" terimi bazen tam bir aynı anda oluşumu vurgulamak için kullanılırken, "simültane" daha geniş olarak, EEG'nin görece daha yavaş kayıt hızı sınırları içinde, sadece kesin olmayan bir tarzda görülebilen, aynı anda oluşumu belirtmek için kullanılabilir.(Fisch ve Spehlmann,1998)

Değişik kanallarda oluşan ve aralarında sabit zaman ilişkisi olmayan dalgalara asenkron dalgalar denilmektedir. Bu genellikle, aynı frekansta olmasalar ya da aynı fazda birleşmeseler bile, değişik alanlarda aynı zamanda ortaya çıkan dalgaları göstermektedir. Eğer dalgalar bir an bir alanda, başka bir zaman da başka bir alanda oluşuyorsa, bunlara bağımsız dalgalar denilmektedir.(Fisch ve Spehlmann,1998)

3.4.8. Israrlılık

Israrlılık (persistans); bir kayıt esnasında bir dalga veya paternin hangi sıklıkla oluştuğunu tanımlanmaktadır. EEG paternlerinin klinik öneminin sıklıkla, sadece ısrarlılıklarına değil, aynı zamanda amplitüdlerine de bağlı olması nedeniyle, ısrarlılık ve amplitüd, genellikle kantite, miktar veya belirginlikle (quantity, amount, prominence) tanımlanmaktadır. Tek tek dalgalar ve kompleksler yüksek, orta ve düşük ısrarlılıkla oluşabilmektedir. Bunların ısrarlılığı en iyi, bir saniye veya bir dakika içindeki ortalama sayıları ile ifade edilmektedir. Belirli aralıklarla ya da düzensiz aralarla oluşabilmektedir. Düzensiz ve seyrek oluşum bazen sporadik olarak da adlandırılmaktadır.(Fisch ve Spehlmann,1998)

3.4.9. Reaktivite

Reaktivite, çeşitli manevralarla bazı normal ve anormal paternlerde ortaya çıkabilen değişiklikleri anlatmaktadır. Bazı paternler, gözlerin açılıp kapatılması, hiperventilasyon, fotik ya da duyusal uyaranlar, uyanıklık düzeyinde değişiklikler, hareket veya diğer manevralar ile kışkırtılabilir ya da artırılabilir, azaltılabilir ya da engellenebilmektedir. Toksik ve metabolik ensefalopatilerde anormal yavaş dalgalar genellikle hastanın uyarılmasıyla azalmaktadır. Hiperventilasyon ve uyuklamayla artarken, yapısal lezyon vakalarında görülen anormal yavaş dalgalar uyarıcı manevralar sırasında genellikle daha az azalabilmekte ya da engellenebilmektedir.(Fisch ve Spehlmann,1998)

3.5. EEG'de Kaydedilen Normal Aktivite

3.5.1. Alfa Aktivitesi

Alfa ritmi 8-13 Hz arasında frekansa sahiptir, ancak yetişkinlerin çoğunda 9-11 Hz arasındadır. Uyanıkken genellikle tipik olarak başın arka bölümlerinde (parietal ve oksipital) bulunur, fakat santral veya temporal bölgelerde de bulunabilir. En iyi ve belirgin olarak gözler kapalı, istirahat halinde görülür. Göz kapanmasının hemen

ardından sıklığı geçici olarak artar. Alfa ritmi görsel dikkat ve sensoryel stimuluslarla geçici olarak ortadan kalkar (Şekil. 3). Normal bireylerde alfa aktivitesi iyi oluşur ve baskındır. İleri yaş, antikonvulzanlar gibi belirli ilaç tedavisi alanlar ile konfuzyondaki hastalar veya herhangi bir serebral patolojisi olan hastalarda yavaşlama olur. Alfa aktivitesinin frekansı çocuklarda ve tirotoksik yaşlılarda artabilir. (Bora vd., 2008)

Alfa aktivitesinin genliğinde sıklıkla iki hemisfer arasında hafif asimetri olabilir. Sağ hemisferde alfa ritminin daha yüksek genlikli olması oldukça sık görülür, ancak bu fark %50'in üzerinde olmamalıdır (Şekil. 4). Bu fark, sağın dominant olmayan hemisfer olması ve kranyum kalınlığındaki farklılıklarla ilişkilendirilmiştir. Daha belirgin amplitüt asimetrisi lateralize edici öneme sahip olabilir, fakat eşlik eden EEG anormalliklerinin yokluğunda değerlendirme zordur, çünkü hem supresyon, hem de amplitüt artışı hemisfer lezyonu tarafında meydana gelebilir. Amplitüt farkı gibi, iki hemisfer arasında 1 Hz. den fazla ısrarlı alfa frekansı farklılığı genelde anormal kabul edilse de başka anormalliklerin eşlik etmediği durumlarda hangi tarafın anormal olduğunu yorumlamak zordur. Tek taraflı alfa ritm cevabının kaybı bazen temporal veya parietal lob lezyonlarında meydana gelebilir. (Bora vd., 2008)

3.5.2. Beta Aktivitesi

Beta aktivitesi, 13 Hz. den daha yüksek frekansa sahip ritmik aktivitedir. Hemisferlerin anterior (frontosantral) kısmında bulunur, ancak özellikle uykuda yaygın yerleşim gösterir ve genellikle 30 μV'dan daha düşük bir amplitüde sahiptir. Farklı normal bireylerde beta ritminde değişiklikler olabilir. 18-25 Hz arasında frekansa sahip olan aktivite uykuya dalma, yüzeysel uyku ve hızlı göz hareketleri (REM) uykusu sırasında genellikle daha dikkat çekicidir. Beta aktivitesi, barbitürat ve benzodiazepinler başta olmak üzere birçok farmakolojik ajandan etkilenebilir. Fokal veya lateralize spontan beta aktivitesi veya asimetrik ilaçla indüklenen hızlı aktivite, lokalize serebral patoloji olasılığını arttırır; fakat bu durumlarda tek serebral hemisferi içeren bir lezyonun ipsilateral veya kontralateral beta aktivitesinin genliğini arttırmış olabileceği akla gelmelidir. Beynin yüzeyinde kaydedici elektrotların büyük çoğunluğunda olan bir kranyum kemik defekti alanı üzerinde beta aktivitesi amplitüt olarak artar.

Jeneralize paroksizmal hızlı aktivite, EEG kayıtlarının yüzde birinden azında görülen nadir bir bulgudur ve kas artefaktı, ilaç etkisi, uyku iğcikleri gibi durumlara bağlı olarak görülebilir. (Bora vd., 2008)

3.5.3. Teta Aktivitesi

Frekansı 4-7 Hz arasındaki aktivite, teta aktivitesi olarak adlandırılır. Yavaş dalga aktivitesi genelllikle çocuklarda belirgindir, büyüme ile birlikte serebral maturasyona paralel olarak trasedeki dağılım oranı azalır. Bazı teta aktiviteleri genç yetişkinlerde özellikle temporal bölgelerde ve hiperventilasyon esnasında bulunabilir, fakat daha yaşlı kişilerde 30jıV'dan daha büyük amplitütlü teta aktivitesi uyku dönemleri hariç daha az görülür. Fokal veya lateralize teta aktivitesi lokalize serebral patolojinin göstergesi olabilir. Daha diffüz dağılmış teta aktivitesi geniş bir nörolojik hastalık yelpazesindeki hastalarda sık bir bulgudur. (Bora vd., 2008)

3.5.4. Delta Aktivitesi

Frekansı 4 Hz. den daha yavaş olan ritmlere delta aktivitesi denir. Bu tür aktivite süt çocuklarında ön plandadır ve daha büyük bireylerde derin uyku (NREM III, IV) sırasında normal bir bulgu olabilir. Uyanık yetişkinlerde istirahat halinde görülen delta aktivitesi her zaman anormal bir bulgudur. (Bora vd.,2008)

3.5.5. Mü Ritmi

Mü ritmi genelikle alfa ritminde olan frekansa sahiptir ve tek veya iki hemisferin santral bölgelerinde görülür. Göz açmadan etkilenmez, hareketin yapılması veya hareketin düşünülmesi ile unilateral veya bilateral olarak bloke olur. Bilateral mü ritmi sıklıkla asenkronizedir ve iki hemisfer arasında amplitüt asimetrilerini açığa çıkarabilir. Dalgaların negatif kısımları genellikle keskin, pozitif bölümleri ise genellikle yuvarlaktır. Mü ritmi sıklıkla kontralateral hareket ile de etkilenen santral lokalize beta

ritmi ile ilgilidir. Birçok durumda mü ritmi tanısal öneme sahip değildir. (Bora vd.,2008)

3.5.6. Lambda Dalgaları

Lambda dalgaları, iyi belirlenmiş bir alandaki nesneye özellikle de dikkat ve ilgileri artmışken bakan normal bireylerde oksipital bölgede oluşabilen elektropozitif keskin (testere görünümlü) dalgalardır. Morfolojik olarak benzer aktivite bazen NREM uykusu esnasında da görülür. Bu potansiyellerin doğası belirsizdir ve şu an için bilinen herhangi bir tanısal önemleri yoktur. Bazen her iki hemisfer arasında asimetrik olabilirler, fakat ortam aydınlatılması ile zemin aktivitesinin amplitüdünde azalma veya göz kapatmaya cevap olarak patolojik önemli keskin dalgalardan ayrılabilir. (Bora vd.,2008)

3.5.7. 14 ve 6 Hz Dikenler

Bu aktiviteler, uykuya dalma ve yüzeyel uyku sırasında, özellikle ergenlerde daha yavaş dalgalar üzerine yerleşmiş olarak meydana gelebilir; genellikle bir saniyeden daha kısa sürede sonlanır. Özellikle bir veya her iki yanda posterior temporal bölgelerde bulunur ve en iyi referans montajlarda görülür. Patolojik öneme sahip değildir. (Bora vd.,2008)

3.5.8. Küçük Keskin Dikenler

Küçük keskin dikenler veya uykunun benign epileptiform geçicileri (transientleri) uykuya geçiş veya yüzeysel uyku sırasında, normal yetişkinlerin yaklaşık dörtte birinde bulunur. Bazen bir yavaş dalganın takip ettiği monofazik veya bifazik dikenler içerir, fakat bu görünüme keskin dalgalar veya zemin aktivitesinin ritmik fokal yavaşlaması eşlik etmez. Genellikle, lokalizasyonu rastgele değişir ve her iki hemisferde bağımsız olarak ortaya çıkarlar, fakat en iyi anteromeziyal temporal bölgelerde görülürler. Zemin aktivitesinde bir bozukluk yoktur. Bu tür değişiklikler normal olarak

tanımlanabilir ve şüpheli epilepsisi olan hastaların değerlendirilmesinde tanısal yararı yoktur. (Bora vd.,2008)

3.5.9. 6 Hz Diken- Dalga Aktivitesi

Genellikle 1 saniyeden kısa sürede sonlanan 6 Hz diken dalga aktivitesinin kısa boşalımları, normal ergen ve genç yetişkinlerde özellikle uyku sırasında görülebilir ve bazen fantom diken-dalgası olarak da adlandırılır. Uykunun daha derin evreleri sırasında açığa çıkar. Genelde asimetrik olan bu boşalımlar, bazı durumlarda bilateral simetrik ve senkrondur. Bu dikenler, yavaş dalga ile kıyaslandığında genellikle daha düşük amplitütlüdür ve farkedilmesi zor olabilir. Amplitütleri yüksek olduğu zaman boşalımların patolojik olma olasılığı daha yüksektir. Ağırlıklı olarak frontal bölgede görülen boşalımlar, aynı anda oksipital olarak elde edilenlere göre epilepsi ile daha çok ilişkilidir. (Bora vd.,2008)

3.5.10. Uykuya Geçişin Ritmik Temporal Başalımları

Uykuya geçiş veya yüzeyel uyku sırasında özellikle genç er-keklerde, bazen çentikli bir görünümle birlikte ritmik keskin karakterli teta boşalımları daha çok orta temporal bölgelerde unilateral veya bilateral olarak meydana gelir. Eğer bilateralse, senkron veya bağımsız olarak bir yandan diğerine yayılabilir. Patolojik önemleri yoktur. (Bora vd.,2008)

3.5.11. Ritmik Teta Boşalımları

Ritmik teta aktivitesi boşalımları, 40 yaşın üstündeki hastalarda, sıklıkla ani başlangıçlı olup, ani bir şekilde kaybolurlar. Yetişkinlerdeki bu subklinik ritmik elektrografik boşalımlar (SREDA) genellikle bilateral ve diffüz olarak dağılmıştır, fakat nadiren fokal veya lateralizedir. Boşalımlar diffüz olduğu zaman sıklıkla parietal ve posterior temporal bölgeler üzerinde belirgindir. İstirahatte, hiperventilasyon esnasında veya uykuya geçiş durumunda oluşabilir. Boşalımlar, 1 veya 2 dakika içerisinde veya

daha fazla sürede sonlanır ve subklinik bir nöbeti anımsatır. Bu tür boşalımlar, tanısal öneme sahip değildir ve epilepsi veya herhangi bir klinik şikâyetle uyumlu değildir. (Bora vd.,2008)

3.5.12. Yaşla İlgili EEG Değişiklikleri

Yaşlanma ile bazı EEG değişiklikleri meydana gelir ve bireyler arasında büyük farklılıklar gösterir. En sık görüleni, genç yetişkinlere kıyasla normal yaşlılarda alfa frekansının yavaşlamasıdır. Aynı zamanda entelektüel azalma ile alfa ritminin yavaşlaması arasında bir ilişki vardır. EEG'deki yavaşlamanın nedeni tam olarak açıklanamamakla birlikte genelde vasküler faktörler sorumlu tutulmaktadır. (Bora vd.,2008)

Elli yaşın üzerindeki bireylerde alfa benzeri aktivite, bir veya iki temporal bölgede görülebilir. Bu bazen oksipital alfa ritminden daha dikkat çekici olup sıklıkla uykuya geçiş sırasında açığa çıkar ve nadiren bilateral, biraz asenkron boşalımlar halinde meydana gelebilir. (Bora vd.,2008)

Yaşlanmanın beta aktivitesi üzerine etkisi daha az bilinmektedir. Bu değişiklikler; analiz ve kayıt yöntemleri arasındaki farklılıklar, tıbbi ve ilaç öyküsünün alınmasında ile entellektüel fonksiyonun saptanmasındaki başarısızlık gibi nedenlerle sınırlanmıştır. (Bora vd.,2008)

Yaşlı kişilerde difüz teta ve delta aktivitesi belirgin olarak artmıştır. Bu yavaşlamanın normal yaşlanma değil, dejeneratif veya vasküler değişikliklerin sonucu olduğu kabul edilmektedir. EEG uyku paternleri de, total uyku zamanındaki azalma ve özellikle evre 3 ve 4 uykusunun süresinde azalma ile birlikte yaşa göre değişiklik gösterir. (Bora vd.,2008)

3.6. Anormal EEG'nin Tanımlanması

Bir EEG eğer; normal aktiviteden farklı olarak A. interiktal paroksismal (nöbet arası) Patternler ve iktal (nöbet anı) patternler içeriyorsa anormal eeg olarak nitelendirilir. En önemli EEG anormallikleri aşağıdaki temel anormal EEG patternlerine ayrılabilir (Safiye Bilgin,2012):

A. İnteriktal Paroksismal (Nöbet Arası) Patternler:

1) Diken (Spike)
2) Keskin (Sharp)
3) Multipl Diken
4) Hızlı Diken Akımları
5) Diken ve Yavaş Dalga Kompleksi
6) Yavaş Diken ve Dalga Kompleksi

B. İktal (Nöbet Anı) Patternler:

1) Jeneralize Tonik Klonik Konvülsiyon
2) Absans
3) Kompleks Parsiyel Nöbetler
4) Basit Parsiyel Nöbetler
5) Myoklonik Nöbetler
6) Tonik Nöbetler
7) Atonik Nöbetler
8) Salaam Ataklar

3.7. EEG Desenleri ile Nörolojik Hastalıklar Arasındaki Bağıntı

EEG'nin epilepsinin tanısında yeri çok önemlidir. Ayrıca spesifik Epileptik sendromların ayrımında da işe yaramaktadır. EEG de epileptik dalgaların varlığı epilepsi tanısını destekler, fakat yokluğu tanıyı dışlamaz. Epileptik olguların %10 ile %40'ında EEG daima normal kalır. Uyku, uykusuz bırakılma, fotik stimulasyon, hiperventilasyon epileptik hastalarda deşarjların ortaya çıkışını kolaylaştırır. West sendromu, Lennox- Gastaut Sendromu, Çocukluğun Absans Epilepsisi, Benign Rolandik Epilepsi, Jüvenil Myoklonik Epilepsi, Temporal lop nöbetleri özel EEG bulguları ile kolaylıkla tanınabilir. Bunlardan belki en tipik olanı çocukluğun absans epilepsisidir. 3- 3.5 Hz' lik diken-dalga deşajı görülür.(ETD,2012c)

Demans ve diffüz ansefalopatiler: Bu tür hastalarda EEG iyi bir laboratuar testidir. Trifazik dalgaların varlığı toksik ve metabolik ensefalopatiye işaret eder. Yüksek voltajlı ve yüksek frekanslı beta aktivitesi hipnotik kullanımda görülür. Creutzfelt-Jakop hastalığında %90 0.5-1 Hz periyodik jeneralize keskin dalga deşarjları görülür. Alzheimer hastalığında zemin ritminde diffüz yavaşlama görülür. Multi-infarkt demanslarda fokal yavaşlamalar izlenir.(ETD,2012c)

Fokal serebral lezyonlar: Yapısal lezyonların tanısında EEG'nin rolü sınırlıdır. Ancak kitle lezyonlarının epileptojenik özelliklerinin değerlendirilmesinde yararlıdır. İnfarktların erken tanısında kısmen yararlı olabilir. Örneğin geçici iskemik atakta BT ve MRG normalken EEG zemin ritmi yavaşlamasını ortaya koyabilir.(ETD,2012c)

Serebral infeksiyonlar: Herpes Simpleks ansefalitli olguların %80' inde fokal yavaşlama ve periyodik lateralizan epileptiform deşarjlar (PLED) izlenir. PLED'ler tipik olarak hastalığın 2-15. günleri arasında görülür. SSPE' de 4-10 sn aralıklarla tekrarlayan yüksek voltajlı, delta dalgaları görülür. AIDS demans kompleksinde EEG de nonspesifik zemin ritmi yavaşlamaları görülür. Eğer fokal yavaşlamalar varsa o durumda AIDS üzerinde yerleşmiş lenfoma, toxoplasma enfeksiyonu gibi olaylar akla gelmelidir.(ETD,2012c)

BÖLÜM 4

EEG SİNYAL ANALİZİ VE MODELLEME METODLARI

EEG işaretlerinde beyin fonksiyonlarıyla ilgili çok miktarda bilgi saklıdır. Bu bilgilerin tıbbi araştırmalarda ve hastalık teşhisinde kullanılabilmesi için, modern parametrik yöntemlerle gerçek zamanda spektral analizinin gerçekleştirilmesi ve otomasyona geçilmesi gerekmektedir.(Batar,2005)

Sinyallerin analiz edilmesi direk sinyalin kendisi veya sinyalin gösteriminin başka boyutlara (zaman, frekans, zaman- ölçek.vb) taşınarak gerçekleştirilmesiyle yapılmaktadır. Amaç; sinyalin bilgi kaybına uğratılmadan bu boyutlardan birisine dönüşümü yapılarak işlenmemiş haldeki verilerden sağlıklı seçilemeyen anlamlı detay bilgilerine ulaşmaktır.(Batar,2005)

Öncelikle dönüşüm işlemine ihtiyaç duyulur. Her hangi bir sinyalden daha fazla bilgi almak için matematiksel dönüşüm işlemi uygulanır. Mevcut bir çok dönüşüm tekniği farklı tip sinyaller ve amaçlar için kullanmaktadır. Hilbert, KZFD, Zak Fourier, Running Fourier, Wigner, Radon ve Wavelet dönüşümü gibi. Her bir dönüşüm tekniğinin kendine göre avantajları ve dezavantajları vardır. Bunlardan en eski olanlardan biriside Fourier Dönüşüm tekniğidir .(Batar,2005)

4.1. Sinyal Analiz Metotlarının Gelişimi

19.Yüzyılda Fourier herhangi bir periyodik fonksiyonun sonsuz sayıdaki kompleks üstel fonksiyonların toplamıyla ifade edilebileceğini söylemiştir. Fourier temelli dönüşüm teknikleri daha sonraki yıllarda Ayrık Fourier Dönüşümü (DFT) ve

Hızlı Fourier dönüşümü (FFT) algoritmalarının geliştirilmesiyle bilgisayarlarda veri isleme için önemli bir basamak atlanmıştır.(Batar,2005)

Denis Gabor 1946 yılında Gabor Dönüşümü olarak bilinen Fourier temelli analiz yöntemini sunmuştur. Bu metot da pencereleme yöntemini kullanarak ki pencere fonksiyonu Gauss Hata fonksiyonudur, işaretin küçük bir parçasını zaman tanım aralığında ele alınmış, işaret zaman ve frekansın fonksiyonu olarak iki boyutta ifade edilmiştir. Böylece durağan olmayan sinyallerin islenmesi daha sağlıklı hale gelmiştir. Gabor Dönüşümünü bugün Kısa Zaman Hızlı Fourier Dönüşümü (KZFD) olarak yöntemin kaynağı olarak görülmektedir.(Batar,2005)

Fourier temelli analiz yaklaşımlarının ilerleyen konularda da bahsedileceği dezavantajları (zaman ve/veya Frekans çözünürlüğü sorunu, her zaman var olacak olan Gibbs olayı..vb) (Robertson, 1998), sinyal analizcileri daha farklı yollara, modern ve parametrik yöntemlere yöneltmiştir.(Batar,2005)

Bu şekilde Fourier serisi yaklaşımıyla başlayan sinyal frekans analizi kavramı Günümüze değin ölçek analizi kavramına ve bununla sıkı sıkıya bağlı dalgacık dönüşümüne kadar gelişmiştir.(Batar,2005)

Alex Grossman ve Jean Morlet bu fonksiyon yapı blokları için ilk defa "Wavelet" yani DALGACIK adını önermişlerdir ve o zamana kadar Littlewood-Paley teorisi olarak Belirtilen teori "Wavelet" DALGACIK TEORİSİ olarak adlandırıla gelmiştir.(Batar,2005)

4.2. Fourier Temelli Sinyal Analizi

19.Yüzyılda Fourier herhangi bir periyodik fonksiyonun sonsuz sayıdaki kompleks üstel fonksiyonların toplamıyla ifade edilebileceğini söylemiştir. Fourier dönüşümü sonuçları frekans bileşenlerinin yerel zaman, başlangıç veya bitiş bilgileri hakkında kesin veriler sağlamaz çünkü bu bilgi Fourier spektrumu boyunca yayılmıştır. Yani Fourier dönüşümü ancak genel bir bakış sağlar. Bunun sebebi Fourier temel fonksiyonlarının (sinüs ve kosinüs) tam dayanaklı olmamalarıdır. Bundan dolayı

durağan sinyallere uygulanmalarında en iyi sonuç verir. Eşitlik 4.1' de görüldüğü gibi herhangi bir t_1 veya t_2 anı entegrale aynı etkide bulunacaktır.(Batar,2005)

$$x(f) = \int_{-\infty}^{\infty} x(t)e^{-jwt}\,dt \quad (4.1)$$

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{jwt}\,df \quad (4.2)$$

4.2.1. Akan fourier dönüşümü

Bir f(t) sinyalinin Akan Fourier Dönüşümü ifadesi; c verilen bir sabit olmak üzere Eşitlik 4.3 ile verilir. F(t,ω), f(t) sinyalinin τ dilimine karşılık gelen f(t+ τ) sinyalinin Fourier dönüşümüdür. Böylelikle durağan olmayan sinyaller için spektrumun değişimi sinyalin durağan kabul edildiği τ dilimleri ile belirlenebilmektedir.

$$F(t,\omega) = \int_{-c}^{c} f(t+\tau)e^{-j\omega\tau}\quad d\tau = e^{j\omega t}\int_{t-c}^{t+c} f(\alpha)e|^{-j\omega\alpha}\,d\alpha \quad (4.3)$$

F(t, ω) nini Ters Dönüşümü ifadesi de $\omega_\circ = \frac{\pi}{c}$ olmak üzere

$$F(t) = \frac{1}{2c}\quad \sum_{m=-\infty}^{\infty}\quad F(t, m\omega_0\ ) \quad (4.4)$$

eşitliği ile verilir.(Batar,2005)

4.2.2. Kısa zaman fourier dönüşümü

Daha önce FD'nin durağan olmayan sinyaller için elverişli olmadığı ifade edilmişti. Denis Gabor, 1946 yılında pencereleme yöntemini kullanarak, işaretin küçük bir parçasını zaman tanım aralığında ele almış, işareti zaman ve frekansın fonksiyonu olarak iki boyutta ifade etmiştir. Bu dönüşüm yönteminde işaretin belirli bir kesiminin durağan olduğu kabul edilebilecek bir pencereden geçirilir ve yerel bir frekans parametresiyle FD işlemi gerçekleştirilir. KZFD ile FD arasında çok az bir fark bulunur. KZFD' de sinyal küçük çerçeveler bölünür ve bu çerçeve anlarında sinyalin durağan olduğu kabul edilir. Durağanlığın geçerli olduğu bu çerçeveler sinyalin bir pencere

fonksiyonu ile çarpılmasıyla elde edilir. FD' nin yerelleştirilmesi fikrine dayanan bu teknik ilgilenilen yerde uygun bir pencere seçilerek dönüşüm işlemi gerçekleştirilir. Durağan olmayan sinyallerin spektral analizi çok iyi zaman ve frekans çözünürlüğü sunması gereken yetenekli işlevlere ihtiyaç duyar. KZFD temel bir pencere fonksiyonundan zaman alanında kaydırma ve frekans parametreleri olmak üzere iki ayrı şekilde türetilir. Burada kullanılan pencere fonksiyonu Eşitlik 4.5 ile ifade edilmiştir. Pencere fonksiyonu sonlu enerjiye sahip olması ve entegralinin alınabilmesi gerekmektedir. KZFD yönteminde zaman ekseni üzerinde bir τ noktasına w(t) pencere

$$g_{w,\tau}(t) = e^{jwt} w(t-\tau) \tag{4.5}$$

Fonksiyonunu yerleştirilerek pencerelenmiş fonksiyonun Fourier Dönüşümü gerçekleştirilir. Daha sonra pencere kaydırılır ve tekrar Fourier dönüşümü alınarak işleme devam edilir. Bu yöntemin matematiksel ifadesi Eşitlik 4.6' de ifade edilmiştir.(Batar,2005)

$$KZFD\ (w,\tau) = \int_{-\infty}^{\infty} f(t) w^*\ (t-\tau) e^{jwt} dt = < g_{w,\tau\ (t),f(t)>} \tag{4.6}$$

KZFD' deki f(t) fonksiyonu tekrar elde edilebilir özelliktedir. Bu durum Eşitlik 4.7 ile gösterilmiştir.

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} KZFD(w,\tau) g_{w,\tau}\ (t) dw, d\tau \tag{4.7}$$

4.2.3. Zak dönüşümü

Bir fonksiyonun Zak dönüşümü aşağıdaki gibi tanımlanır .(Batar,2005)

$$Zf(x,\omega) = \sum_l f(x+1) e^{-i\omega l} \tag{4.8}$$

4.2.4. Ayrık fourier dönüşümü

Fonksiyonlar teorik olarak tanımlı hallerde Fourier dönüşümleri rahatlıkla hesaplanabilir. Ancak uygulamada sinyallerin kesin fonksiyonel ifadeleri yoktur ve islenmeleri için analog sinyallerden örneklenmiş sınırlı sayıda sayısal ayrık dizileri mevcuttur. Bu yüzden daha önce bahsedilen şekilde Fourier dönüşümleri hesaplanamaz.(Batar,2005)

Ayrıca bütün frekans boyutunun analog olarak gösterimi sonsuz sayıda örneklenmiş işareti gerektirmektedir bu ise uygulamada mümkün değildir. Sayısal işaretler için Fourier dönüşümünün hesaplanılması için belirli sınırlamalı içindeki yaklaşıklıklarla verilebilir. Bir f[k] ayrık dizisinin N örneği için tanımlanan bu yeni dönüşüm Ayrık Fourier dönüşümü (AFD) olarak adlandırılır.(Batar,2005)

Tersi de alınabilen bu dönüşümün önemli özellikleri vardır. Ayrık Fourier Temelli dönüşümler dizinin periyodik olduğunu kabul ederler dolayısıyla bir ayrık zaman sinyali periyodik ise bunun yaklaşık Fourier dönüşümü AFD' dir . Özellikle iki AFD' nin çarpımı bunlara karsı düsen dizilerin ayrık-zaman boyutunda konvolusyon toplamıdır. Ayrıca sayısal ortamdaki bir çok spektral analiz yöntemi AFD' ye dayanmaktadır.(Batar,2005)

AFD, f(k), k = 0, 1, ..., N-1, gibi bir sonlu diziyi , F(n), n = 0, 1, … , N – 1.gibi diğer bir sonlu diziye esleyen önemli bir operatördür. Normalize edilmiş örnekleme frekansı 2π olmak üzere;

$$F[n] = AFD\,\{f[k]\} = \sum_{k=0}^{N-1} f[k]\, e^{-j2\pi kn/N} \qquad (4.9)$$

Ters Ayrık Fourier dönüşümü (TAFD) ise F[n]'yi yeniden f[k] dizisine dönüştürür:

$$f[k] = TAFD\,\{F[n]\} = \frac{1}{N}\sum_{n=0}^{N-1} F[n]\, e^{j2\pi kn/N} \qquad (4.10)$$

4.2.5. Hızlı fourier dönüşümü

AFD hesaplanmasında bugün kullanılmakta olan verimli ve etkin bir yaklaşım hızlı Fourier dönüşüm algoritmalarıdır. Eşitlik 4.9' da ilk dönemlerde fark edilmeyen ve sonuçlardaki periyodik sayısal tekrarlama özelliğine dayanan dairesel katlama dan (circular convolution) yararlanılarak geliştirilen birbirinden farklı algoritmalar mevcuttur. Her ne kadar farklı bir dönüşüm olarak adlandırılmışsa da Hızlı Fourier Dönüşümü (HFD) AFD' den farklı değildir. Sadece kullanılan hesaplama yöntemi FD tahmini için HFD'yi cazip kılar.(Batar,2005)

Son olarak Ayrık zamanlı bir f[k] işaretinin w[k] ayrık pencere dizisi kullanılarak AFD ifadesini aşağıdaki denklemle verebiliriz:

$$F[n] = AFD\{F[K]\} = \sum_{k=0}^{N-1} f[k]w[k]e^{-j2\pi kn/N} \qquad (4.11)$$

4.5. Dalgacık Dönüşümü (Wavelet) ile Sinyal Analizi

KZFD'nin sabit uzunluklu pencereler kullanması ve belirsizlik ilkesi kısıtlaması nedeniyle istenilen eş zamanlı zaman-frekans çözünürlüğü sağlayamaması gibi temel eksiklerle pencere uzunluğu değişken, çeşitli frekanslarda seviye seviye çözünürlük sağlayabilen ve hem zaman bilgisini hem de frekans bilgisini gösteren bir çözüm araştırılmaya başlanmıştır. Bulunan çözümlerden biri *Dalgacık Dönüşümü* 'dür.(Fidan,2006)

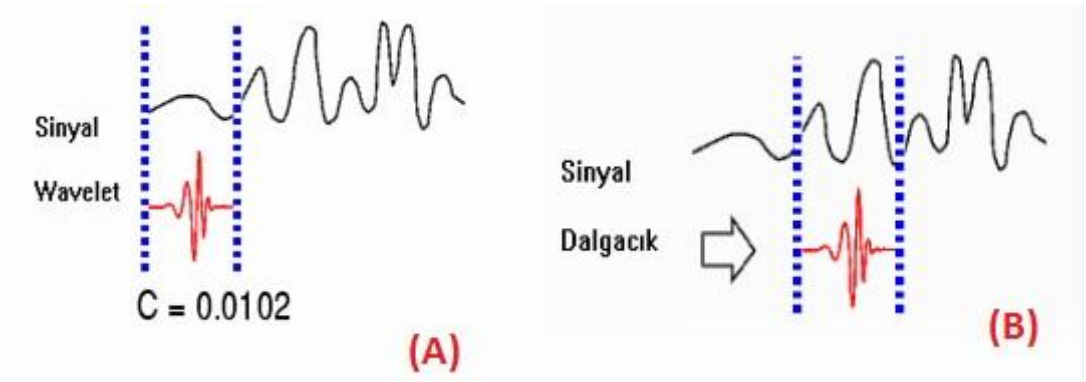
Dalgacık Analizi, sinyali analiz için bileşenlerine ayırırken, diğer bir deyişle sinyali *ayrıştırırken* (decomposition) baz fonksiyonları ile ifade etmesiyle Fourier analizine benzer. Fourier dönüşümünde, sinyal değişik frekanslardaki sinüzoidal serilere ayrılır, dalgacık dönüşümünde sinyal "ana dalgacık"ın ölçeklendirilmiş ve kaydırılmış versiyonları olan kendi "dalgacıklarına" ayrılır. Dalgacık analizinin avantajı baz fonksiyonuna *ölçekleme* (dilation, scaling) ve *öteleme* (translation) özelliği kazandırmasıdır. Dalgacık dönüşümünde sinyaller analiz edilirken analiz penceresinin uzunluğu ve konumu kendiliğinden değişir. KZFD'nde ise analiz penceresi kaydırılabildiği halde uzunluğu analiz boyunca sabit kalır.(Fidan,2006)

Ana dalgacığa uygulanan dönüşüm ve genişleme işlemleri, dalgacık ile analiz edilen sinyalin belirlenen bölümü ile arasındaki bağlantıyı gösteren dalgacık katsayılarını hesaplamak için uygulanır.(Fidan,2006)

Dalgacık dönüşümleri analiz yöntemlerine göre Sürekli Dalgacık Dönüşümü (SDD) ve ayrık Dalgacık Dönüşümü (ADD) olarak ikiye ayrılmaktadır. Sayısal uygulamalarda ayrık dalgacık dönüşümü, analog uygulamalarda da sürekli dalgacık dönüşümü kullanılmaktadır.(Fidan,2006)

4.5.1. Sürekli Dalgacık Dönüşümü

Sürekli Dalgacık Dönüşümü çalışma prensibini açıklayacak olursak; dalgacık dönüşümünde önce kullanılacak olan ana dalgacık seçilir, bu ana dalgacık bizim verilerimizin içinde bulunduğu sinyal ile karşılaştırılır. Dalgacık sinyali ile sinyalin ilk kısmının birbirine ne kadar benzediği bir C katsayısı hesaplanır. Bu katsayı ne kadar büyük ise ana dalgacık ve sinyal o zamanlarda o kadar benzerdir denir. Şekil 4 (A)'da orijinal sinyal ile dalgacık sinyali pencere metodu ile karşılaştırılmıştır ve C değeri hesaplanmıştır (Demir,2008)

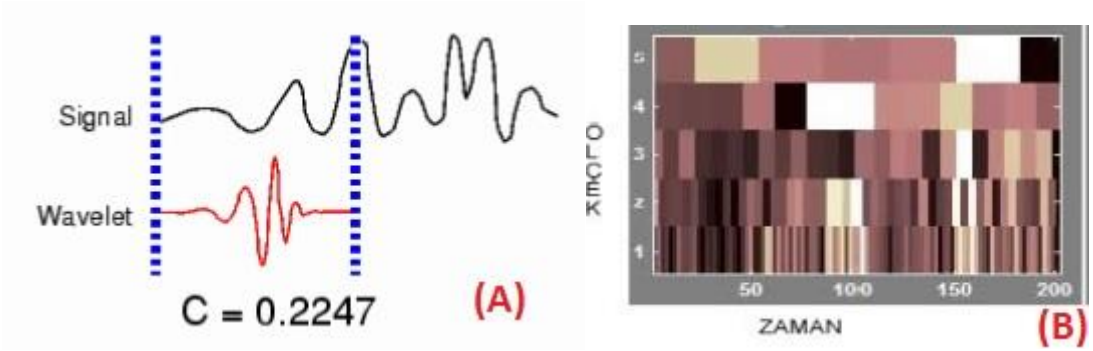


Şekil 4. (A)Sinyal ile küçük ölçekli dalgacığın karşılaştırılması. (B)Sinyal ile kaydırılmış küçük ölçekli dalgacığın karşılaştırılması.

Şekil 4(B)'de dalgacık sinyali sağa doğru kaydırılır ve sinyalin diğer kısmıyla tekrar karşılaştırılır. Biz bu olaya bir pencere sağa kaydırma da diyebiliriz ve C katsayısı tekrar

hesaplanır yani benzeşme katsayısı . Sinyaller birbirlerine ne kadar benzer ise bu katsayı o kadar büyük çıkar.(Demir,2008)

Ana dalgacık bir pencere sağa kaydırma işlemi sinyalin sonuna kadar devam ettirilir ve tüm katsayılar elde edilir. Şekil 4(A)'da görüldüğü gibi ana dalgacığın ölçeği artırılır ve tüm sinyal baştan sona tekrar bölümler halinde (pencereler) sinyalle karşılaştırılır ve katsayılar elde edilir.(Demir,2008)



Şekil 5 (A) Sinyal ile kaydırılmış büyük ölçekli dalgacığın birleşimi. (B) İki boyutlu ölçek- zaman eğrisi.

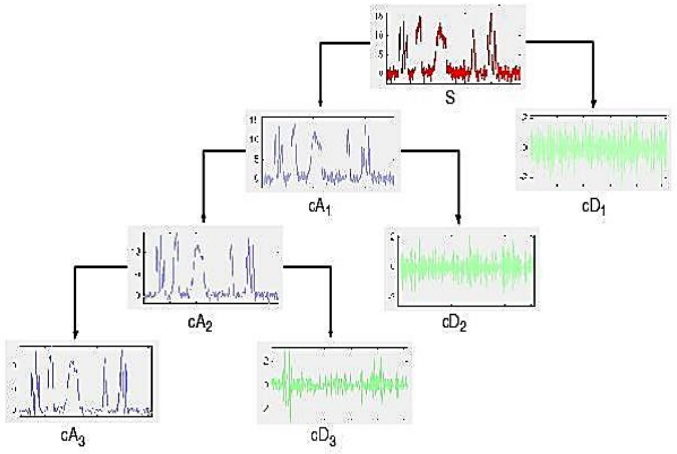
Bu işlemlerin sonunda sinyalin farklı zamanlarında, farklı ölçeklerde ki katsayıları elde ederiz. Bu bize, sinyalin ölçek-zaman eğrisini verir ve aşağıdaki eğrileri elde ederiz (Şekil 4(B)). Sinyalde hiçbir değişme, bozulma söz konusu değildir sadece sinyalin analizi yapılmıştır, tüm bilgiler sinyalde mevcuttur.(Demir,2008)

4.5.2. Ayrık Dalgacık Dönüşümü

Dalgacık dönüşümünde tüm ölçek aralığında analiz yapılırsa çok büyük veri yığınları oluşur ve çok fazla işlem yapmak gerektirir. Bu istenmeyen bir durumdur bu yüzden belirli ölçek grupları tespit eder ve bu aralıkta analizleri yaparsak, bu analize Ayrık Dalgacık Dönüşümü (ADD), (Discrete Wavelet Transform) denir. Pratikte en çok kullanılan yöntem, ölçek ve konum değerleri için ikinin kuvveti olacak şekilde seçilmesidir. Çalışma sistemi ve matematiksel kuram olarak her şey SDD'ndeki gibidir.(Demir,2008)

EEG sinyallerinde ve birçok sinyalde olduğu gibi, düşük frekanslı bileşenler önemli değerler içerirler. Sinyal içindeki önemli bileşenler bu alçak frekans bileşenleridir. Tabi bu iki bileşen sinyalin kendisini oluşturur, yani sinyalin içinde ikisi de vardır. Bu alçak frekans bileşeni, sinyalin tanınma özelliklerini içerir. İnsan beyninin ürettiği dalgalar olan Alfa, Beta, Teta ve Delta sinyalleri de alçak frekans sınırı içerisindedir.(Demir,2008)

Alt bileşenlere ayrıştırma işlemi, aynı sinyal üzerinde birkaç kez, hatta ihtiyaca göre daha fazla gerçekleştirilebilir, genelde alçak geçiren çıkışları yani yaklaşım çıkışları tekrar ADD'ne tabi tutulur. Burada anlamlı bir sinyal çözümü üretilene kadar ADD yani alt bileşenlerine ayrıştırma işlemine devam edilir. Bu işleme dalgacık ayrıştırma ağacı de denir. Şekil 6'da ayrıştırma işlemi blok diyagramlar kullanılarak gösterilmiştir.(Demir,2008)



Şekil 6. Dalgacık ayrıştırma ağacı

Şekil 6'dan da görüldüğü üzere, Dalgacık ayrıştırma ağacı, ana sinyal alt bileşenlerin toplamından oluşur: S = A3 + D3 + D2 + D1. ADD'ne tabi tutulan sinyal (S) hiçbir kayba uğramaz, sadece ana dalgacıkla çarpılarak, bileşenlerine ayrıştırılmıştır. Buradaki A_3 yaklaşım bileşeni, temel bileşendir. Yaklaşımlar, yüksek ağırlıklı, düşük frekanslı bileşenlerdir. Detaylar ise düşük ağırlıklı, yüksek frekanslı bileşenlerdir. Dalgacık dönüşümünde amaç görüldüğü üzere, sinyali yaklaşım ve detaylarına ayrıştırmaktır. Elde edilen yaklaşım tekrar ikinci seviye yaklaşım ve detayına ayrıştırılır, bu işlem istenilen sonuç elde edilene kadar tekrarlanabilir.(Demir,2008)

BÖLÜM 5

VERİ MADENCİLİĞİ (DATA MINING)

5.1. Veri Madenciliği Nedir?

Veri madenciliği kavramını anlayabilmek için işin en başında kelimelerin yalın anlamlarından yola çıkılabilir. Madencilik yeryüzünün gizli ve kıymetli kaynaklarının açığa çıkarılması süreci olup, bu kelimenin veri kelimesi ile ilişkilendirilmesi ise veri yığınları içerisinde ilk bakışta fark edilemeyen kıymetli bilgilerin bulunması ve çıkartılması fikrini uyandırmaktadır.(Irmak,2009)

Bir bilim disiplininin tanımlanması çoğu zaman tartışmalı bir iştir; araştırmacılar kendi çalışma alanlarının kesin aralığı ve sınırları konusunda genellikle aynı fikirde değildirler. Aşağıda, veri madenciliği konusunda literatürde yer alan bazı tanımlar verilmektedir.(Irmak,2009):

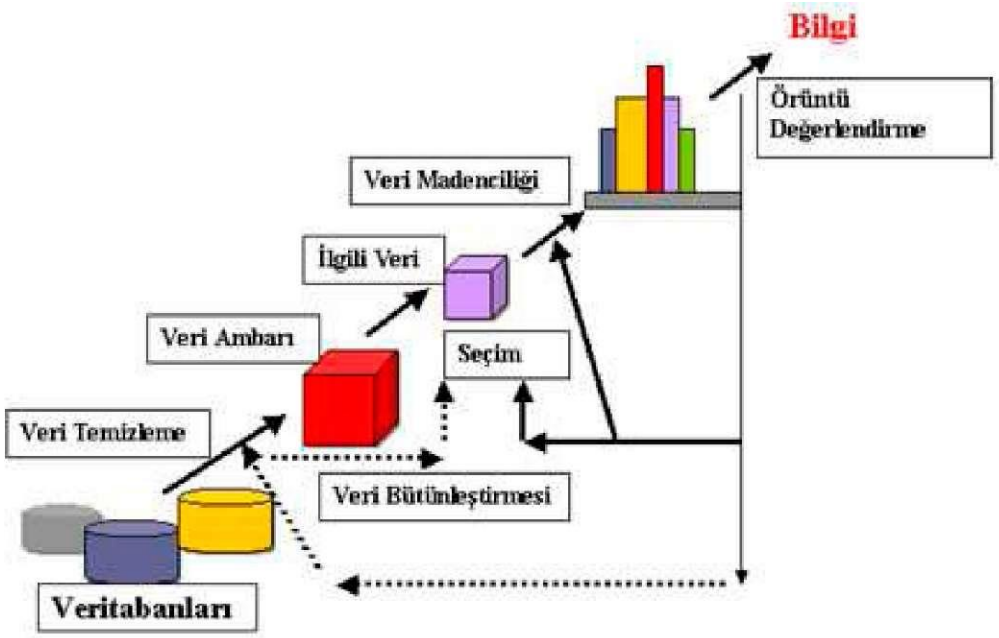
1. Cabena, vd (1998), veri madenciliği, geniş veritabanlarından bilgi çıkarımını hedeflemek için makine öğrenimi, örüntü tanıma, istatistik, veritabanı ve görselleştirme tekniklerini bir araya getiren disiplinler arası bir alandır .
2. Hand (2001), veri madenciliği, (genellikle büyük) gözlemsel veri setlerinin veri sahibi için anlaşılabilir ve faydalı olması amacıyla, tahmin edilemeyen ilişkilerin bulunması için analiz edilmesi ve bunların sözel yollarla özetlenmesidir.
3. Tang ve MacLennan (2005),veri madenciliği, otomatik veya yarı-otomatik biçimlerde verinin analiz edilerek gizli örüntülerin bulunmasıdır.
4. Witten ve Frank (2005), veri madenciliği veride var olan örüntüleri keşfetme sürecidir. Süreç otomatik veya (daha çok) yarı otomatiktir. Keşfedilen örüntüler

anlamlı olmalıdır ve genellikle ekonomik avantaj olmak üzere fayda sağlamalıdır.

5. Tan, vd., (2006), veri madenciliği, büyük veri depolarında faydalı bilgilerin otomatik olarak keşfedilmesi sürecidir.
6. Gartner Group (2007), veri madenciliği, veri ambarlarında depolanan büyük miktarlardaki verinin istatistiksel ve matematiksel tekniklerle birlikte örüntü tanıma teknolojilerinin de kullanılarak incelenmesi yoluyla anlamlı yeni ilişkiler, örüntüler ve eğilimler bulunması sürecidir.

Literatürde yer alan tanımlardan ve veri madenciliği süreçlerinde yaşanılan deneyimlerden yola çıkılarak şu kapsayıcı tanımlama yapılabilir;

Veri madenciliği, büyük veri setlerinde, veri tabanlarında veya veri ambarlarında bulunan veriler arasında var olan, bilinmeyen, klasik yöntemlerle görülemeyen ve sıradan olmayan ilişkileri, örüntüleri, belirli yapıları veya eğilimleri ortaya çıkarmak amacıyla istatistik, matematik, makine öğrenimi ve bilgisayar uygulamaları alanlarının birleşimi tekniklerin kullanılarak analiz edilmesi ve sonuçların anlamlı bir şekilde özetlenmesi ve görselleştirilmesi sürecidir.(Irmak,2009)



Şekil 7. Veri madenciliği aşamaları

Şekil 7.'de veri madenciliğine ait aşamalar verilmiştir. Süreç bu aşmalara uygun olarak sürdürülmelidir. Veri ambarı, organizasyonun ihtiyaçları ile uyumlu büyük miktarlardaki verinin kolay erişilebilir bir yapıda tutulmasını sağlayan bilgisayar tabanlı depolama sistemleridir. Veri ambarları organizasyonel veriye kolay bir şekilde ulaşılmasını sağlayan yapılardır. Veri ambarları 1990'lı yıllarda ortaya çıkmıştır. Veri ambarları veriyi kullanılabilir trend, ilişki ve profillerde sınıflandırmazlar, sadece potansiyel bilgiye sahip veritabanlarıdırlar. Veride saklı bilgiyi keşfetmeyi sağlayan ise veri madenciliği gibi tekniklerdir. Veri ambarından veriyi çekebilmek için hangi verinin gerekli olduğunu ve bu verinin nerede olduğunu tespit etmek önemlidir. Çoğunlukla gerekli veri, farklı sistemler üzerinde olup, farklı formatlardadır. Bu nedenle, ilk aşamada veri temizleme ve düzenleme işlemi gerçekleştirilmelidir. Veri ambarının yaratıcısı W.H. Immon'a göre veri ambarı verinin temizlendiği, birleştirildiği ve yeniden düzenlendiği merkezi ve entegre bir depodur. (Albayrak,2008)

5.2. Veri Madenciliği Tarihçesi

İnsanoğlu geçmişten bugüne her zaman verileri yorumlayıp bilgi edinmeye çalışmıştır ve bunun için çeşitli donanımlar oluşturmuştur. Bu donanımlar bilginin taşınmasını sağlamıştır. Zamanla her alanda bilgi toplanmaya başlanmış ve kronolojik olarak gelişimi Tablo 2.'de özetlenmiştir(Farboudi,2009)

Gelişim Adımları	Cevaplanan Karar Problemi	Kullanılabilen Teknolojiler	Ürün Sağlayıcıları	Karakteristikler
Veri Toplama (1960'lar)	"Benim toplam karım geçen 5 yılda ne kadardı?"	Bilgisayarlar. Teypler. Diskler	IBM. CDC	Geriye Dönük. Statik Veri Dağıtımı
Veri Erişimi (1980'ler)	"İngiltere'de geçen mart ayında birim satışları ne kadardı?"	ilişkisel Veritabanları, SQL. ODBC	Oracle. Sybase. Informix. IBM. Microsoft	Kalıt Düzeyinde Geriye Dönük. Dinamik Veri Dağıtımı
Veri Ambarlama ve Karar Destek Sistemleri (1990'lar)	"ingiltere'de geçen mart ayında birim satışları ne kadardı?"	OLAP. Çok Boyutlu Veritabanı Sistemleri. Veri Ambarları	Pilot. Comshare. Arbot. Cognos. Microstrategy	Çoklu Düzeylerde. Geriye Dönük Dinamik Veri Daıtımı
Veri Madenciliği (Bugün)	Gelecek ay Boston'daki birim satışları muhtemelenne olabilir, niçin?"	ileri Düzeyde Algoritmalar. Çok İşlemcili Bilgisayarlar. Büyük Veritabanları	Pilot. Lockheed. IBM. SGI. SPSS. SAS. Microsoft vs.	Geleceğe Dönük. Proaktif Enformasyon Dağıtımı

Tablo 2. Veri Madenciliği Tarihçesi

5.3. Veri Madenciliğinin Gelişimini Etkileyen Faktörler

Temel olarak veri madenciliği ve gelişimini beş ana faktör etkilemektedir. Bunlar(Kayaalp,2007) :

1. Veri: Veri madenciliğinin gelişmesindeki en önemli faktördür. Son yirmi yılda sayısal verinin hızla artması, veri madenciliğindeki gelişmeleri hızlandırmıştır. Verilerin bu üstel artışına karşın, verilerle uğraşan bilim adamları, mühendisler ve istatistikçilerin sayısı aynı oranda artamamıştır. Problemin çözümü, verileri analiz etme yöntemlerinin ve tekniklerinin geliştirilmesidir.
2. Donanım: Veri madenciliği, sayısal ve istatistiksel olarak büyük veri kümeleri üzerinde yoğun işlemler yapmak için planlanmıştır. Gelişen bellek kapasiteleri ve artan işlemci hızları sayesinde, son birkaç yıla kadar madencilik yapılamayacak kadar büyük veri ambarları üzerinde çalışmaya olanak sağlamıştır.
3. Bilgisayar ağları: Yeni nesil Internet teknolojileri, çok yüksek hızlarda veri transferine izin vermektedir. Bilgisayar ağları aracılığıyla, dağıtık veri tabanlarına ulaşmak, verileri analiz etmek ve farklı algoritmaları kullanmak mümkün olmaktadır. 2000'li yılların başlarında hayal olarak görünen kavramlar günümüzde kullanılabilen teknolojilerdir.
4. Bilimsel hesaplamalar: Günümüz bilim adamları ve mühendisleri, simülasyonu bilimin üçüncü yolu olarak görmekteler. Veri madenciliği ve bilgi keşfi, teori, deney ve simülasyonu birbirine bağlamada önemli bir rol almaktadır.
5. Ticari eğilimler: Günümüzde, işletmeler rekabet ortamında varlıklarını koruyabilmek için daha hızlı hareket etmeli, daha yüksek kalitede hizmet sunmalı, bütün bunları yaparken de minimum maliyeti ve en az insan gücünü göz önünde bulundurmalıdır. Veri madenciliği sayesinde müşterilerin ve müşteri faaliyetlerinin yarattığı fırsatlar daha kolay tespit edilebilmekte ve riskler daha açık görülebilmektedir.

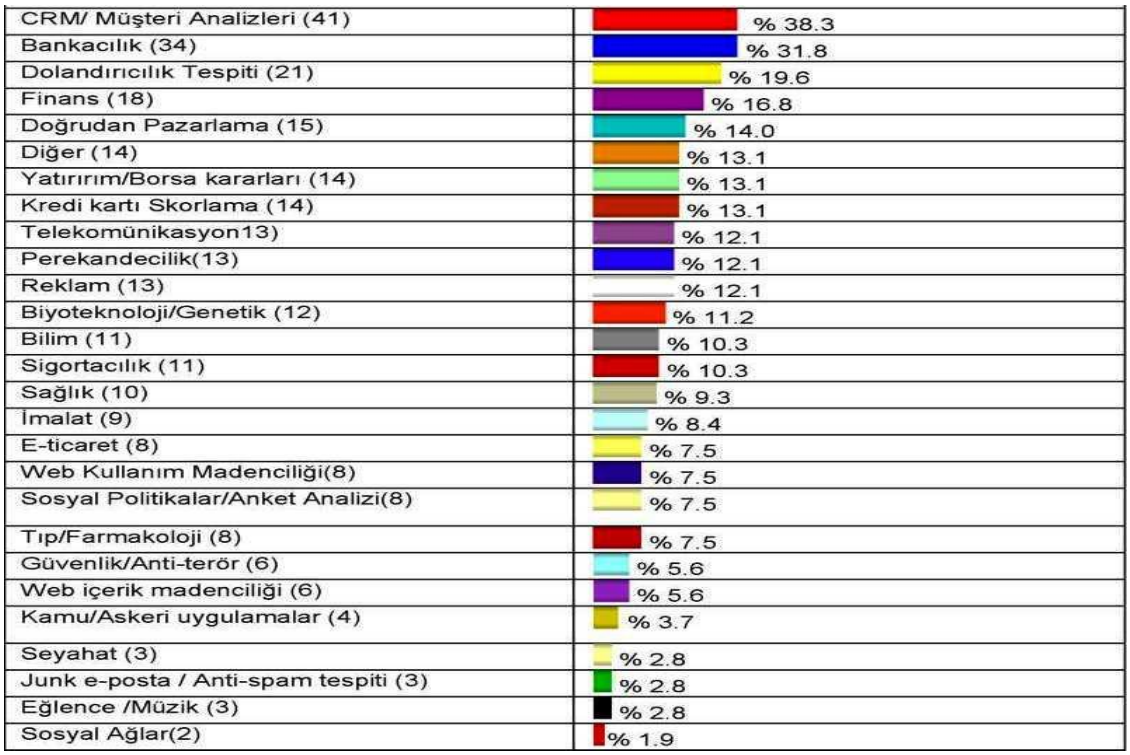
5.4. Veri Madenciliği Uygulama Alanları

Günümüzde veri madenciliği pek çok alanda kullanılmaktadır. Bu alanlardan bankacılık, pazarlama, sigortacılık ve sağlık gibi sektörler başı çekmektedirler. Genel olarak veri madenciliğinin sektörlere göre hangi amaçla kullanıldığı aşağıda sayılmıştır.(Akman,2010)

a. Sağlık ve Farmakoloji: İlaç geliştirme, hastalıkların teşhisi, tedavi sürecinin belirlenmesi alanlarında kullanılmaktadır.
b. Biyoloji: DNA sıra analizi ile hastalıklara neden olan gen sıralamasını belirlemek amacıyla kullanılmaktadır (Microarray veri analizi).
c. Pazarlama Yönetimi: Müşterilerin demografik özellikleri arasındaki bağlantıların kurulmasında, satın alma eğilimlerinin belirlenmesinde, pazarlama kampanyalarının planlanmasında, mevcut müşterilerin elde tutulması ve yeni müşterilerin kazanılması için geliştirilecek pazarlama stratejilerinin oluşturulmasında, pazar sepeti ve çapraz satış analizlerinde, müşteri ilişkileri yönetiminde ve satış tahminlerinde kullanılmaktadır.
d. Bilişim ve Mühendislik: İnternet işlemleri dolandırıcılığının tespit edilmesinde, bilgisayar sistemlerine ve bilgisayar ağlarına yetkisiz girilmesinin tespit edilmesinde, parmak izi ve yüz şeklinden kimlik tespitinde ve yapay zeka uygulamalarında kullanılmaktadır.
e. Bankacılık: Farklı finansal göstergeler arasındaki gizli korelasyonların bulunmasında, kredi kartı dolandırıcılıklarının tespitinde, kredi taleplerinin değerlendirilmesinde, usulsüzlük tespiti, risk analizleri ve risk yönetiminde kullanılmaktadır.
f. Meteoroloji ve Atmosfer Bilimleri: Bölgesel iklim ve yağış haritaları oluşturma, hava tahminleri yapma amacıyla kullanılmaktadır.
g. Sigortacılık: Yeni poliçe talep edecek müşterilerin tahmin edilmesinde, sigorta dolandırıcılıklarının tespitinde ve riskli müşteri tipinin belirlenmesinde kullanılmaktadır.
h. Perakendecilik: Satış noktası veri analizleri, alış-veriş sepeti analizleri, tedarik ve mağaza yerleşim optimizasyonunda kullanılmaktadır.

i. Borsa: Hisse senedi fiyat tahmini, genel piyasa analizleri, alım-satım stratejilerinin optimizasyonunda kullanılmaktadır.
j. Endüstri: Kalite kontrol analizlerinde, lojistik ve üretim süreçlerinin optimizasyonunda kullanılmaktadır.
k. Telekomünikasyon: Kalite ve iyileştirme analizlerinde, hatların yoğunluk tahminlerinde ve telefon dolandırıcılığının tespit edilmesinde kullanılmaktadır.

Veri madenciliği uygulama alanları burada sayılmayan, eğitim/öğretim, güvenlik gibi daha başka birçok alanda da yapılmaktadır. www.kdnuggets.com sitesinin veri madenciliğinin uygulama alanlarını ve oranlarını belirlemek için, web sitesi üzerinden analistlere yönelik olarak yaptığı anket sonuçları Şekil 8'de yer almaktadır. (Akman,2010)

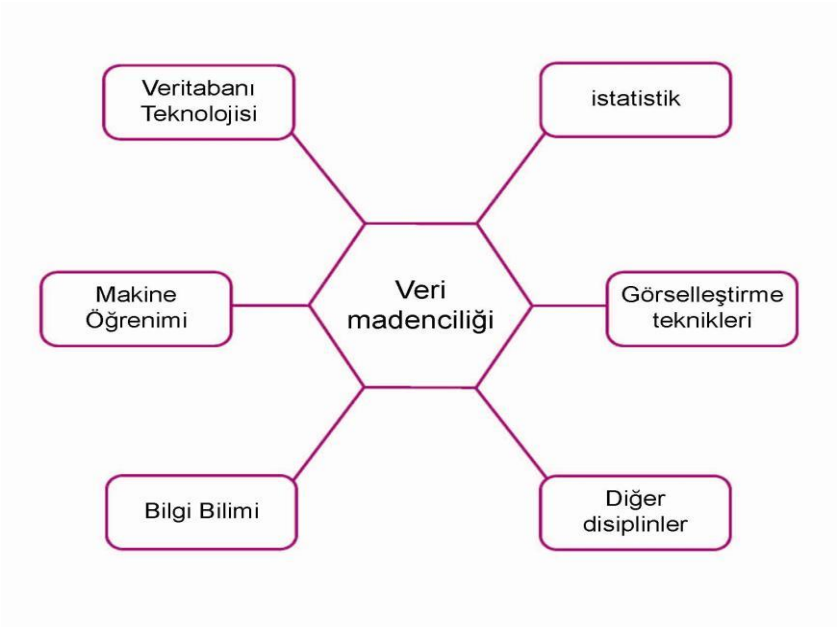


Şekil 8. Veri Madenciliğinin Kullanıldığı Alanlara Göre Dağılımı

5.5. Veri Madenciliği ve Diğer Disiplinler

Veri madenciliği, makine öğrenmesi, örüntü tanıma, veritabanı teknolojileri, istatistik, yapay zeka, uzman sistemler, veri görselleştirme (data visualization)

alanlarının bir kesişim noktası olarak doğmuş ve bu doğrultuda gelişmesini sürdürmektedir. Bu yapı temel olarak Şekil 8'de görüldüğü gibi sembolize edilebilir.(Farboudi,2009)



Şekil 9. Veri Madenciliği ve Diğer Disiplinler

Makine öğrenmesi, örüntü tanıma ve istatistik alanları, veri madenciliği örüntü keşfetme aşamasında, yapay zeka teknolojileri, bulunan örüntülerin yorumlama aşamasında banı teknolojileri, eldeki verileri depolama, ayrıştırma, temizleme, sorgulama işlemleri aşamasında veri görselleştirme ise, raporlama ve insan beyni için anlamlı sembollere çevirme aşamasında yardımcı olmaktadır.(Farboudi,2009)

5.6. Veri Madenciliği Gereksinimleri

Veri madenciliği sürecinin gereksinimleri aşağıdaki belirtilebilir:

1. Erişilebilir veri,
2. Etkin erişim yöntemleri,
3. Açık problem tanımı,
4. Etkin algoritmalar,

5. Yüksek performanslı uygulama sunucusu,
6. Sonuç oluşturmada esnekliktir.

Veri madenciliğinin diğer bir gereksinimi temizlenmiş veridir. Veri madenciliğinde kullanılacak veri yanlış sonuçlar üretmeye yol açabilecek aykırı değerler veriden temizlenmelidir. Doğru veri mevcut değilse ve verinin limitleri bilinmiyorsa; kullanılan yazılımın yanlış sonuçlar üretmesi kaçınılmazdır. (Akbulut,2006)

5.7. Veri Madenciliği Uygulanmasında Karşılaşılan Güçlükler

Veri madenciliği girdi olarak kullanılacak ham veriyi veri tabanlarından alır. Bu da veri tabanlarının dinamik, eksiksiz, geniş ve net veri içermemesi durumunda sorunlar doğurur. Diğer sorunlar da verinin konu ile uyumsuzluğundan doğabilir. Sınıflandırmak gerekirse başlıca sorunlar aşağıdadır:

a. Sınırlı bilgi: Veri tabanları genel olarak veri madenciliği dışındaki amaçlar için tasarlanmışlardır. Öğrenme görevini kolaylaştıracak bazı özellikler bulunmayabilir.

b. Gürültü ve kayıp değerler: Veri özellikleri ya da sınıflarındaki hatalara gürültü adı verilir. Veri tabanlarındaki eksik bilgi ve bu yanlışlardan dolayı veri madenciliği amacına tam olarak ulaşmayabilir. Bu bilgi yanlışlığı, ölçüm hatalarından, ya da öznel yaklaşımdan olabilir.

c. Belirsizlik: Yanlışlıkların şiddeti ve verideki gürültünün derecesi ile ilgilidir. Veri tahmini bir keşif sisteminde önemli bir husustur.

d. Ebat, güncellemeler ve konu dışı sahalar: Veri tabanlarındaki bilgiler, veri eklendikçe ya da silindikçe değişebilir. Veri madenciliği perspektifinden bakıldığında, kuralların hala aynı kalıp kalmadığı ve istikrarlılığı problemi ortaya çıkar. Öğrenme sistemi, kimi verilerin zamanla değişmesine ve keşif

sisteminin verinin zamansızlığına karşın zaman duyarlı olmalıdır.(Albayrak,2008)

5.8. Veri Madenciliği Türleri

En üst düzeyde, veri madenciliğinin 3 türü vardır. Bunlar (Farboudi,2009);

a. Doğrudan veri madenciliği: Bu yöntemde, örneğin hekim, belirli bazı bilgilerin edinilmesiyle (örneğin tıkalı damarların bulunması) ilgilidir. Birçok tıbbi uygulama bu kategoriye girer.
b. Varsayım deneme ve varsayımı daha iyi hale getirme: Kullanıcı bazı varsayımlar üretir ve bunun sistem tarafından doğrulanmasını, değiştirilmesini ve daha uygun olabilecek yenilerinin önerilmesini bekler.
c. Dolaylı ya da saf veri madenciliği: Sistemde hiçbir kısıtlamanın ve kullanıcının bulacağı bilgi ya da ne tür bir bilgiyle ilgilenildiğine ilişkin bir beklentinin olmadığı en genel senaryo budur. Bu aynı zamanda en güç yöntemdir.

5.9. Veri Madenciliği Yazılımları

Veri madenciliği türlerinin uygulamasında ise yazılım firmaları tarafından gerçekleştirilen çeşitli araçlardan yararlanılmaktadır. Bu araçlar aşağıdaki şekilde özetlenebilir:

5.9.1. Spss

Merkezi Chicago‟da bulunan SPSS (Statistical Package for Social Sciences) 1967 yılından bu yana verilerdeki gizli bilgileri keşfetme ve stratejik karar desteği sağlama yönünde ileri analitik çözümler sunmaktadır. SPSS‟in veri madenciliği metadolojisi olarak kabul ettiği CRISP DM (CRoss Industry Standart Processing for Data Mining) %50‟nin üzerinde bir kullanıma sahiptir. İnternet kayıtlarına ve elde edilen verilere gelişmiş veri madenciliği teknikleri uygulanarak, kullanıcılar ile birebir ilişki kurmayı sağlayacak öngörüler elde edilebilir. Bu aşamada SPSS çözümlerine, teknolojilerine ve danışmanlığına başvurarak, güvenilir sonuçlar elde etme yolunda bir

adım atmış oluruz. SPSS veri madenciliği çalışmalarına kendi yeteneğini ve tecrübesini getirerek, öğrenme süresini azaltacak, çalışmalara en hızlı şekilde başlamamızı sağlayacaktır. (Farboudi,2009)

5.9.2. Clementine

Clementine SPSS firmasının veri madenciliği için geliştirmiş olduğu bir modüldür. SPSS istatistiksel bir araçtır. Clementine'nin SPSS içinde bir modül olarak kullanılması kullanıcıların SPSS'in istatistiksel fonksiyonlarından faydalanmasına imkan verir. Yapay sinir ağları ve kural tümevarım yöntemlerini kullanır. Clementine müşteri hizmetleri yönetimi, kimya sektöründe maddelerin aşındırıcılık tahmininde ve bankacılık alanında kredi kartı dolandırıcılıkları gibi konularda kendine uygulama alanı bulmuştur.(Doğan,2007)

5.9.3. Sas

SAS"ın (Statistical Analysis Software) dünya çapında 112 ülkede 44000"i aşkın kullanıcısı bulunmaktadır. Kullanımı SPSS programına göre biraz daha zordur. SAS programında komut yazmak gerekir. Veriler üzerinde gerekli istatistik tekniklerini kullanarak tahmini sonuçlar verir. SAS araştırma, kamu, perakende, sigorta, bankacılık, medya, eğitim ve telekomünikasyon sektörlerinde kullanılmaktadır. (Farboudi,2009)

5.9.4. Enterprise miner

SAS firmasının veri madenciliği aracıdır. SAS'ın Veri ambarı ve ÇAİ (çevrimiçi analitik işleme) araçlarıyla bütünleşik çalışabilmektedir. Enterprise Miner karar ağaçları, yapay sinir ağları, regresyon analizi, 2-aşama modelleri (two-stage models), kümeleme, zaman serileri, ilişkilendirme, vb. veri madenciliği sorgularını ele alabilmektedir. Grafiksel arayüzü sayesinde kullanım kolaylığı sağlar ve kullanıcılar uygulamanın karmaşıklığından habersiz bir şekilde sadece girdi ve çıktılara yoğunlaşabilirler. 2 katmanlı mimariyi kullanır. İstemci bilgisayardaki yazılım

gereksinimi Windows 98, 2000 ve NT'dir. Sunucu bilgisayardaki yazılım gereksinimi Windows 98, 2000 ve NT ile Linux'dır .(Doğan,2007)

5.9.5. Kxen

KXEN (knowledge extraction engine), bir bilgi çıkarma motorudur. Veri madenciliği motorunu kullanmak için veri madenciliği araçlarını veri işleme akışına kolaylıkla ekleyebilen tek programdır .(Farboudi,2009)

5.9.6. Insightful miner

Basit veri madenciliği projelerini yapacak olan sıradan çözümleyiciler için mevcut olan en iyi programlardan biridir. İnsightful miner, S-Plus kullananlar için iyi bir veri madenciliği aracıdır. Çünkü S-Plus fonksiyonlarının tüm kütüphaneleri bu programla kullanılabilir. Bu sayede Statistica Data Miner kadar olmasada zengin istatistiksel çözümleme algoritmalarına sahiptir (Farboudi,2009)

5.9.7. Affinium model

Bu program piyasadaki yanıt (response) modelleme ürünleri içinde kullanımı en kolay programdır. İstatistiksel ve grafiksel temeli zayıf olan veri madencisi ve istatistikçi olmayan kişiler için en iyi programdır. Diğer programların kullanıcılar tarafından elle yapılan bir çok veri madenciliği işlemleri, algoritma seçimi dahil modelleme motoru tarafından otomatik olarak yapılır. Kullanıcı sadece hızlıdan kapsamlıya kadar çözümleme seviyesini seçmelidir. Program mevcut en iyi modeli koruyarak, modelleri az sayıdan çok sayıda algoritma ve parametre kümelerinden kurar. Dört değişken modelleme uygulaması vardır; yanıt modelleyici, çapraz satıcı, müşteri bölümleyici ve müşteri değerlendiriciler. Bu uygulamalar fonksiyon açısından birbirine çok benzer, sadece modeli oluştururken kullanılan terimler açısından farklılık gösterir . (Farboudi,2009)

5.9.8. Statistica Data Miner

STATISTICA Data Miner, KXEN gibi kendine has kategoride bir programdır. Veri madenciliği projesindeki tüm görevleri kolaylaştırmadaki başarısı ve bir çok işlemi başarıyla gerçekleştirmesi açsından eşsizdir. Diğer programların kullanımı daha kolay olabilir (Insightful Miner gibi) ya da daha otomatik olabilirler (Affinium Model ya da KXEN gibi) ancak hiçbir veri madenciliği programı STATISTICA Data Miner kadar fazla araç sunamaz. (Farboudi,2009)

5.9.9. Inlen

İlişkisel veri tabanından aldığı verileri makine öğrenimi teknikleriyle işledikten sonra ortaya çıkan sonuçları Veri tabanına yazmaktadır. Üretilen bilgi kesimi, basit ya da bileşik olabilmektedir(Doğan,2007). INLEN aracında dört işleç vardır:

1. Veri tabanı yönetim işleci: Veri tabanı sorgularını yazmak için geliştirilen bir işleçtir.
2. Bilgi yönetim işleci: Üretilen bilgiyi yönetmek için kullanılır.
3. Bilgi üretim işleci: Veri tabanından bilgi almak ve makine öğrenimi algoritmalarını çağırmak için kullanılır.
4. Makrolar: INLEN işleçlerini bir sırada tanımlamayı ve tek bir işleç gibi kullanabilmeyi sağlar.

5.9.10. DBMiner

Kanada Simon Fraser Üniversitesi tarafından geliştirilen bir sistemdir. DBMiner sınıflama, kümeleme, eşleştirme ve sıra örüntüleri sorgularını yapabilecek veri madenciliği algoritmalarını kullanır. DBMiner çevrimiçi analitik işleme özelliğiyle veri madenciliği algoritmalarının bütünleşik çalışabilme özelliği sayesinde ön plana çıkmaktadır. Bu özellik OLAM (Online Analytical Mining) olarak anılır. DBMiner OLAP ve veri madenciliği yöntemlerini dinamik bir şekilde seçebilme imkânına sahiptir. Kullanıcının kolay kullanabileceği bir ara yüze sahiptir. Bu ara yüz sayesinde

elde edilen sonuçlar çok yönlü bir soyutlama kullanılarak gösterilebilmektedir. (Doğan,2007)

5.9.11. Darwin

Darwin Oracle firmasının veri madenciliği aracıdır. Darwin regresyon ağaçları, karar ağaçları, kümeleme, yapay sinir ağları, Bayesian öğrenme, k-yakınlığında komşuluk gibi birçok algoritmayı destekleyen bir veri madenciliği aracıdır. Paralel sunucular için geliştirilmiş bir veri madenciliği sistemidir. Darwin kullanımı kolay bir ara yüze sahiptir. Darwin veri madenciliği algoritmalarından CART, StarTree, StarNet ve StarMatch'i kullanır. (Doğan,2007)

5.9.12. WEKA

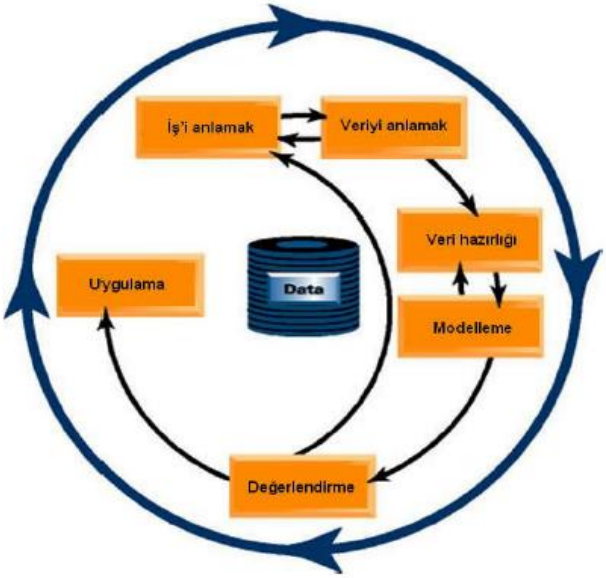
WEKA bir proje olarak başlayıp bugün dünya üzerinde birçok insan tarafından kullanılmaya başlanan bir Veri Madenciliği uygulaması geliştirme programıdır. WEKA java platformu üzerinde geliştirilmiş açık kodlu bir programdır. WEKA çalıştırıldıktan Application menüsünde çalışılabilecek modlar listelenmektedir. Bunlar komut modunda çalışmayı sağlayan Simple CLI, projeyi adım adım görsel ortamda gerçekleştirmeyi sağlayan Explorer ve projeyi sürükle bırak yöntemiyle gerçekleştirmeyi sağlayan KnowledgeFlow seçenekleridir. Arff, Csv, C4.5 formatında bulunan dosyalar WEKA'da import edilebilir. Herhangi bir text dosyadaki verileri WEKA ile işlemek olanaksızdır. Ayrıca Jdbc kullanılarak veritabanına bağlanıp burada da işlemler yapılabilir. WEKA'nın içerisinde Veri İşleme, Veri Sınıflandırma, Veri Kümeleme, Veri İlişkilendirme özellikleri mevcuttur. Bu adımdan sonra yapılacak olan projenin amacına göre açılan sayfadaki uygun tabdaki (Sınıflandırma, Kümeleme, İlişkilendirme) uygun algoritma veya algoritmalar seçilerek veriler üzerine uygulanmakta ve en doğru sonucu veren algoritma seçilebilmektedir. Explorer seçeneği seçildikten sonra üzerinde çalışılacak verilerin seçilmesi, bu veriler üzerinde temizleme ve dönüştürme işlemlerinin gerçekleştirilebilir. (Akademik Bilişim,2012)

Ürün Adı	IBM Intelligent Miner	Oracle Darwin	SAS Enterprise Miner	AngOSS Knowledge Seeker	NCR Teraminer Stats	WEKA
Platform	AIX 4.1, NVS, AS/400, Windows	Windows	Macintosh. Windows, Unix	Windows, Unix	Windows. Unix	Macintosh Windows Unix
Karar Ağacı	X	X	X	X		X
Sinir Ağları	X	X	X		X	
Zaman Serileri	X		X			
Tahmin	X	X	X	X		X
Kümeleme	X		X			X
Birliktelik	X		X			X
Göıselleştirme	X	X	X	X	X	X

Tablo 3.Veri Madenciliği Yazılımlarının Kıyaslanması

5.10. Veri Madenciliği Süreci

Daha detaylı süreç tanımı ise veri madenciliğinin uluslararası düzeyde standardı olarak kabul edilmiş, CRISP-DM (CRoss Industry Process for Data Mining) ile yapılmaktadır. Veri madenciliği projelerinin hızlı, daha verimli ve daha az maliyetli gerçekleştirilmesi için geliştirilmiş olan bu süreç altı adımdan oluşmaktadır. (Argüden ve Erşahin,2008)



Şekil 10. CRİSP-DM Veri Madenciliği Süreci

5.10.1. İşi ve İş Ortamını Anlamak

İlk adım veri madenciliği çalışmasının hangi amaç için yapılacağının net olarak tanımlanmasıdır. Amaç; cevap aranılan sorunun üzerine odaklanmalı, net biçimde ifade edilmeli ve sonuç değerlendirme kriterleri tanımlanmalıdır. (Argüden ve Erşahin,2008)

a. İş hedeflerini algılamak: Çalışmanın temel amacının belirlenmesi ve bu amacın mümkün olduğunca ikincil amaçlardan ayrıştırılarak net olarak tanımlanması aşamasıdır.
b. Durumu değerlendirmek : Veri madenciliğinin temel amacı verim artırmaktır. Bu amaç elde edilecek sonuçlar kadar sürecin kendisi için de geçerlidir. Çalışma sonucunda elde edilecek faydayı değerlendirmek (yanlış kararların maliyetleri ve doğru kararların getirilerine ilişkin öngörüler) önemli bir gerekliliktir.

5.10.2. Veriyi Anlamak

İkinci adım ilk verilerin toplanması, mevcut verilerin uygunluğunun değerlendirilmesi, modeli oluşturmak için gerekli farklı veri ihtiyaçlarının tespit

edilmesi, sahip olunan kayıt sayısının yeterliliği gibi veri kalite ve yeterliliğine yönelik düşünce sürecinden geçilmesi aşamasıdır. (Argüden ve Erşahin,2008)

a. Başlangıç verilerini toplamak: Proje kaynaklarında tanımlanmış olan başlangıç verilerinin toplanması aşamasıdır.
b. Veriyi tanımlamak: Toplanan verinin tanımlanması ve ihtiyaçları karşılama yeterliliğinin değerlendirilmesi aşamasıdır.
c. Veriyi keşfetmek: Başlangıç aşamasında toplanan veriler ile başlangıç hipotezlerinin oluşturulması, limitli bir şekilde veriden çıkarımlar yapılması aşamasıdır.
d. Verinin kalitesini belirlemek: Veri tam mı, doğru mu, hatalar içeriyor mu, hatalar içeriyorsa ne tür hatalar içeriyor, veride eksik bölümler var mı şeklindeki sorular ile verinin kalitesinin tespit edilmesi aşamasıdır.

5.10.3. Veri Hazırlama

Bu aşama başlangıç verilerinin, çalışmalara temel oluşturacak final verilere dönüştürülmesi aşamasıdır. (Argüden ve Erşahin,2008)

a. Veri Setini Tanımlamak: Modelin kurulacağı, tanımlanan soru için gerekli olduğu düşünülen veri setinin (veriler ve bu verilerin toplanacağı veri kaynakları) belirlenmesi aşamasıdır.
b. Veriyi Seçmek: Yapılacak analizde kullanılacak verilerin belirlenmesi aşamasıdır. Değerlendirme sırasında verinin hedefler ile ilişkisine, kalitesine, teknik limitlere dikkat etmek gerekir.
c. Veriyi Temizlemek: Gürültülü ve tutarsız verileri çıkararak verinin kalitesini artırma aşamasıdır. Yanlış girişten veya istisnalardan kaynaklanan verilerin, değerlendirmeden çıkarılması tercih edilir.
d. Veriyi Kurmak: Verileri tanımlayan mevcut değişkenlerde modifikasyonlar yaparak model için daha kullanılır değişken setleri oluşturma aşamasıdır.
e. Veri Birleştirmek: Veri madenciliğinde kullanılacak verilerin farklı kaynaklardan toplanması, doğal olarak veri uyumsuzluklarına neden olmaktadır. Bu

adımda farklı kaynaklardan toplanan verilerin uyumsuzluklar mümkün olduğu ölçüde giderilerek, tek bir veri tabanında toplanması amaçlanır.

f. Veri Formatlamak: Veri seti oluşturulduktan sonra kullanılacak modele göre anlam değişikliği yapmayacak format düzenlemelerinin yapılması aşamasıdır.

5.10.4. Modelleme

Benzer veri madenciliği problemleri için birden çok çözüm tekniği olabilmektedir. Bazı teknikler verilerde spesifik ihtiyaçlar duyarlar. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir. (Argüden ve Erşahin,2008)

a. Model Tekniğini Seçmek: Kullanılacak veri madenciliği fonksiyonun ve algoritmasının belirlenmesi aşamasıdır.
b. Model Test Tasarımı Yapmak: Modeli işletip sonuçları elde etmeye başlamadan önce, modelin kalitesini ve geçerliliğini test etmek gereklidir. Örneğin öngörü fonksiyonlarından sınıflandırma fonksiyonunda hata oranlarını kalite göstergesi olarak kullanılır.
c. Modeli Kurmak: Model için kullanılacak algoritmanın/yöntemin/tekniğin hazırlanan veri üzerinde çalıştırılması aşamasıdır. Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi, bir başka uygulamanın alt parçası olarak da kullanılabilir.
d. Modeli Değerlendirmek: Başarı kriterleri, daha önceki tecrübeler ve test sonuçlarına göre modelin değerlendirilmesi aşamasıdır. Tüm projenin değerlendirilmesinden çok modelin teknik değerlendirilmesi amaçlanır.

5.10.5. Değerlendirme

Bu aşamaya gelindiğinde kurulmuş bir model vardır. Bu aşama, modelin nihai olarak sunulmasından önce modeli yoğun olarak değerlendirilmesi ve iş hedefleri ile uyumlu olup olmadığının kontrol edilmesini amaçlar. (Argüden ve Erşahin,2008)

a. Sonuçları Değerlendirmek: Ön değerlendirme aşamaları modelin geçerliliği ve

uygunluğu konusunda değerlendirme sunarken, bu aşamada modelin iş hedeflerini ne ölçüde karşıladığı değerlendirilir.

b. Süreci Değerlendirmek: Kalite güvence aşamasıdır. Modelin iş hedeflerini karşılamaya yeterli olduğu kararını aldıktan sonra, modelin doğru kurulup kurulmadığı, sadece eldeki verilerden mi yararlanıldığı, gelecekte kullanılabilecek farklı verilerin neler olabileceği gibi konularda değerlendirmeler yapılmalıdır.
c. Gelecek Adımları Planlamak: Projenin geldiği noktanın yeterli olup olmadığı, ek çalışma gerekliliğinin değerlendirilmesi aşamasıdır. Atılacak başka adımlar nelerdir, bunu gerçekleştirebilecek bütçemiz var mı, eğer devam etmek mantıklı ise nereden devam etmeliyiz gibi konularda değerlendirmeler yapılmalıdır.

5.10.6. Yayma

Modelin tamamlanmış olması projenin nihai sonucu değildir. Modelin amacı veriler hakkında bilinenleri artırmak dahi olsa, elde edilen veri kullanılacak biçimde organize edilmeli ve sunulmalıdır. Genellikle gerçek verilerden örneklerin sunulması şeklinde olur. (Argüden ve Erşahin,2008)

a. Yayma Planını Oluşturmak: Sonuçları değerlendirilerek, yayma stratejisinin oluşturulması aşamasıdır.
b. Takip ve Bakımı Planlamak: Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmesini gerektirecektir. Uzun süre yanlış veri kullanarak çalışmanın önüne geçmek için bakım çok önemlidir.
c. Final Raporu Hazırlamak: Yapılan çalışmanın başkaları tarafından da tekrarlanabilirliğini sağlamak ve sonuçlarını karar vericilere aktarabilmek üzere hazırlanan rapordur. Çalışmanın üçüncü taraflarca denetlenebilmesini ve güvenilirliğini sağlamak açısından önem taşır.
d. Projeyi Değerlendirmek: Yapılan çalışmaya dayandırılan kararların ve sonuçların belli bir zaman sonrasında beklentilerle karşılaştırılması ve

gerektiğine çalışmanın yenilenmesi aşamalarını içerir.

5.11. Veri Madenciliği ve Makine Öğrenmesi Kavramı

Bilgi çağında, her türlü karar biriminin en önemli girdisi bilgidir. Ekonomik birim olarak firma açısından rekabet avantajı sayılabilecek bilgiyi incelemek, modellemek ve karar destek sistemlerinde kullanabilmek için güçlü araçlara ihtiyaç vardır. (Albayrak,2008)

Kuramsal tabana oturtulmuş bilgi destek sistemleri, bilişim teknolojilerinde, özellikle son yıl ardaki hem yazılım hem de donanım açısından gelişmelere paralel olarak, büyük miktardaki verileri işleme ve çıktı olarak elde edilen bilgiyi karar destek sistemlerinde kullanabilme olanağına kavuşmuştur .(Albayrak,2008)

İnsan ve makine uzman, veri madenciliğinin birbirini tamamlayan en önemli bileşenleridir. İnsan uzman, veri tabanının hazırlanması, problemin tanımlanması ve hedeflerin belirlenmesinde aktif olan etkendir. Makine uzman ise verileri, hedeflerle uyuşan örüntüleri saptamak amacıyla taramak ve bu örüntülere uygun kuralları öğrenmekten sorumludur. (Albayrak,2008)

Veri madenciliğinde veri elektronik ortamda saklanmakta, verideki örüntüler otomatik olarak aranmakta, doğrulanmakta ve tahmin veya sınıflandırma amacıyla kullanılmak üzere kurallar öğrenilmektedir. İnsan uzmanın bilgi çıkarma sürecindeki diğer bir sorumluluğuysa, makine uzmanın öğrendiği bilginin incelenmesi, sorgulanması, sınanması ve tutarsız ve anlamsız bilgilerin ayıklanması, sorgulama ve sınama sonucunda veri tabanının düzenlenmesi, hatalı bilgi ve verilerin düzeltilmesidir. Bununla birlikte, veri madenciliğinin asıl amacı olarak düşünüldüğünde, insan uzmanı, makine uzmanın bilgi işleme yeteneği ve kapasitesine dayanarak sürece olabildiğince az dâhil etmek ve değerli bir kaynak olarak zamanın daha önemli alanlarda kullanılabilmesini sağlamaktır. (Albayrak,2008)

Makine öğrenmesi, yapay zekânın veri madenciliği ve istatistikle ilintili bir dalıdır. Veri madenciliğinde olduğu gibi, makine öğrenmesinde de, analize konu olan alana ait veriler, örnekler, daha açık bir ifadeyle, örneklere ait özellikleri tanımlayan öznitelik değerleri olarak sisteme iletilir ve makine uzman ya da programın bu verilere ait örüntüyü ya da yapıyı öğrenmesi istenir. Öğrenme sürecinin en önemli özelliği, veri setine eklenen her yeni bilgi için, makine uzmanın önceki veri yapılarını öğrenmiş olarak, örüntüdeki değişmelerin ayırımına varması ve bu değişimlerin getirdiği bilginin kural setine eklemesinin varsayılmış olmasıdır. Böylece, makine uzmanın yeni durumlarla karşılaştığında, en azından potansiyel performansının artırılması amaçlanmaktadır. Makine öğrenmesi yöntemleri, büyük veri setlerinde rahatlıkla kullanılabilen ve daha çok bilgiyi kolaylıkla işlemeye yatkın güçlü yöntemlerdir. Bu buluşsal tahmin yöntemleri, istatistiksel yöntemlerden farklı olarak, sıklıkla güçlü ampirik sonuçlar üretebilmektedir. (Albayrak,2008)

Öğrenme yöntemlerinin çoğunda, "öğrenilen problemin" yapısal tanımları araştırılır [Bu tanımlar, genellikle oldukça karmaşık ve karar ağaçlarının öğrendiğine benzer kural setleri biçiminde ifade edilmektedir. Diğer taraftan, örneğin yapay sinir ağları gibi yöntemler, böyle tanımlar yapmazlar. Makine öğrenmesi uygulamalarında tanımlar, genellikle algoritmanın yeni örnekleri sınıflandırma performansından daha önemlidir; tanımların kuramsal anlamlılık koşulu sağlandığında, yüksek sınıflandırma performansı algoritmanın tercih edilmesinde ikinci düzeyde aranacak ölçüttür. (Albayrak,2008)

Veri madenciliği ve makine öğrenmesi yöntemlerinin temel özellikleri şöyle özetlenebilir (Albayrak,2008):

1. Her iki yöntem de kuramsal olmaktan çok ampirik öğrenme/sınıflandırma problemleriyle ilgilenir.
2. Verideki yapısal örüntüleri bulmak ve tanımlamak, bu veriyi açıklayacak aracı, yani algoritmayı ve ilintili kural kümesini (kümelerini) bulmak ve bu bilgiyi tahmin veya sınıflandırma amacıyla kullanmak makine öğrenmesinin asıl amacıdır.

3. Veri seti örneklerden oluşmaktadır. Her iki yöntemin çıktısı, yeni örnekler üzerinde yapılan tahmin/sınıflandırma ya da öğrenilen kurallardır.

4. Elde edilen bilginin, tanım alanı bağlamında kuramsallaştırılması önemli bir hedeftir.

5. Makine öğrenmesi ve veri madenciliği yöntemlerinde, veriden bilgi çıkarma yarım ya da tam otomatik süreçlerdir.

6. Karar destek sistemleri açısından, her iki yöntemde de, makine uzmanın bilgi işleme kapasitesine dayanılarak, insan uzmanın bu alandaki eksiklileri giderilmek ve rekabet avantajı sağlamak amaçlanmaktadır.

5.12. Veri Madenciliği Modelleri

Veri madenciliği modelleri eldeki veri türlerine ve elde edilecek sonuçların kullanım amacına göre farklılık gösterir. Temelde veri madenciliği iki kategoride incelenir.

- Tanımlayıcı (Descriptive)

- Tahmin edici veya öngörüsel (Predictive)

Tanımlayıcı veri madenciliği, veritabanındaki verinin genel karakterini, mevcut durumu ortaya çıkarmaya yönelik yöntemleri ön plana çıkarır. Bu modellerde karar vermeye rehberlik etmede kullanılabilecek mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. X-Y aralığında geliri, evi ve arabası olan, ayrıca çocukları okulda olan aileler ile çocuğu olmayan ve geliri X-Y aralığından düşük olan ailelerin satın alma örüntülerinin birbirlerine benzerlik gösterdiğini belirlenmesi tanımlayıcı modellere bir örnektir.(Farboudi,2009)

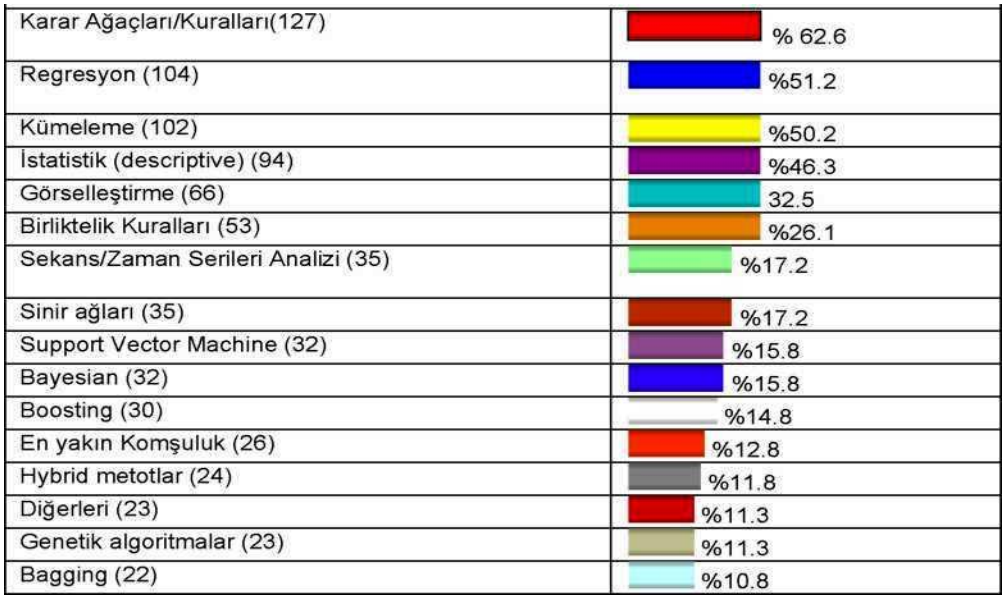
Tahmin edici modellerde ise, veriler geleceğe yönelik tahmin yapma, sonuç çıkarma amaçlı işlemlerde kullanılır. Bu modellerde sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır. Örneğin bir banka önceki dönemlerde vermiş olduğu kredilere ilişkin gerekli tüm

verilere sahip olabilir. Bu verilerde bağımsız değişkenler kredi alan müşterinin özellikleri, bağımlı değişken ise kredinin geri ödenip ödenmediğidir. Bu verilere uygun olarak kurulan model, daha sonraki kredi taleplerinde müşteri özelliklerine göre verilecek olan kredinin geri ödenip ödenmeyeceğinin tahmininde kullanılmaktadır .(Farboudi,2009)

Veri madenciliği teknikleri kullandıkları veri yapılarına ve keşfedebildikleri örüntü biçimlerine göre kategorilere ayrılır. Birçok kaynak veri madenciliği teknikleri için farklı gruplandırmalar yapmıştır. Bunlardan en yaygın kabul göreni J.Han'ın ortaya sürdüğü kategoridir. Bu kategori aşağıdaki gibidir .(Farboudi,2009)

- Tanımlama ve Ayrımlama (Characterization and Discrimination)
- Birliktelik Analizi (Association Analysis)
- Sınıflandırma ve Öngörü (Classification and Prediction)
- Kümeleme Analizi (Cluster Analysis)
- Sıradışık (İstisna) Analizi (Outlier Analysis)
- Evrimsel Analiz (Evolution Analysis)

www.kdnugets.com sitesinin en çok kullanılan veri madenciliği teknikleriyle ilgili 2007 yılında, web sitesi üzerinden analistlere yönelik olarak yaptığı anket sonuçları Şekil 11'de verilmiştir.(Akman,2010)



Şekil 11. En Çok Kullanılan Veri Madenciliği Teknikleri

5.12.1. Tanımlama ve Ayrımlama

Veriler gösterdikleri ortak özelliklere göre genelleştirilmiş sınıflara ayrılabilir. Bir firma müşteri portföyünü alışveriş ortalaması belirli bir miktardan yüksek olan müşterileri " zengin ", diğerlerini ise " orta halli " ya da " fakir " olarak tanımlayabilir. Bu tür genellemeler veri kümesini elemanlarının ortak özellikleri ya da veri kümesinin diğer veri kümeleri ile olan farklılıklarını yansıtacak şekilde yapılabilmektedir. (Farboudi,2009)

a. Tanımlama (Characterization): Bir veri kümesinin elemanlarının genel özelliklerini özetlemek amaçlı kullanılır. Örneğin bir alışveriş merkezinde bu yıl satış oranı %25'in üzerinde artan mallar ifadesi bir Tanımlama işlemidir.

b. Ayrımlama (Discrimination): Bir veri kümesinin diğer bir veri kümesinden farklarını ortaya çıkarma işlemidir. Örneğin bu yıl satış oranı %10 artan mallar ile satış oranı %15 azalan malların karşılaştırılması Ayrımlama tabanlı veri madenciliğidir.

Her iki tür veri madenciliği yöntemi birbirine çok benzer yöntemler kullanırlar. Ayrıca her iki yöntemle elde edilen sonuçlar pasta grafiği, sütun grafiği, eğriler ve çok boyutlu küpler ile sunulurlar.

5.12.2. Birliktelik Analizi

Birliktelik analizi bir veri kümesinde kendiliğinden, sıklıkla gerçekleşen, birlikte ya da ayrı süre içinde alınma, yapılma, oluşma gibi etkileri keşfetme temeline dayanır. Yöntem bankacılık işlemlerinin analizinde ya da Pazar sepeti analizi yönteminde yaygın olarak kullanılır. Pazar sepeti analizi, bir alışveriş sırasında ya da birbirini izleyen alışverişlerde müşterinin hangi mal veya hizmetleri satın alma eğiliminde olduğunu belirlenmesiyle müşteriye daha fazla ürün satılması yollarından biridir. Pazar sepeti analiziyle örneğin müşteriler bira satın aldığında %75 ihtimalle cipste alırlar şeklinde kampanyalar oluşturarak satışlar arttırılabilir. (Farboudi,2009)

Birliktelik analizi yalnızca mal ve hizmetlerin birlikte satın alınması için değil

aynı zamanda hangi koşulları sağlayan müşterilerin hangi ürünleri alacağı hakkında da çözümler getirmektedir. Örneğin bir banka kredi kartı kayıtları incelendiğinde yaşları 20 ile 29 arasında değişen müşterilerden, gelirleri 700 ile 900 YTL arasında değişen müşterilerin bilgisayar satın aldıkları görülmüştür. Bu kural bir gün için birliktelik analizi yöntemlerinde şöyle ifade edilir.(Farboudi,2009)

Yaş (X,"20...29")^{A}Gelir (X,"700...900") ---> alır (X,"bilgisayar")

5.12.3. Sınıflandırma ve Öngörü Yöntemi

Sınıflandırma işlemi insan düşünce yapısına en uygun veri madenciliği yöntemidir. İnsanoğlu çevresindeki nesneleri ve olayları daha iyi anlamak ve başkalarına anlatabilmek için hemen her şeyi sınıflandırma eğilimindedir. Örneğin, insanları davranışlarına göre, hayvanları türlerine göre, evleri görünüşlerine göre sınıflandırmaktadır. (Farboudi,2009)

Veri madenciliğinde sınıflandırma yöntemi, eldeki mevcut verileri önceden belirlenen bir özelliğe göre sınıflara ayırmak ve yeni eklenecek verilerin hangi sınıfa dahil olacağını tayin etme işlemidir. Diğer bir değişle, yeni karşılaşan bir girdinin hangi sınıfa dahil olacağına karar verme işlemidir. (Farboudi,2009)

Sınıflandırma teriminin örüntü tanımında kaynakları vardır. Amaç, yeni bir nesnenin, belirli sınıflar içinde hangi sınıfa ait olup olmadığını belirleyecek bir sınıflayıcı (Classifier) oluşturmaktadır. Sınıflayıcı, gerilme modellerinin özel bir biçimi olarak kabul edilebilir. Sınıflandırma işlemine, bankaların kredi başvuruları için; düşük, orta veya yüksek riskli olarak sınıflandırılması, bir okula yeni gelen öğrencilerin hangi sınıfta eğitim görmesi gerektiğinin belirlenmesi örnekleri verilebilir. (Farboudi,2009)

Öngörü işlemi sınıflandırma işlemine çok benzer. Ancak öngörü işleminde sınıflandırma, gelecek için tahmin edilen belirli bir davranışa ya da belirli bir değere göre yapılır. Öngörü işleminde yapılan sınıflandırmanın doğru olup olmadığını test etmenin tek yolu " bekle ve gör " prensibidir. Öngörü işlemine örnek olarak deprem

tahmini, bir turizm şirketi müşterilerinden hangilerinin bu yaz yurt dışında tatil yapmak isteyeceğinin belirlenmesi verilebilir. (Farboudi,2009)

Bu iki yöntem aslında birbirinden farklı amaçları da olsa aynı tekniklerle kullanıldığı için tek başlık içerisine alabiliriz. Bu iki yöntem arasındaki bağlantı tahmin edilen değerlerin sınıflanmış bir yapıya sahip olmasıdır. (Farboudi,2009)

Sınıflandırma yöntemi iki adımdan oluşmaktadır. İlk adımda gözlenmiş veriler, sınıflandırma algoritması kullanarak sınıflandırma kuralları oluşturulur. İkinci adımda ise oluşturulan sınıflandırma kuralları kullanarak veriler sınıflandırılır. Öngörü modellerinde sürekli veriler alınarak oluşturulan kurallara göre sonuçlandırılır. Bu modellerde kullanılan algoritmalardan bazıları Karar Ağaçları (Decision Tree), Yapay Sinir Ağları (Neural Networks), K-en Yakın Komşu (K- Nearest Neighbour), Genetik Algoritmalar, Naive Bayesian Sınıflandırma, Bellek Tabanlı Nedenleme (Memory Based Reasoning) ve Hatayı Geri Yayma'dır (Backpropagation). Veri madenciliği uygulamasında ayrık nitelik değerlerini tahmin etmek için sınıflandırma ve sürekli nitelik değerlerini tahmin etmek için ise öngörü yapılır. (Farboudi,2009)

5.12.3.1. Karar ağaçları

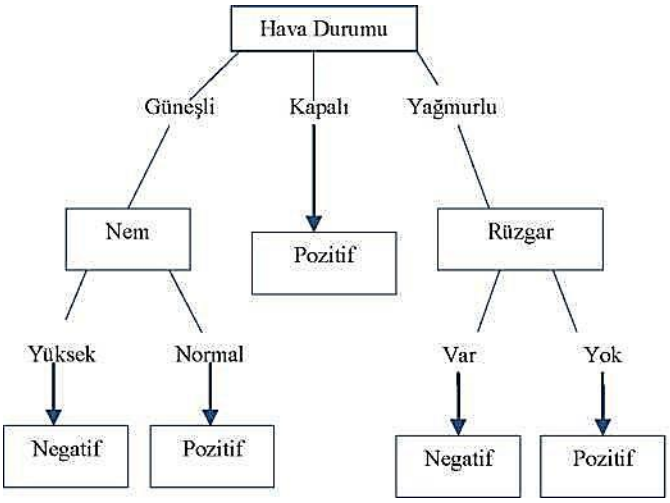
Veri setindeki ilişkileri ve kuralları ortaya çıkarmak için, veri setini sistematik olarak alt kümelere ayıran analitik bir araç olan karar ağaçları , karar alma sürecinde seçenek sayısının fazla olduğu veya ardışık karar almanın söz konusu olduğu durumlarda tercih edilmektedir. Kuruluşlarının ve yorumlanmalarının kolay olması, veri tabanı sistemleri ile kolayca entegre edilebilmeleri ve güvenliklerinin daha iyi olması dolayısıyla yaygın kullanıma sahip olan karar ağaçları problemin çözümünü verdiği gibi, o çözümün tercih edilme sebebini de açıklamaktadır. Tablo 4.'de tenis maçı oynama kararı alınabilmesi için oluşturulacak karar ağacı için hava durumu seti görülmektedir (Esen,2009):

No	Hava Durumu	Sıcaklık	Nem	Rüzgâr	Sınıf
1	Güneşli	Sıcak	Yüksek	Yok	Negatif
2	Güneşli	Sıcak	Yüksek	Var	Negatif
3	Kapalı	Sıcak	Yüksek	Yok	Pozitif
4	Yağmurlu	Ilıman	Yüksek	Yok	Pozitif

5	Yağmurlu	Soğuk	Normal	Yok	Pozitif
6	Yağmurlu	Soğuk	Normal	Var	Negatif
7	Kapalı	Soğuk	Normal	Var	Pozitif
8	Güneşli	Ilıman	Yüksek	Yok	Negatif
9	Güneşli	Soğuk	Normal	Yok	Pozitif
10	Yağmu	Ilıman	Normal	Yok	Pozitif

Tablo 4. Hava Durumu Veri Seti

Tablo 4'deki veri setine ait karar ağacı aşağıdaki gibi oluşturulmaktadır:



Şekil 12.Örnek Bir Karar Ağacı Yapısı

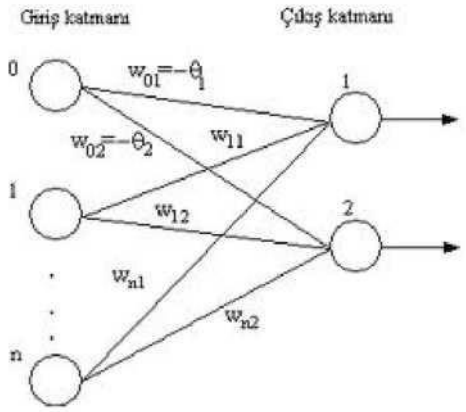
Karar ağaçları; kütüphanelerde zaman ve maliyet kaybını önlemek için kitapların sınıflandırılması, tümör karakterizasyonu, uzay görüntülerinin sınıflandırılması, gelecekteki olayların tahmin edilebilmesi için kurallar oluşturulması, bireylerin kredi geçmişlerinden hareketle kredi kararlarının verilmesi, üretim hattında ürün hatalarına yol açan değişkenlerin belirlenmesi, hangi değişkenlerin satışları etkilediğinin belirlenmesi gibi değişik alanlara uygulanabilmektedir(Esen,2009).

5.12.3.2. Yapay sinir ağları

Yapay sinir ağları, beyin hücreleri olan nöronların çalışma prensibini modelleyen, öğrenebilen algoritmalardır. İleriye yönelik ya da geri besleme alabilen yenilemeli ağlar olmak üzere iki tür yapısı vardır. Yapay sinir ağları (YSA), ağırlıklı

bağlantılar denilen tek yönlü iletişim kanalları sayesinde birbiriyle haberleşen, her biri kendi hafızasına sahip birçok nörondan oluşan paralel ve dağıtık bilgi işleme yapılarıdır. YSA'lar gerçek dünyaya ait ilişkileri tanıyabilir, sınıflandırma, kestirim ve işlev uydurma gibi görevleri yerine getirebilirler. Desen tanıma tekniğinin gerekliliği, gerçek dünya ile bilgisayar ilişkisinin başlaması ile ortaya çıkmıştır. Bu durumda YSA'nın çok güçlü örnek tanıma tekniği olarak ortaya çıkmasına ve gelişmesine neden olmuştur. (Şentürk,2011)

Yapı kurulduktan sonra sinir ağacı eğitilir. Giriş verilerine karşılık çıkış verileri alınır. Bu değer gerçek değerlerle karşılaştırılır ve ağın içerisindeki nöron fonksiyonlarının bu sonuçtaki hata miktarına göre ayarlanması sağlanır. Bu şekilde birçok değer ağa verilir ve ağın eldeki verinin yapısının öğrenilmesi sağlanır. Öğrenme işlemi tamamlandıktan sonra sinir ağı kullanıma hazır hale gelir (Yılmaz, 2006). Bir sinir ağı yapısı aşağıdaki gibidir (Şentürk,2011):



Şekil 13. Örnek Bir Yapay Sinir Ağı Yapısı

Yapay Sinir Ağlarının Türleri (Bilginin Adresi,2012)

* Özdüzenleyici Haritalar (Self-Organizing Maps (SOM))
* Katmanlı Sinir Ağları (Convolutional Neural Networks)
* Radial Basis Function (RBF) Network
* İleri Beslemeli Sinir Ağları (Feedforward Neural Networks (FFNN))
* Yinelenen Sinir Ağları (Recurrent Neural Networks (RNN))

5.12.3.2.1 Radial basis function (RBF) network

İleri beslemeli ağ ile karşılaştırıldığında, radyal taban fonksiyonu (RTF) ağ sonraki en çok kullanılan ağ modelidir. Adından da anlaşılacağı gibi, bu ağ radyal fonksiyonları kullanarak yapar. Radyal Taban Fonksiyon Ağları Bu ağ modeli için bir öğretici giriş yapmaktadır. (Bilginin Adresi,2012)

İleri beslemeli ağlar ve RBF ağları sorunlar aynı tip için kullanılır ve komutları ve seçenekleri çok benzer olabilir. Bu nedenle, yerine birçok bilgi yinelenen, burada sunum ileri beslemeli ağlar mümkünse bölümüne karşılık gelen yere referanslar vererek, kısa olacak. (Bilginin Adresi,2012).

Mathematica sembolik işlem kapasitesi kullanılarak, RBF ağ yapısı, diğer sinir ağ yazılım araçları mümkün olandan çok daha doğal olarak modifiye edilebilir. Bazı özellikler Bölüm 5, İleri Beslemeli Sinir Ağı listelenir. (Bilginin Adresi,2012)

5.12.3.3. Bellek tabanlı yöntemler

Bellek tabanlı yöntemler, günümüzde bilgisayar teknolojilerinin gelişimi ile önem kazanmış bir veri madenciliği tekniğidir. Bu tekniğin içinde en yaygın olarak bilineni k-en yakın komşu algoritmasıdır (Esen,2009).

Sınıflandırmada kullanılan söz konusu algoritmaya göre sınıflandırma sırasında çıkarılan özelliklerden, sınıflandırılmak istenen yeni bireyin daha önceki bireylerden k tanesine yakınlığına bakılmasıdır (Esen,2009).

Örneğin k=3 için yeni bir eleman sınıflandırılmak istendiğinde eski sınıflandırılmış elemanlardan en yakın üç tanesi alınıp, elemanlar hangi sınıfa dâhilse, yeni eleman o sınıfa dâhil edilir. K sayısı, hatalı sınıflandırmayı azaltmak için büyük seçilmeli, önceden sınıflandırılmış örnek sayısına göre küçük tutularak yakın komşuların etkileri arttırılmalıdır (Esen,2009).

Bellek Tabanlı Yöntemler; kredi skorlama sistemleri, kalite ve performans tahminleme, işgücü kullanım stratejileri, sistem performans yönetimi gibi çeşitli uygulama alanlarına sahiptir (Esen,2009).

5.12.3.4. k-en yakın komşu

En yakın komşu tahmin yöntemi, veri madenciliğinde kullanılan en eski yöntemlerden biridir. En yakın komşu yöntemi kümelemeye oldukça benzer ve bir noktanın en yakınındaki noktayla beraber aynı kümeye dâhil olacağı prensibine dayanır (Şentürk,2011).

Ortak en yakın komşu kümesinin iki farklı tanımı vardır ve bu algoritmalar farklı algoritmaların yaratılmasına neden olmuştur. Ortak en yakın komşu kümelemesinin ilk tanımına göre oluşturulan algoritmanın adımları(Şentürk,2011),

1. Her nokta için en yakın komşusu bulunur.
2. Her noktanın kendisi ve komşusundan oluşan çift karşılaştırılır. Karşılaştırmada şu iki nokta dikkate alınır ve üçüncü adıma geçilir.
    a. Eğer iki nokta belli bir komşu sayısından fazla ise,
    b. İki nokta da birbirinin k-en yakın komşularına dâhilse,
3. Bu iki nokta ve bulundukları kümeler birleştirilir.

Farklı yoğunluktaki kümeleri ayırt edebilen bu kümeleme tekniğinin geçişli bir yapısı vardır. Eğer p noktasından q noktasına ortak komşu varsa ve aynı zamanda q noktasının r noktasına ortak komşusu varsa, p, q ve r noktaları aynı kümeye atanırlar. Böylece, algoritma farklı boyutlarda ve şekillerdeki kümeleri de fark edebilir. Ortak en yakın komşu kümesinin ikinci tanımı, noktaların en yakın komşularının sıralanarak sıra sayılarının toplanmasıyla elde edilen ortak komşuluk değerine dayanır(Şentürk,2011).

Ortak komşuluk değerinin birbirine yakın olduğu noktalar ortak en yakın komşu sayılırlar. Bu tanımdan yola çıkarak oluşturulan algoritmanın adımları(Şentürk,2011);

1. Her nokta için k-en yakın komşular bulunur.
2. Her noktanın bütün k-en yakın komşuları için nokta ve komşusu arasındaki ortak komşuluk değeri hesaplanır.
3. En düşük ortak komşuluk değerine sahip olan kümeler birleştirilir.
4. Üçüncü adım istenen küme sayısı elde edilinceye ya da kümeler artık birleştirilinceye kadar devam edilir.

Bu tanıma göre yapılan kümeleme, farklı yoğunluk, büyüklük ve şekillerdeki kümeleri fark edebilir.

5.12.3.5. Genetik algoritmalar

Temellerini Darwin'in evrim teorisinden alan genetik algoritmalar (GA), karmaşık çok boyutlu arama uzayında en iyinin hayatta kalması ilkesine göre bütünsel en iyi çözümü aramaktadır (Esen,2009).

İlk defa Michigan Üniversitesi'nde John Holland ve çalışma arkadaşları tarafından geliştirilmiştir. Holland, araştırmalarını, arama ve optimumu bulma için, doğal seçme ve genetik evrimden yola çıkarak yapmıştır. 1975 yılında yayınlanan "Adaptation in Natural and Artificial Systems" kitabında mekanik öğrenme konusunda görüşler belirtmiş, Darwin'in evrim teorisinde belirttiği canlılarda yaşanan genetik süreci bilgisayar ortamında gerçekleştirmeyi düşünmüştür(Esen,2009).

Yıllar boyu süregelen genetik mühendisliği ve biyoloji alanında yapılan çalışmalar sonucu ortaya atılmış bir teknik olan GA'da amaç, canlıların yapılarında var olan bir takım özellikleri sanal ortamda taklit ederek modeller geliştirmektir. Otomatik programlama, öğrenme kabiliyetli makineler, ekonomi, ekoloji, planlama, üretim hattı yerleşimi gibi alanlarda da uygulanan GA, veri madenciliği çalışmalarında da başarılı bir şekilde kullanılmaktadır. GA işlemleri adımları kısaca şöyle açıklanabilir (Esen,2009):

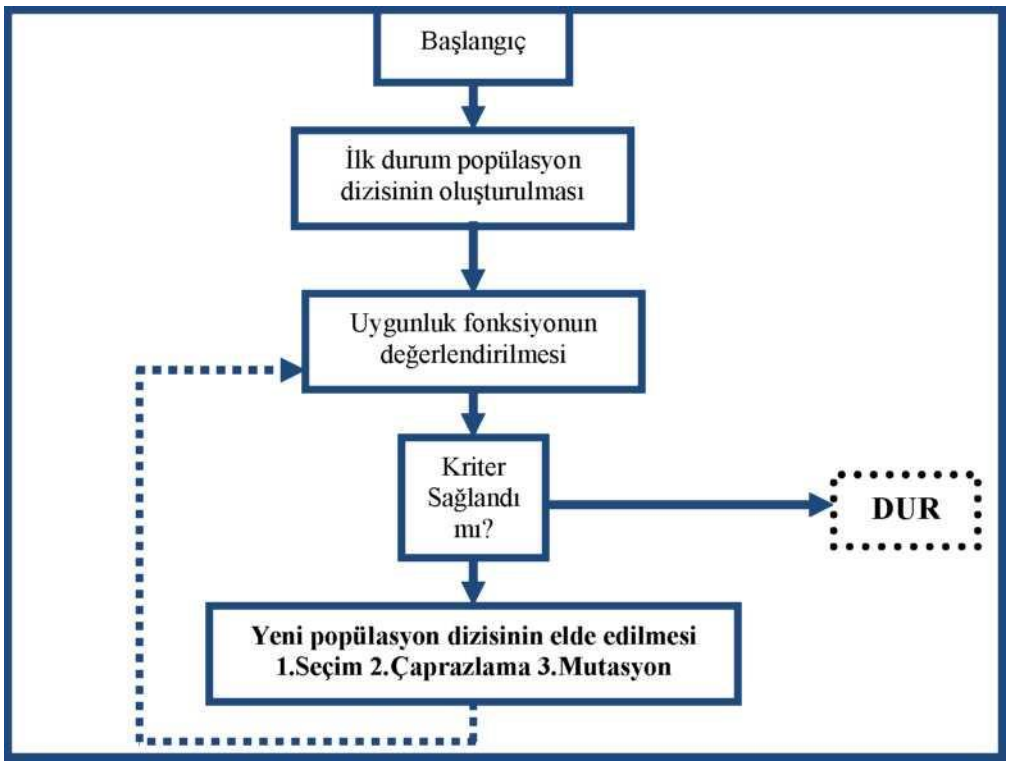
a. Kodlama Süreci: Arama uzayındaki tüm mümkün çözümleri dizi olarak kodlanması şeklinde tanımlanan söz konusu süreç, probleme özgü bilgilerin genetik algoritmanın kullanacağı şekle çevrilmesidir.
b. İlk Popülâsyonun Oluşturulması: Olası çözümlerin kodlandığı bir çözüm grubu oluşturulur. Çözüm grubuna popülasyon, kodlara da kromozom adı verilir.
c. Uygunluk Değerinin Hesaplanması: Bir çözüm kuşağı oluşturulduktan sonra ilk olarak çözümdeki değerlerin uygunluk değerini hesaplamak gereklidir. Uygunluk değeri, yaşama ve çoğalma şansını, diğer kuşaklardaki temsil edilme

oranını göstermektedir.

d. Çoğalma İşlemi: Bir grup dizinin belirli bir olasılık değerine göre tesadüfi olarak seçilerek uygunluk değerinin hesaplanmasından sonra, değeri en yüksek olan çözüm kuşağı seçilip, yeni çözümler üretilmesi aşamasıdır.

e. Çaprazlama ve Mutasyon İşlemleri: Çaprazlama, çözüm havuzunda bulunan iyi uygunluk değerine sahip iki kromozomun iyi özelliklerini birleştirerek daha iyi sonuçlar elde etme sürecidir. Mutasyon ise, mevcut kromozomlardan yeni kromozomlar elde edilmesi sürecidir. Kromozomlarda tesadüfi olarak belirlenen bir noktada yapılan değişiklik olarak da tanımlanan mutasyon, bir daha elde edilemeyebilir iyi bir çözümün kaybına karşı koruma sağlamaktadır.

f. Yeni Kuşağın Oluşması ve Döngünün Durdurulması: Çoğalma, çaprazlama ve mutasyon işlemlerinden sonra yeni kuşak oluşmakta, söz konusu kuşak bir sonraki kuşağın ebeveynleri olmaktadır. Bu süreç bir hedefe ulaşılıncaya kadar veya önceden belirlenen kuşak sayısı kadar ya da başka bir durdurma kriteri sağlanana kadar devam etmektedir.



Şekil 14.Genetik Algoritma Süreci

Finansal problemler, montaj hattı dengeleme problemleri, çizelgeleme problemleri, portföy optimizasyonu, tesis yerleşim problemleri, atama ve taşıma problemleri, gezgin satıcı problemleri gibi geniş bir uygulama alanına sahip olan genetik algoritmalar; bilgisayar ağlarının tasarımı, mekanik öğrenme ve ekonomik-sosyal sistem modelleri geliştirilmesi amacıyla da kullanılmaktadır(Esen,2009).

5.12.3.6. Naive Bayes ile sınıflandırma yöntemi

Eldeki verilerin belirlenmiş olan sınıflara ait olma olasılıklarını öngörür. İstatistikteki Bayes teoremine dayanır. Bu teorem; belirsizlik taşıyan herhangi bir durumun modelinin oluşturularak, bu durumla ilgili evrensel doğrular ve gerçekçi gözlemler doğrultusunda belli sonuçlar elde edilmesine olanak sağlar. Belirsizlik taşıyan durumlarda karar verme konusunda çok kullanışlıdır (Altıntaş,2010).

Bayes yöntemi koşullu olasılık durumları ile ilgilidir. Her hangi bir koşullu olasılık durumu aşağıda belirtilen eşitlik şeklinde tanımlanır :

$$P(X=x \mid Y=y) = R \qquad (5.1)$$

Bu ifade; "Eğer Y = y doğru ise, X = x olma olasılığı R'dir" anlamına gelmektedir. X ve Y 'nin alabileceği değerlerin her kombinasyonu için koşullu olasılıkları belirleyen tabloya koşullu olasılık dağılımı adı verilir ve P(X|Y) ile ifade edilir. Bayes Kuralı şu şekilde tanımlanır (Altıntaş,2010).

$$P(XIY) = P(YIX) \times P(X) / P(Y) \qquad (5.2)$$

Bu ifade; Y'nin gerçekleşmesi halinde X'in gerçekleşme ihtimalinin ne olduğunu belirtmektedir. Bu değeri bulabilmek için "X'in gerçekleştiği durumlarda Y'nin gerçekleşme ihtimali" ile X'in gerçekleşme ihtimalini çarpmak ve bunu Y'nin gerçekleşme ihtimaline bölmek gereklidir (Altıntaş,2010).

Belirsiz durumlarda tahmin ve sınıflandırma yapmak için kullanılır. En önemli dezavantajı değişkenler arası ilişkinin modellenmemesi ve değişkenlerin birbirinden tamamen bağımsız olduğu varsayımıdır (Altıntaş,2010).

5.12.3.7. Bayes Ağları (Bayes Net) yöntemi

Bir Bayes Ağı, bir grup rastsal değişkeni ve onların koşullu bağlılıklarını yönlü dönüşsüz grafiklerle ifade eden, olasılığa dayalı bir grafiksel modeldir. Örneğin, bir Bayes ağı, hastalıklar ve belirtileri arasında olası ilişkileri gösterebilir. Verilen belirtiler ile ağ, değişik hastalıkların varlık olasılıklarını hesaplayabilir(Şentürk,2011).

Daha resmi bir tanımla, Bayes ağları, düğümleri (nodes) rastsal değişkenleri Bayes bağlamında gösteren yönlü dönüşsüz grafiklerdir. Bayes bağlamından kasıt, gözlemlenen miktarlar, gizli değişkenler, bilinmeyen parametreler veya hipotezler olabilir. Kenarlar koşullu bağlılıkları, bağlantılı olmayan düğümler ise koşullu olarak birbirinden bağımsız olan değişkenleri gösterir. Her bir düğüm, düğümün alt/üst değişkenleri için bir grup değeri girdi olarak alan olasılık fonksiyonları ile ilişkilendirirler ve düğüm tarafından gösterilen değişkenin olasılığını verirler(Şentürk,2011).

Bayes ağlarının daha geniş hali belirsiz karar ağaçlarıdır (uncertain decision trees). Bayes ağlarını daha iyi anlayabilmek için Bayes teorilerini hatırlatmakta yarar vardır(Şentürk,2011):

$$P(Aİ\backslash B)= \frac{p(B\backslash Ai).p(Ai)}{p(B\backslash A_1)+\cdots+p(B\backslash A_n)} = \frac{p(B.A_i)}{p(B)} \qquad (5.3)$$

(5.3) eşitliği Bayes teorimidir. P(Ai) terimine Ai için önsel olasılık veya marjinal olasılık denir. Önseldir, çünkü B olayı hakkında önceden herhangi bir bilgi içermemektedir. P(Ai\B) terimi, verilmiş bir B olayı için Ai'nin koşullu olasılığıdır. P(B\Ai), verilmiş bir Ai için B'nin koşullu olasılığıdır. Son olarak P(B) terimi ise B olayı için önsel olasılıktır(Şentürk,2011).

5.12.3.8. Vaka tabanlı nedenleşme yöntemi

Sınıflandırmak için yeni bir vaka verildiğinde, vaka tabanlı nedenleyici benzer bir eğitim vakasının olup olmadığını kontrol eden, bulduğunda aynı çözümü uygulayan, böyle bir vaka bulunamazsa, yeni vakanın bileşenlerine benzeyen eğitim vakalarını araştıran bir yöntemdir. (Farboudi,2009)

5.12.3.9. Kaba küme teorisi yöntemi

Kaba küme teorisi sınıflandırmada, kesin olmayan ya da gürültülü veri içindeki yapısal ilişkileri keşfetmek için kullanılmaktadır. Kesikli değerli niteliklere uygulanır. Sürekli değerli nitelikler, kaba küme uygulanmadan önce kesiklileştirilmelidir (Farboudi,2009)

5.12.4. Kümeleme analizi

Kümeleme işlemi sınıflandırma ve öngörü işleminin aksine, veri kümesini önceden sınıflara ayırmaz, bunu yerine veriler dağılımına göre irdelenerek doğal sınıflandırmalar oluşturulur. Kümeleme işleminin sınıflandırma işleminden en önemli farkı önceden belirlenmiş sınıflar ya da sınıf tanımları (etiketleri) olmamasıdır. Bu yüzden kümeleme işlemi gözetimsiz (Unsupervised) veri madenciliği yöntemidir. Kümeleme işlemi sonunda elde edilen kümeler kullanılan yöntemlerin giriş parametrelerine bağımlı olsa da, giriş parametrelerinden bağımsız kümeleme teknikleri geliştirme çalışmaları sürmektedir (Farboudi,2009).

Başka bir değişle bu modelde temel amaç, çok boyutlu veriler içindeki doğal grupları (kümeleri) bulmaktır. Nesneler, birbirine benzemiyorsa (aynı ölçüye göre) ve başka kümelerdeki nesnelere benzemiyorsa, aynı kümeye alınabilir. Kümelerde, alan bilgisinin kümeleme mekanizmalarıyla nasıl birleştirileceğiyle ilgilenilir. Nüfus bilimi ve astronomi alanında kullanımları yaygındır. (Farboudi,2009)

Kümeleme işleminde temel prensip, sınıf içi benzerliği maksimum, sınıflar arası benzerliği minimum yapmaktır. Bir kümeleme yönteminin kalitesi bu prensibi sağlaması ile doğru orantıdadır. Kümeleme analizi sadece veri madenciliğinde değil, örüntü tanıma, görüntü işlemi, coğrafi bilgi sistemleri gibi birçok alanda yoğun olarak kullanılmaktadır. (Farboudi,2009)

Literatürde birçok kümeleme algoritması bulunmaktadır. Kullanılacak olan kümeleme algoritmasının seçimi, veri tipine ve çalışmanın amacına bağlıdır.Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir. (Farboudi,2009)

a. Bölme Yöntemleri (Partitioning Methods)

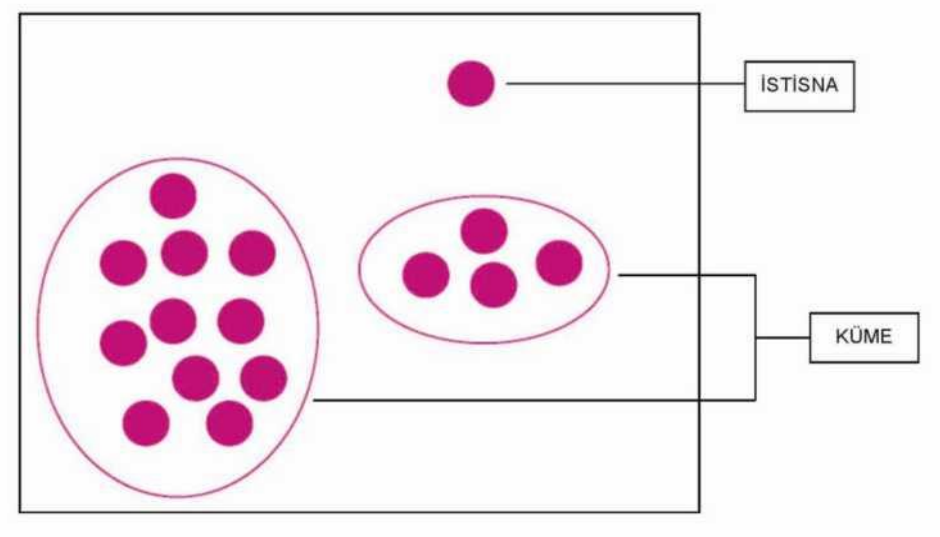
b. Hiyerarşik Yöntemler (Hierarchical Methods)
c. Yoğunluk Tabanlı Yöntemler (Density-Based Methods)
d. Izgara Tabanlı Yöntemler (Grid-Based Methods)
e. Model Tabanlı Yöntemler (Model-Based Methods).

5.12.5. İstisna analizi

Bir veri kümesinde verilerin genel davranışlarından veya veri dağılım modellerinden farklılık gösteren nesnelere sıra dışı (Outlier) denir. Bir çok veri madenciliği yöntemi istisnaları gürültü veya aşırı durumlar olarak görür, bu yüzden dikkate almaz. Fakat bazı durumlarda istisna noktalar değerlerine göre çok daha fazla bilgi içerir. Örneğin kredi kartı veya sigorta sahtekarlıkların tespitinde, tıp biliminde yeni bir hastalığın başlangıcını tespit etmede istisnalar analiz edilir. İstisna analizinde iki yöntem söz konusudur (Farboudi,2009).

a. İstatistik Tabanlı Yöntem: Dağılım analizi ya da standart sapma hesabı gibi istatistiksel yöntemlerle istisna olabilecek noktalar tespit edilir, fakat çok büyük veri yığınlarında yoğun hesaplama gücü gerektirdikleri için performansları sınırlıdır.

b. Yoğunluk Tabanlı Yöntem: Bu yöntemde her noktanın çevresindeki komşuları ile olan yakınlığı hesaplanır. Yakınlık hesaplamada genelde Öklit uzaklığı kullanılsa da veri türüne göre yakınlık hesaplama yöntemi farklılık gösterebilir. Bu yöntemin temel prensibi " yeterince komşu olmayan noktaları " tespit etmektir. Bu durum şekil 15.'de görülmektedir.



Şekil 15. Şekil 16. İstisna Analizi ve Küme Oluşumları

5.12.6. Evrimsel analiz

Evrimsel analiz, zamanla davranışları değişen nesnelerin düzenlilik (Regularity) ya da eğilimlerini (Trends) ortaya çıkarmayı amaçlar. Evrimsel analiz tanımlama, ayrımlama, birliktelik analizi, sınıflama ve kümeleme metotlarını içerse de asıl amacı verinin zaman ile olan ilişkisini ortaya çıkarmaktır. Bunun için zaman serileri (time series), ardışık ve periyodiklik örüntüsü bulma, benzerlik analizi gibi metotları kullanır (Farboudi,2009).

Evrimsel analiz veri madenciliği kategorileri içine dahil edilse de birçok kaynakta bağımsız bir kategori olarak yer almaktadır (Han & Kamber, 2006). Evrimsel analizin kullanıldığı her bir yöntem evrimsel analiz adı altında değil, kendi başına bağımsız bir teknik olarak kabul görmektedir(Farboudi,2009).

Bulunan örüntünün önemli olması için:

- İnsanlar tarafından kolayca anlaşılabilir,
- Test verisi veya yeni veriler üzerinde belli oranda geçerli,
- Yeni, yararlı ve kullanabilir olması gerekir.

5.13. Tıbbi Verilerle Veri Madenciliği

Tıp alanında bilginin kullanım şeklinde meydana gelen değişiklikler sağlık bakım hizmetini verenleri etkilemiştir, sağlık bakım hizmetinin verilmesinde bilgisayar kullanımı, bilginin paylaşım-ekip yaklaşımını, veri ve bilgi temelli uygulama gibi kavramlar yaygınlaşmaya başlamıştır. Bilgisayarlar hasta bakım hizmetlerinin destekleme, sağlık bakım hizmetlerinin kalitesinin değerlendirilmesi gibi doğrudan sağlık bakım hizmetlerinin sunulmasında kullanılmasının yanı sıra, teşhis koyma, tedavi süreçleri, yönetim, planlama ve tıbbi araştırmalar gibi yönetsel ve akademik fonksiyonların yerine getirilmesinde daha fazla kullanılmaya başlanılmıştır. (Aslan,2008)

Tıp alanında bulunan mevcut veri oldukça fazla ve hayati öneme sahiptir. Hastane bilgi sistemleri sayesinde bu veriler düzenli olarak tutulmaktadır. Hayati öneme sahip olan bu verilerden daha fazla yararlanmak mümkündür. Hastane Bilgi sistemlerinden veya diğer tıbbi veri toplayan sistemlerden alınan veriler üzerinde yapılan veri madenciliği çalışmaları hem uzmanlar için hem hastane yönetimi için hem de hastaların daha kaliteli bir hizmet almalarında etkin rol oynamaktadır (Aslan,2008)

5.13.1. Tıbbi verilerle veri ambarının oluşturulması

Tıbbi veriler üzerinde çalışma yapmak bu verileri iyi tanımakla mümkündür. Tıbbi verilerin yorumlanmasında uzmanların önemli bir rolü vardır. Bu nedenle uzman görüşleri ile işlemler arasındaki bağlantı iyi kurulmalıdır. Tıp alanında belirli bir standardın olmayışı ve var olan standartlar arasında tam bir uyumun olmaması nedeniyle, bu alanında bir veri ambarının oluşturulması oldukça zor bir işlemdir. Bu nedenle veri ambarı oluştururken farklı kaynaklardan toplanan veriler arasındaki standart uyumu da göz önüne alınmalıdır. (Aslan,2008)

Bunu yanı sıra tıp alanındaki terimlerin hem karışık hem de birbirine yaklaşık olması da veri ambarı oluşumunu negatif yönde etkilemektedir. Tıp alanındaki veri

genellikle farklı kaynaklarda toplanmaktadır. Örneğin hastanın laboratuvar İle ilgili verileri ile hastanın teşhis bilgileri farklı kaynaklarda ve farklı şekillerde tutulmaktadır. (Aslan,2008)

5.13.2. Tıbbi verilerin özellikleri

Tıbbi verilerin veri madenciliğinde etkin olarak kullanılabilmesi için bu verilerin özelliklerinin çok iyi bilinmesi ve yapılan işlemlerin bu özellikleri dikkate alınarak yapılması gerekmektedir. Bunlar (Aslan,2008):

1. Çok sayıda yordam, görüntülemeyi bir tanı aracı olarak kullanmaktadır. Bu nedenle, görüntülerden oluşan veri tabanlarında etkin bir veri madenciliği gerçekleştirebilmek için yöntemler geliştirmek gerekmektedir. Bu da sayısal veritabanlarındaki veri madenciliğinden hem daha farklı, hem de daha zordur.
2. Tıbbi veritabanları, her zaman heterojendir. Örneğin, bir organa ait görüntü, her zaman, hekimin yorumu (klinik izlenim, tanı) gibi, başka klinik bilgilerle bir aradadır. Bu ise, bu tür verilerin çözümlemesi için yeni araçlar ve yüksek kapasiteli veri depolama aygıtları gerektirir.
3. Hekimler, görüntüler, sinyaller ya da diğer klinik bilgilerle ilgili yorumlarını, standartlaştırılması çok güç olan serbest metinler olarak yazmaktadır. Örneğin aynı hastalık açıklanırken bile farklı adla kullanılmaktadır. Tıbbi kavramlar arasındaki ilişkileri açıklamak için de farklı dilbilgisi yapıları kullanılmaktadır.
4. Verinin sahibi kimdir? Her yıl milyarlarca baytlık tıbbi veri üretilmekte, bu veriler, heterojen veritabanlarında saklanmakta ve sağlık kurumları tarafından ortak bir biçim ya da ilkeye göre düzenlenmeden yayılmaktadır. Hastaya ait bilgilerin sahibinin kim olduğu sorusu hâlâ yanıtlanmamıştır. Hastalar mı, hekimler mi, sigorta kurumları mı?
5. Hekimlere ya da sağlık hizmeti veren diğer kişilere yönelik davalardır. Örneğin, hekimler gereksiz testler yüzünden dava konusu olabilmektedir. Böyle bir ortamda da hekimler ya da sağlık hizmeti veren diğer kişiler, verilerini bu veriler üzerinde çalışacak olan kişi ya da kurumlara aktarmakta gönülsüz davranmaktadır.

6. Gizlilik, güvenlik ve hasta mahremiyetiyle ilgili vb. konular. Veri internet üzerinden elektronik olarak aktarıldığından, güvenli değildir. Bu nedenle veri bir kurum içinde bir birimden diğerine aktarılacak olsa da dikkatli bir biçimde şifrelenmelidir.

BÖLÜM 6

GERÇEKLEŞTİRİLEN UYGULAMA

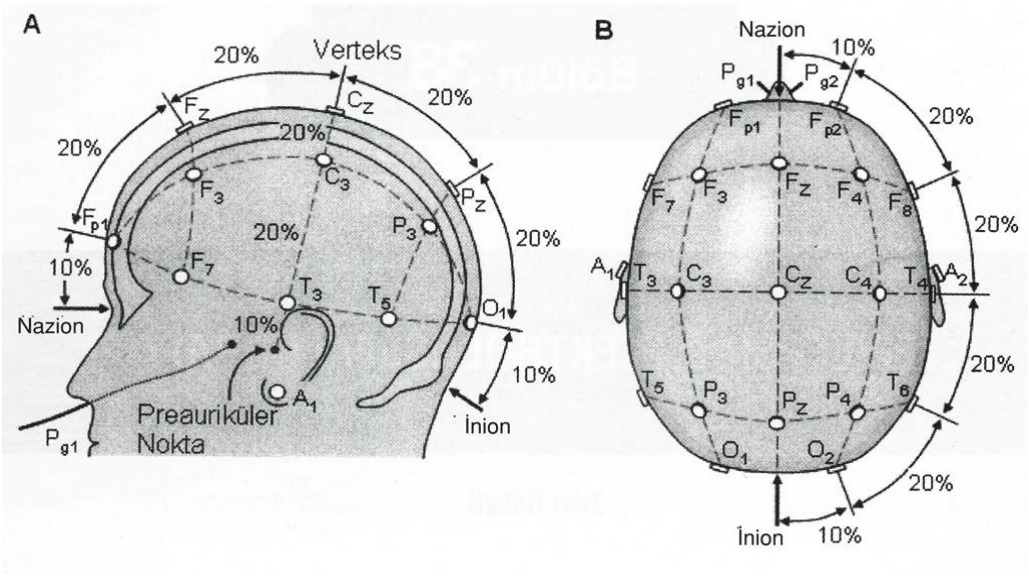
Bu bölümde yapılan uygulama çalışması araştırması ve bulgularına yer verilmiştir. Uygulamada veri madenciliği süreci adımları tek tek ele alınmış ve elde edilen bulgular belirtilmiştir. Tez uygulama çalışmasının temel amacı olan EEG sinyalleri kullanılarak, epileptiform aktiviteye Weka veri madenciliği aracı ile veri madenciliği yöntemlerininuygulanmasıdır.

6.1. Çalışmada Kullanılan Verilerin Elde Edilmesi

Beynin yapısı gereği bu organa fazla müdahale edilememektedir, dolayısıyla bu organın hastalık teşhis ve tedavisini bir hayli zorlaştırmaktadır. Bundan dolayı bu organın hastalıklarının teşhisinde birçok teknolojik cihaz geliştirilmiş ve kullanılmaktadır. Elektroansefalografidir (EEG) boyutu gereği nöroloji kliniklerinde sürekli olarak kullanılmakta ve beyin organının hastalıklarının teşhis ve tedavisinde en çok başvurulan cihaz olma özelliğini korumaktadır. (Demir,2008)

EEG kafatasına üzerine ile yerleştirilen metal elektrotlar kaydedilir. Elektrotlar iletken bir madde ile kaplı olarak kafa derisine yerleştirilir ve başlık veya kafa bandı basıncı ile yerinde tutulur. Şekil 17'de görüldüğü gibi her bir elektrotun yeri; nazion, inion, sağ ve sol preauriküler noktalardan yapılan standart ölçümlerle belirlenerek uluslararası 10-20 sistemine göre yerleştirilir. Bu sistemde yetişkindeki komşu elektrotlar, birbirlerinden 5-7 cm uzaklıktadır. Elektrotlar arasındaki potansiyel farklılıklar yükseltilir ve hareket eden bir kâğıt üzerinde birçok yazıcı tarafından kaydedilir. Günümüzde; gelişmiş dijital EEG kayıtlar sistemleri daha sonra incelenmek

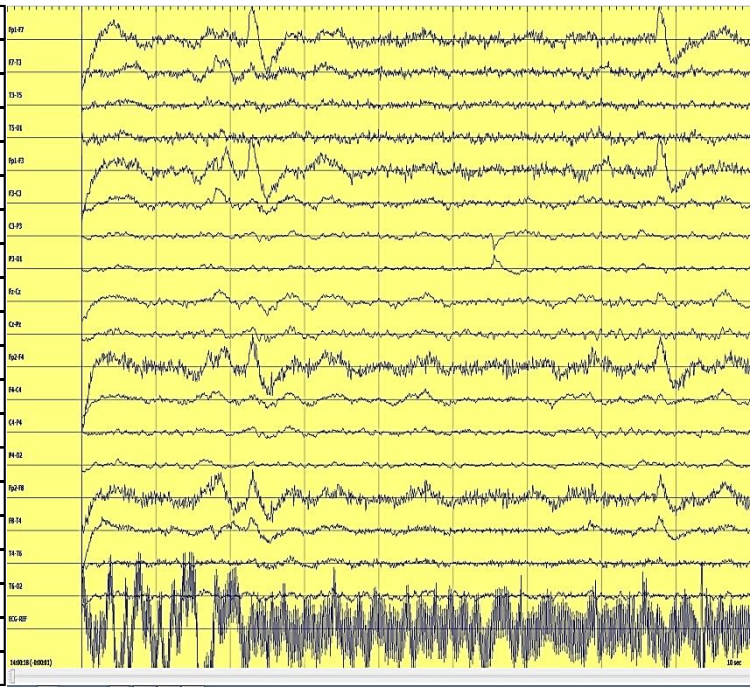
üzere sabit diske kayıt yapabilmekte hatta otomatik skorlama, artefakt temizliği yapabilmektedir.



Şekil 17. Uluslararası 10-20 Sistemi ve Elektrotların Yerleşimi

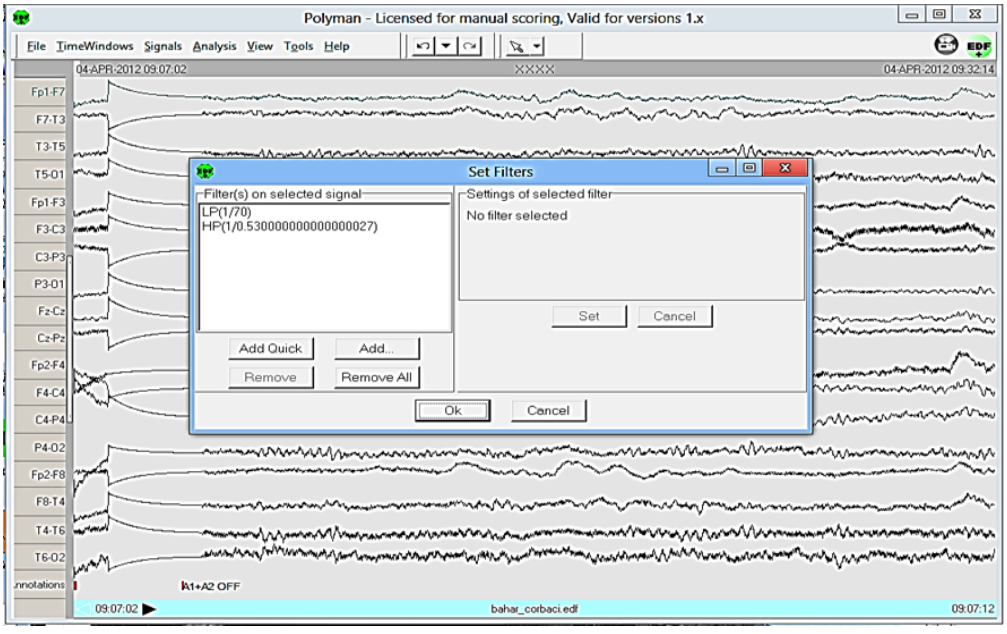
Kayıt yaklaşımları, elektrotlar arasındaki potansiyel farklılık ölçümü (bipolar derivasyon) veya tüm elektrotlar ile bilinen bir referans nokta arasındaki ölçüm (monopolar derivasyon) şeklinde değişebilir. Her bir tekniğin avantajları ve dezavantajları vardır, fakat rutin yaklaşımlar için en azından bu montajlardan ikisinin alındığı EEG trasesi elde edilmelidir. Genellikle montajlar eşdeğer elektrotların sıraları arasında longitudinal veya transvers olarak geçecek şekilde yerleştirilir. Şekil 18'de verilen örnek trasede montajlar longitudinaldir.(Bora vd.,2008)

Kanal No	Kanal Adı
1	Fp1-F7
2	F7-T3
3	T3-T5
4	T5-O1
5	Fp1-F3
6	F3-C3
7	C3-P3
8	P3-O1
9	Fz-Cz
10	Cz-Pz
11	Fp2-F4
12	F4-C4
13	C4-P4
14	P4-O2
15	Fp2-F8
16	F8-T4
17	T4-T6
18	T6-O2
19	ECG-REF



Şekil 18.Örnek EEG Trasesi Üzerinde Montajlar

Çalışmada kullanılan EEG verileri; Şekil 17'de görüldüğü gibi uluslararası 10-20 sitemine göre kafanın saçlı derine yerleştirilen elektrotlardan alınan sinyallerin yükseltilip bilgisayara kaydedilmesiyle oluşmuştur. Daha sonra bu kayıtlara Şekil 18'deki gibi sırasıyla longitudinal montajlar yapılmıştır.



Şekil 19. Polyman Programı Yardımıyla LP ve HP Filtrelerinin Uygulanması

Örnekleme frekansı 200 Hz olan verilerin artefaktlarının temizlenmesi amacıyla Polyman programı yardımıyla Şekil 19'da görüldüğü gibi 0,53 Hz'lik high pass ve 70 Hz'lik low pass filtrelere tabi tutulmuştur. Veriler veritabanına aktarılırken ayrıca nöroloji uzmanlarının göz ile kontrollerinde de geçirilmiştir.

Çalışmada kullanılan veriler seçilirken göz önünde bulundurulan kriterler;

1. Çocuklarda ve yaşlılarda EEG frekanslarının ve dalga biçimlerinin çok çeşitli olması ve bu çeşitliliğin değerlendirmeyi etkilememesi için epilepsi hastası ve kontrol grubunun yaşları 19 ila 60 arasında olanlar seçilmiştir.
2. Uykuda EEG değişimlerinin değerlendirmeyi etkilememesi için hasta uyanık ve rahat pozisyonda çekilmiş olmasına dikkat edilmiştir.
3. Nöroloji uzmanlarının teşhis koyarken kullandıkları uzunlukta yani rutin olarak 25-30 dakikalık kayıtlar alınmıştır.
4. Alınan verilerdeki epilepsi hastası ve kontrol grubu kişilerin yaş ve cinsiyetleri birbirine yakın seçilmiştir. (Bkz. Şekil 20)
5. 25'i epilepsi hastası, 25'i kontrol grubu olmak üzere eşit sayıda kayıt alınmıştır.
6. Alınan verilerin gerçekleştirilen yazılım tarafından okunabilmesi için edf (European Data Format) formatındaki veriler kullanılmıştır.

DOSYA_ADI	SONUÇ	DOSYA_ADI	YAŞ	SONUÇ
EP024	frontotemporal bölgelerde orta derecede aktif yavaş dalga aktivitesinde paroksismal b	NR009	20	Normal EEG
EP010	sol frontal bölgede orta derecede aktif paroksismal bozukluk ile uyumludur.	NR021	20	Normal EEG
EP022	Bu EEG'de oldukça aktif subkortikal deşarj izlendiği ve pirimer jenaralize epilepsi lehin	NR007	21	Normal EEG
EP016	sağ hemisfer arka kesimlerde hafif şiddette bioelektrik aktivite bozukluğunun varlığı ile	NR038	21	Normal EEG
EP002	, hafif ve yaygın organizasyon bozukluğunu düşündürmektedir.	NR006	22	Normal EEG
EP006	hafif ve yaygın organizasyon bozukluğu ile birlikte sağ frontal bölgede paroksismal ano	NR016	23	Normal EEG
EP015	hafif ve yaygın organizasyon bozukluğu ile uyumludur.	NR008	24	Normal EEG
EP019	sol temporal bölgede hafif derecede aktif paroksismal bozukluk ile uyumludur.	NR037	26	Normal EEG
EP005	hafif ve yaygın organizasyon bozukluğu ile birlikte sağda ve frontal bölgelerde daha bel	NR013	27	Normal EEG
EP008	sol temporal bölgede organizasyon bozukluğu ile birlikte epileptiform bir anomalinin v	NR018	27	Normal EEG
EP004	sol temporal bölgede daha belirgin olmak üzere sol hemisferde ağır ve yaygın organizas	NR011	28	Normal EEG
EP023	sağ frontotemporal bölgede odak izlenimi veren epileptiform bir etkinlikleuyumludur.	NR019	28	Normal EEG
EP025	sol frontal bölgede organizasyon bozukluğu ile uyumludur.	NR015	30	Normal EEG
EP012	solda daha aktif olmak üzere her iki temporal bölgede birbirinden bağımsız epileptiform	NR012	33	Normal EEG
EP011	muhtemelen orta hat yapılarından kaynaklanan paroksismal anomali ile uyumludur.	NR017	39	Normal EEG
EP001	yaygın organizasyon bozukluğu ile uyumludur.	NR031	40	Normal EEG
EP007	her iki frototemporal bölgeye lokalize hafif derecede aktif paroksismal anomali ile uyu	NR002	43	Normal EEG
EP021	sol hemisferde orta derecede aktif bioelektrik disorganizasyon ile uyumludur.	NR034	43	Normal EEG
EP003	sağ frontosentral bölgede orta derecede aktif paroksismal bozukluk ile uyumludur.	NR036	43	Normal EEG
EP013	sağ frontosantral bölgede hafif derecede aktivite gözteren paroksismal bozukluk ile uy	NR032	47	Normal EEG
EP017	hafif ve yaygın organizasyon bozukluğu ile birlikte sağ frontotemporal bölgede hafif de	NR026	49	Normal EEG
EP014	Bu bulgular hafif ve yaygın organizasyon bozukluğu ile uyumludur.	NR033	51	Normal EEG
EP018	sol temporal bölgede odak izlenimi veren hafif derecede aktif paroksismal bir anomali	NR039	54	Normal EEG
EP009	sol frontotemporal bölgede epileptojenik odak düşündürmektedir.	NR025	58	Normal EEG
EP020	sol temporal bölgede hafif derecede aktif paroksismal bozukluk ile uyumludur.	NR029	60	Normal EEG

Şekil 20. Epilepsi Hastası Olanlar ile Kontrol Gruplarının Karşılaştırılması

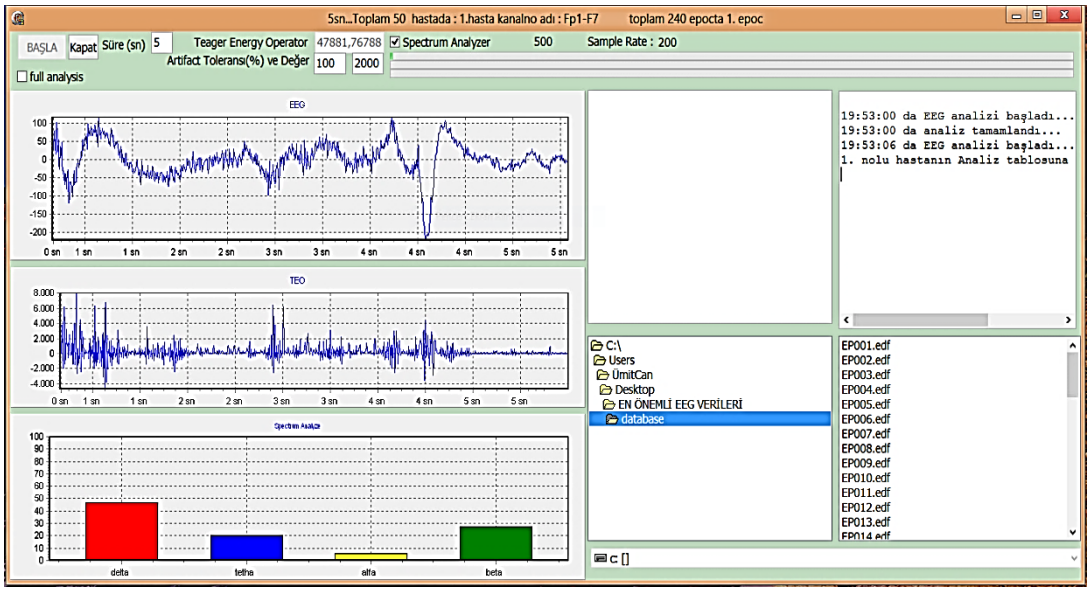
6.2. Verilerden Öznitelik Parametrelerinin Seçilmesi

EEG işareti geniş bir frekans bandına sahip olmakla birlikte, klinik ve fizyolojik ilgi 0.5 ile 30 Hz arasına yoğunlaşmıştır. Bu frekans aralığı belli frekans bantlarına ayrılmıştır. Bu yaklaşım EEG spektrumunun temelde delta (0.5-4 Hz), teta (4-8 Hz), alfa (8-13 Hz) ve beta (13-30 Hz) şeklinde dört frekans bandı içerdiği gözlemlerine dayanmaktadır. (Albayrak, 2008).

Frekans bantlarına ayırma işlemi ise ayrık dalgacık dönüşümü yardımıyla olur. Dalgacık dönüşümleri tüm sinyal frekans-zaman bilgisini korurlar. Bu sebeplerden dolayı EEG gibi durağan olmayan gerçek doğal sinyallerin dalgacık temelli metotlarla işlenmesi geleneksel metotlardan daha iyi sonuç verirler. (Batar,2005)

Teşhis işlemine yardımcı olan bir önemli faktör de EEG sinyalini alt frekans bantlarına veya özniteliklerine göre ayırarak analiz etmektir. Bu analizi bazı EEG cihazlarının üzerindeki yazılımlar yapmakla birlikte bazen bu sinyal EEG cihazından alınarak farklı bilgisayarda da yapılabilmektedir.

Şekil 21'de arayüzü görülen analiz yazılımı, edf formatındaki verileri okuyabilmekte ve bu veriler üzerinde ayrık dalgacık dönüşümü yardımıyla alt frekans bantlarına ayırma işlemini yapabilmektedir.(Umut, 2011)



Şekil 21.Verilerin Analizini Gerçekleştiren Yazılım

19 kanaldan alınmış 25-30 dakikalık edf formatındaki kayıtlarda epilepsi hastaları EP001,kontrol grubu normal kişilerin ise NR001 ile başlayan şekilde adlandırıldığı görülmektedir.

Yazılım database klasörü altındaki bu kayıtları okuyup her kaydın tüm kanallarında delta (0.5-4 Hz), teta (4-8 Hz), alfa (8-13 Hz) ve beta (13-30 Hz) alt frekans bantlarının görülme yüzdesini, toplam gücü ve artifakt yüzdesini analiz adındaki Access tablosuna yazmaktadır.

Analiz

Hasta_no	Kanal_ad	Epoc_no	Toplamguc	artifactyuzde	Delta (0,5Hz	Tetha (4Hz-8	Alfa (8Hz-12	Beta (12Hz-3
1	F3-C3	0	152139187,94	0	31,934229037	19,303648309	14,645700263	34,116422391
1	C3-P3	0	142094458,02	0	27,624564282	21,195583268	16,487482984	34,692369465
1	P3-O1	0	191761567,19	0	22,060129840	20,390987715	17,469613996	40,079268448
1	Fz-Cz	0	159348958,74	0,0025	34,485389120	20,085717468	13,569950820	31,858942592
1	Cz-Pz	0	140996742,43	0	32,281965491	21,852250868	15,735461388	30,130322253
1	Fp2-F4	0	188685619,34	0	33,68274917	14,611607554	11,351786929	40,353856348
1	F4-C4	0	140224922,34	0	30,672163924	19,718904474	14,174824222	35,43410738
1	C4-P4	0	111874045,58	0	29,003260028	21,271967242	16,035870871	33,688901859
1	P4-O2	0	143628065,85	0	23,343271712	20,577314403	18,861266611	37,218147274
1	Fp2-F8	0	171451462,88	0	28,694312474	14,397732915	11,158218125	45,749736486
1	F8-T4	0	194811911,17	0	26,889107763	14,483500896	10,731903188	47,895488154
1	T4-T6	0	164764049,08	0	16,993270276	13,954429547	12,065080344	56,987219833
1	T6-O2	0	185258398,13	0	27,628025154	21,404686798	16,560890784	34,406397264
1	ECG-REF	0	1256139870,3	0,150416666667	24,009587693	14,019906804	13,17081304	48,799692463
2	Fp1-F7	0	236559746,08	0,030196629213	40,835411304	15,038853626	8,3187673461	35,806967724
2	F7-T3	0	159143737,11	0,005617977528	42,189955863	15,116554953	8,4389190927	34,254570092
2	T3-T5	0	98219874,07	0,003511235955	25,170122016	17,108359251	11,608333258	46,113185475
2	T5-O1	0	122576749,94	0,006320224719	24,567227472	17,462391238	12,133181167	45,837200123
2	Fp1-F3	0	260190017,25	0,072331460674	38,100027775	14,170003027	8,7566528798	38,973316318
2	F3-C3	0	143181734,85	0,039325842697	36,126405172	17,70496569	11,674163361	34,494465777
2	C3-P3	0	119246165,92	0,058988764045	35,504589234	21,865265665	13,716350199	28,913794902
2	P3-O1	0	100612791,02	0,056179775281	35,135456814	21,526223943	13,450044022	29,888275221

Şekil 22. Analiz Tablosu

Literatürde bulunan bazı çalışmalarda kullanılan öznitelik parametreleri Srinivasan(2005)baskın frekansın belirlenmesi, ana enerji bölgesinde ortalama gücün hesaplanması, normalize edilmiş spectral entropi, diken ritmi, bağıl diken genliği seçilmiştir.

Yukarıda sayılan parametreler farklı çalışmalarda farklı bilim adamları tarafından da kullanılmıştır. Analiz parametreleri olarak Güler ve Übeyli'nin de kullandığı alttaki öznitelik parametrelerini kullanmıştır. Bunlar:

1. Her bir alt banttaki katsayıların maksimumu,
2. Her bir alt banttaki katsayıların minimumu,
3. Her bir alt banttaki katsayıların standart sapması,

4. Her bir alt banttaki katsayıların ortalamasıdır.

Albayrak(2008) tez çalışmasında Freiburg Üniversitesinin aldığı tek kanallı ve kısa süreli kayıtlar üzerine uyguladığı aşağıdaki istatistiksel yöntemleri öznitelik parametresi olarak almıştır. Bunlar;

1. Ortalama
2. Karesel Ortalama
3. Maksimum değer
4. Minimum Dğer
5. Standart Sapma
6. Varyans
7. Medyan
8. Sıfır geçiş sayısı

Adı geçen eserlerde kullanılan öznitelik parametreleri bizim oluşturmuş olduğumuz veri setiyle örtüşmemektedir. Çalışmada 19 kanallı 30 dakikalık kayıt incelendiği için seçilecek öznitelik parametreleri bu verileri temsil edebilmesi gereklidir. Bu yüzden çalışma için önemli olan 18 kanalın her birinin delta (0.5-4 Hz), teta (4-8 Hz), alfa (8-13 Hz) ve beta (13-30 Hz) yüzdeleri ve hastalık durumları içeren öznitelik parametresi seçilmiştir. Bu parametreler;

1. Delta1,Delta2,…, Delta18
2. Teta1,Teta2,…, Teta18
3. Alfa1,Alfa2,…Alfa18
4. Beta1, Beta2, …, Beta18
5. Hastalık {Yes,No}

Öznitelik parametrelerini açıklayacak olursak Delta1 olarak adlandırdığımız öznitelik; 1. kanalda (Fp1-F7) tüm kayıt boyunca görülen Delta altbantının yüzdesini ifade etmektedir. Diğer kanallar ve altbantları da hesaba katarsak altbantların ifade eden 72,hastalık durumunu gösteren 1 olmak üzere toplamda 73 öznitelik mevcuttur.

Ek1'de bu öznitelik parametreleri ve bu parametrelere karşılık gelen değerlerin bulunduğu öznitelik tablosu yer almaktadır.

6.3. Weka'da Model Kurma

Çalışmada kullanılan veri setlerinin yukarıda sayılan yetmiş üç ayrı öznitelik parametresi çıkarıldıktan sonra, veri madenciliği süreci ile analizlerine ve farklı algoritmalar ile yapılmış analizlerin karşılaştırılması yapılmıştır.

Tıbbi verilerin analizinde veri madenciliği algoritmalarının uygulamalarını gerçekleştirmek için WEKA (Waikato Environment for Knowledge Analysis) programı kullanılmıştır. WEKA, veri madenciliği için geliştirilmiş sınıflandırıcı makine öğrenmesi tekniklerini içeren bir analiz programıdır. Veri kümesini WEKA yazılımında kullanabilmek için ARFF(Attribute-Relation File Format) dosya formatına çevirmeliyiz. ARFF(Özellik-İlişki Dosya Biçimi) dosyası bir ASCII metin dosyasıdır. ARFF dosyaları, WEKA makine öğrenmesi yazılımı ile kullanılabilmesi için Waikato Üniversitesi Bilgisayar Mühendisliği Bölümünde Makine Öğrenmesi projesiyle geliştirilmiştir.(Uzun,2005)

Bu çalışmada kullanılan EEG veri kümesinin ARFF dosya formatındaki bir parçası şekil 24'de gösterilmektedir.

```
@RELATION ep

@ATTRIBUTE class        {Yes, No}
@ATTRIBUTE Delta1       REAL
@ATTRIBUTE Teta1        REAL
@ATTRIBUTE Alfa1        REAL
@ATTRIBUTE Beta1        REAL
@ATTRIBUTE Delta2       REAL
@ATTRIBUTE Teta2        REAL
@ATTRIBUTE Alfa2        REAL
@ATTRIBUTE Beta2        REAL

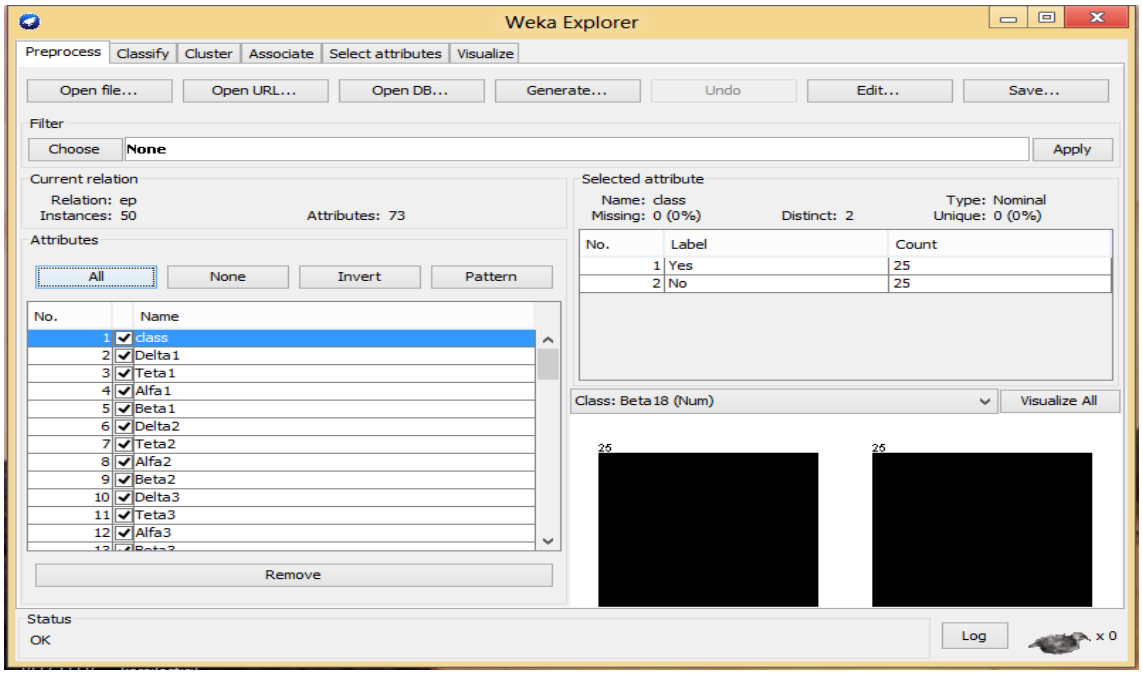
@DATA
Yes;36,56818986;16,13310177;10,91881227;36,3798961;39,2340163;15,38013889;
Yes;40,8354113;15,03885363;8,318767346;35,80696772;42,18995586;15,11655495
Yes;38,67914;17,85372126;11,90411754;31,5630212;30,73011478;18,20232149;15
Yes;48,85560823;20,1133239;8,103266952;22,92780092;48,51030034;19,77669224
Yes;46,41894819;24,34794011;8,444045023;20,78906668;36,77373912;26,2168986
Yes;47,01457168;23,1785922;8,035471311;21,7713648;47,40652743;22,02296601;
Yes;27,95374297;13,42952459;10,60557712;48,01115532;23,80430202;15,5531989
Yes;32,51858863;21,92848852;19,0109124;26,54201045;33,90180144;22,61121946
Yes;35,01228947;24,77468945;14,84833008;25,364691;24,665938;28,46313621;19
Yes;35,12930194;15,93010411;9,518629553;39,4219644;39,1378336;18,56548615;
```

Şekil 23. ARFF Dosya Formatındaki Veri Kümesi Parçası

WEKA bir proje olarak başlayıp bugün dünya üzerinde birçok insan tarafından kullanılmaya başlanan bir Veri Madenciliği uygulaması geliştirme programıdır. WEKA java platformu üzerinde geliştirilmiş açık kodlu bir programdır. WEKA çalıştırıldıktan Application menüsünde çalışılabilecek modlar listelenmektedir. Bunlar komut modunda çalışmayı sağlayan Simple CLI, projeyi adım adım görsel ortamda gerçekleştirmeyi sağlayan Explorer ve projeyi sürükle bırak yöntemiyle gerçekleştirmeyi sağlayan KnowledgeFlow seçenekleridir.

WEKA'da; önişleme (preprocess), sınıflama (classify), kümeleme (cluster), birliktelik kuralları (associate), nitelik seçme (select attribute) ve görselleştirme (visualize) panelleri bulunmaktadır. Şekil 24'deki ön işleme paneli; bilgi keşfinin başlangıç noktasıdır. Epilepsi.arff veri dosyamız bu panelden yüklendi. Sınıflandırma algoritmalarında tüm öznitelik değerlerini kullanacağımız için hepsini seçtik.

Ön işleme kısmı bittikten sonra sınıflandırma paneline geçiyoruz. Bu panel veri kümesi üzerinde WEKA'nın sınıflandırma algoritmalarının çalıştırıldığı paneldir. Sınıflandırma paneli, karar ağaçlarından kurallara; fonksiyonlardan Bayes ağlarına birçok sınıflandırma algoritmasını içermektedir.



Şekil 24. WEKA'da Ön İşleme Paneli

Weka veri modelleme algoritmalarından sınıflandırma algoritmaları iş problemi için uygun bulunmuştur. Çünkü epilepsi hastalığı tespitinde eldeki hastalık özelliklerine göre hastalık sınıfının belirlenmesi gerekmektedir. Model kurulurken verilerin belli bir kısmı eğitilmiş (training), eğitilen verilerin oluşturduğu örüntüler kullanılarak geri kalan veriler test edilmiştir. Örüntü bulma ve test veri sınıflarının tahmini işlemlerini Weka sınıflandırma algoritmaları yapmıştır. Verileri işlemede tahmin edilecek alan için doğruluk performansı yüksek olan algoritma ele alınmıştır. Seçilen algoritma, geçmiş veriler üzerinde bulduğu örüntüleri kullanarak yeni girilen hasta özelliklerine göre kişinin epilepsi olup olmadığını belirlemiştir.

Sınıflandırma algoritmalarının performanslarını test etmek için tüm algoritmalar tek tek denenmiştir. Bir algoritmayla denemeler yapılırken tüm giriş alanları tahmin edici alanlar olarak; sırayla sadece bir çıkış alanı tahmin edilen alan olarak belirlenerek Weka algoritmaları çalıştırılmıştır.

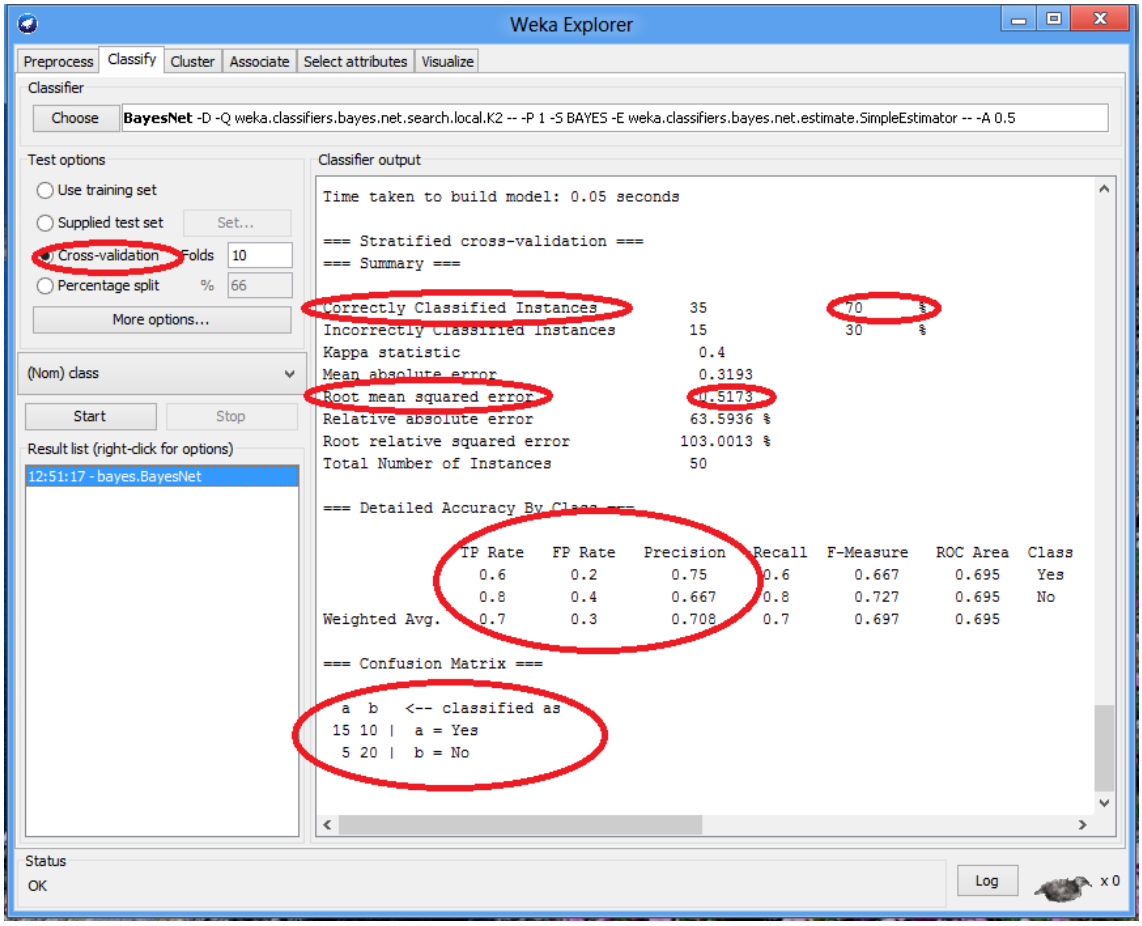
Weka sınıflandırma algoritmalarını deneme aşamasında; veriler bazı algoritmaların işlemesine uygun olmadığından bu algoritmalar çalıştırılamamıştır. Bu algoritmalar karşılaştırmada değerlendirmeye alınmamıştır.

Çalışmada kullanılan Weka 3.7 sürümünde yer alan sınıflandırıcılar sırayla seçilmiştir. Test seçeneği olarak tüm sınıflandırıcılar için 10 kat çaprazlama (10 fold cross-validation) yöntemi kullanılmıştır. Bu yöntemde veri kümesi 10 parçaya bölünür ve 10 parçadan 1 tanesi test için geriye kalan 9 tanesi eğitim için kullanılır.(Uzun,2005) Böylece 10 kat çaprazlanmış epilepsi veri kümesi üzerine Weka'da bulunan sınıflandırma algoritmaları uygulanmış ve sonuçlarına da Tablo 7'de yer verilmiştir.

6.4. Model Değerlendirme

Doğru model seçimi için uygulama çalışmasında doğruluk yüzdesi ve RMSE hata değeri kriterleri değerlendirilmiştir.

Bir model değerlendirilirken göz önünde bulundurulan başarı ölçütleri; örnek veri setinde en yüksek doğruluk oranını veren bayes.BayesNet algoritmasının sonuçlarının üzerinden anlatılmıştır.



Şekil 25. Weka Classifier (Sınıflandırıcı) Paneli

50 kaydın %90'i modelin eğitilmesi için kullanılmıştır:

50 * (90 / 100) = 45

50-45 = 5. 5 veri test amaçlı kullanılmıştır.

Tablo5'de 50 adet kaydın düzensizlik matrisi görülmektedir.

a	b	<==classified as
15	10	a=Yes
5	20	b=No

Tablo 5. Düzensizlik Matrisi

Tablo 5'e göre gerçekte 25 "Yes" değerli test verisinin 15'i "Yes" 10'u "No" ve gerçekte 25 "No" değerli test verisinin 20'si "No" 5'i "Yes" olarak tahmin edilmiştir.

Buna göre toplam 10+5=15 veri yanlış sınıflandırılmış olup diğer 15+20=35 veri doğru sınıflandırılmıştır. Buna göre bu modelin doğru sınıflandırma yüzdesi şöyle hesaplanmıştır:

Doğruluk Yüzdesi = (35 / 50) * 100 = %70'dir.

Tablo 5'te sınıflandırıcı çıkış ekranında görülen sınıflara göre 50 verinin detaylandırılmış doğruluk tablosu Tablo 6'da görüldüğü gibidir.

TP Rate	FP Rate	Precision	Class
0,6	0,2	0,75	Yes
0,8	0,4	0.667	No
0,7	0,3	0,708	

Tablo 6. Sınıflara Göre Doğruluk Çizelgesi

Tablo 5'e göre Tablo 6'daki değerler aşağıdaki şekilde hesaplanmıştır:

TP Rate(Doğru Hassaslık Oranı)	a sınıfı için	15 / 25	=	0,6
	b sınıfı için	20 / 25	=	0,8
	ortalama	35 / 50	=	0,7
FP Rate(Yanlış Hassaslık Oranı)	a sınıfı için	5 / 25	=	0,2
	b sınıfı için	10 / 25	=	0,4
	ortalama	15 / 50	=	0,3
Precision(Hassaslık)	a sınıfı için	15 / 20	=	0,75
	b sınıfı için	20 / 30	=	0,667
	ortalama	(0,75+0,667)/2	=	0,708

Algoritmaların doğru tahmin etme yüzdeleri de eşit olduğunda farklı bir kriterle karşılaştırma yapılması yoluna gidilmiştir. Bunun için en yaygın kullanılan doğruluk ölçüsünden biri olan Root Mean Square Error (RMSE) değerlerine bakılmıştır. RMSE yöntemi; interpolasyon, faz hatası ve genel değişkenlik veya anormallik durumları için hassas bir yöntemdir. (Demirel,2008)

RMSE, hata karelerinin ortalamasının karekökü değeridir. Sözkonusu hata ise, her bir veri örneğinin gerçek değeriyle tahmin edilen değeri arasındaki değerin farkının alınmasıyla elde edilir. RMSE değeri, 0 ile sonsuz arasında olup değer ne kadar küçükse hata o kadar azdır. (Demirel,2008)

Weka'nın izin verdiği tüm sınıflandırma algoritmalarının doğru sınıflandırma yüzdeleri, yanlış sınıflandırma yüzdeleri ve RMSE değerleri ayrı ayrı bulunmuş ve Tablo 7'de karşılaştırılmıştır.

Çalışmada Weka programında en yüksek doğruluk oranını veren ilk on algoritmaya bakılacak olursa;

1. bayes.BayesNet
2. functions.RBFNetwork
3. meta.Bagging
4. meta.MultiBoostAB
5. bayes.NaiveBayesMultinomialUpdateable
6. meta.LogitBoost
7. bayes.BayesianLogisticRegression
8. bayes.ComplementNaiveBayes
9. bayes.NaiveBayes
10. bayes.NaiveBayesMultinomial

olduğu görülür.

Sonuç olarak; epilepsi hastalığının teşhisi için oluşturulan veri setinde Bayes tabanlı sınıflandırma algoritmalarının daha başarılı olduğu açıkça görülmektedir. Ayrıca en başarılı on algoritmanın Weka programından elde edilen detaylı sonuçları da EK2'de yer almaktadır.

Çıkış Alan Adı	Epilepsi		
Algoritma_Adı	DY	HY	RMSE
bayes.BayesNet	70	30	0.5173
functions.RBFNetwork	68	32	0.4726
meta.Bagging	68	32	0.4884
meta.MultiBoostAB	68	32	0.5562
bayes.NaiveBayesMultinomialUpdateable	66	34	0.5835
meta.LogitBoost	66	34	0.5326
bayes.BayesianLogisticRegression	64	36	0.6
bayes.ComplementNaiveBayes	64	36	0.6
bayes.NaiveBayes	64	36	0.5934
bayes.NaiveBayesMultinomial	64	36	0.5843
bayes.NaiveBayesSimple	64	36	0.593
bayes.NaiveBayesUpdateable	64	36	0.5934
functions.SPegasos	64	36	0.6
meta.RandomCommittee	64	36	0.4883
meta.randomSubSpace	64	36	0.5177
trees.RandomForest	64	36	0.5006
functions.SMO	62	38	0.6164
meta.Dagging	62	38	0.4699
meta.Decorate	62	38	0.4705
functions.Logistic	60	40	0.6086
functions.MultilayerPerceptron	60	40	0.5838
meta.ClassificationviaClustering	60	40	0.6325
meta.MultiClassClasifier	60	40	0.6086
misc.HyperPipes	60	40	0.4939
rules.Ridor	60	40	0.6325
trees.NBTree	60	40	0.5539
trees.SimpleCart	60	40	0.5187
meta.AttributeSelectedClassifier	58	42	0.5801
meta.FilteredClassifier	58	42	0.5502
meta.RotationForest	58	42	0.5084
rules.Jrip	58	42	0.5396
trees.FT	58	42	0.5811
functions.VotedPerceptron	56	44	0.6465
lazy.IB1	56	44	0.6633
lazy.Ibk	56	44	0.6494
meta.AdaBoostM1	56	44	0.5721
rules.DecisionTable	56	44	0.5163
rules.DTNB	56	44	0.5307
trees.DecisionStump	56	44	0.532
functions.SimpleLogistics	54	46	0.5176
lazy.Kstar	54	46	0.6771
meta.ThresholdSelector	54	46	0.5994
rules.Nnge	54	46	0.6782
rules.PART	54	46	0.6731
trees.ADTree	54	46	0.5765
trees.BFTree	54	46	0.6031
trees.LMT	54	46	0.5496
trees.RandomTree	54	46	0.6782
lazy.LWL	52	48	0.552
meta.END	52	48	0.6802
rules.OneR	52	48	0.6928
trees.J48	52	48	0.6802
trees.J48graft	52	48	0.6802
trees.LADTree	52	48	0.6454
trees.REPTree	52	48	0.5459

DY=Doğruluk Yüzdesi, HY=Hata Yüzdesi, RMSE=Root Mean Sqare Error

Tablo 7. Weka 'daki Sınıflandırma Algoritmalarının Başarı Oranları

BÖLÜM 7

SONUÇLAR VE ÖNERİLER

Bazı çalışmalarda yararlanılan EEG kayıtları tek kanaldan alınmış kısa süreli kayıtlar incelendiği görülmüştür. Oysa tam yeri bilinemeyen, beynin farklı yerlerinde meydana gelebilen epilepsi hastalığının varlığının tespiti bu kayıtlarla gerçekleştirmek oldukça zordur. Gerçekte kullanılan kayıtlar 19 kanal ve en az 25 ila 30 dakikalık bir EEG çekimi sonucu uzman bir nörologun incelemesiyle karar verilebilir. Yapılan çalışmalarda %90'ın üzerinde doğruluk oranı elde edilmesinin başlıca sebebi budur. Diğer çalışmalarda ise Subaşı (2005), Güler vd. (2005), Batar (2005), Güler ve Übeyli (2005), Srinivasan vd. (2005), Alkan (2005), Subaşı (2007), Toprak (2007) çalışmalarında yapay sinir ağı temelli modeller ve farklı öznitelik parametrelerinin kullanılmıştır.

Gerçek EEG verilere en yakın 19 kanallı,25-30 dakikalık rutin kayıtlar kendi gerçekleştirmiş olduğumuz yazılımla frekans altbantlarına ayrılma işlemini başarıyla yerine getirdiği görülmüş ve veri madenciliğinde kullanılmak üzere yazılımdan elde edilen değerler kullanılmıştır. Veri madenciliği sürecinde seçilen 73 tane öznitelik kullanılarak Weka'nın veri setine uygun tüm sınıflandırma algoritmaları tek tek denenmiştir. Bu denemenin sonucu algoritmaların doğruluk yüzdelerine göre Tablo 7'de sıralanmıştır. Özellikle Bayes tabanlı algoritmalar yüksek başarı oranlarıyla dikkat çekmiştir.

Yaptığımız çalışmalarda kullandığımız yöntemlerin benzerliği nedeniyle çalışmalarının başarı oranını yüksek olmasının sebebini değerlendirdikten sonra kendi çalışmamda kullandığım çok çeşitli algoritmalar neticesinde %70'e kadar başarı sağlandığını görülmektedir.

Bu başarı oranını;

1. Kullanılan kayıtlardaki artefaktlar daha iyi temizlenir,
2. 100 ve üzerinde kayıt kullanılır,
3. Alınan veri setlerine göre tekrar öznitelik parametrelerinin tekrardan belirlenmesiyle arttırılabilir.

Gelecekte bu konuda çalışacak araştırmacılar, en yüksek doğruluk oranını verdiği tespit edilen algoritmalar yardımıyla özel amaçlı bir yazılım geliştirerek epileptik aktiviteye kolay, hızlı ve doğru bir şekilde teşhis konulmasını sağlayabilirler.

KAYNAKLAR

Akademik Bilişim Konferansları, Weka'da Örnek Uygulama,Erişim Tarihi: 01.07.2012. htpp://www.ab.org.tr/ab09/bildiri/42.pdf

Akbulut, S., Veri Madenciliği Teknikleri ile Bir Kozmetik Markanın Ayrılan Müşteri Analizi ve Müşteri Segmentasyonu, Yüksek Lisans Tezi, Gazi Ün., Fen Bil. Ens., Ankara, 2006.

Akman, M., Veri Madenciliğine Genel Bakış ve Random Forests Yönteminin İncelenmesi: Sağlık Alanında Bir Uygulama, Yüksek Lisans Tezi, Ankara Üni.,Sağlık Bil. Ens., 2010

Akpınar, H.,Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, Erişim Tarihi: 03.07.2012.http://www.isletme.istanbul.edu.tr/surekli_yayinlar/dergiler/nisan2000/1.htm

Albayrak, M., EEG Sinyallerindeki Epileptiform Aktivitenin Veri Madenciliği Süreci ile Tespiti, Doktora Tezi, Sakarya Üni.,Fen Bil. Ens., 2008

Argüden, Y., Erşahin, B., Veri Madenciliği: Veriden Bilgiye,Masraftan Değere, Arge Danışmanlık, 2008

Aslan, E., EMG İşaretlerinin İncelenmesi ve Veri Madenciliği Uygulaması, Yüksek Lisans Tezi, Sakarya Üni.,Fen Bil. Ens., 2008

Batar, H., EEG İşaretlerinin Dalgacık Analizi Yöntemleri Kullanılarak Yapay Sinir Ağları İle Sınıflandırılması, Yüksek Lisans Tezi, Kahramanmaraş Sütçü İmam Üni.,Fen Bil. Ens., 2005

Baykal., N., Tıp Bilişimi Güz Okulu, Erişim Tarihi: 11.06.2006. www.metutech.metu.edu.tr/download/File/TR/3-SUNU-Nazife_Baykal.pdf

Bayram, E., Customer Segmentation and Churn Modeling in Wireless Communication, Yüksek Lisans Tezi, Boğaziçi Ün. Fen Bil. Ens., İstanbul, 2001.

Bora İ., Yeni S.N., Gürses C. ve ark., Epilepsi, Nobel Tıp Kitapları, 2008, sf.475-499

Bilginin Adresi, Yapay Sinir Ağları , Erişim Tarihi:20.06.2012, http://bilgininadresi.net/Madde/58071/Yapay%20Sinir%20A%C4%9Flar%C4%B1%20(ANN)

Bransten, L., Technology - Power Tools - Looking For Patterns: Data Mining Enables Companies To Better Manage The Ream of Statistics They Collect; The Goal: Spot The Unexpected, Wall Street Journal, 27 (12): 16- 20, 1999.

Davis, B., Data Mining Transformed, Information Week, 751: 86, 1999.

Demir, Ö., EEG Dalgalarının Wavelet(Dalgacık) Dönüşümü ile Değerlendirilmesi, Yüksek Lisans Tezi, Dumlupınar Üni.,Fen Bil. Ens., 2008

Demirel, B., Meme Kanseri Tedavi Yöntemlerinin Veri Madenciliği ile Belirlenmesi, Yüksek Lisans Tezi, Süleyman Demirel Üni.,Fen Bil. Ens., 2008

Doğan, Ş., Veri Madenciliği Kullanılarak Biyokimya Verilerinden Hastalık Teşhisi, Yüksek Lisans Tezi, Fırat Üni.,Fen Bil. Ens., 2007

Dumouchel, W., Bayesian Data Mining in Large Frequency Tables, With An Application To The FDA Spontaneous, American Statistician, 53 (3): 177, 1999.

Epilepsi Rehberi, Epilepsi Hakkında, Erişim Tarihi:20.06.2012, http://www.epilepsirehberi.com/tr_linkpage_1.htm

Epilepsy Center Australia Epilepsi Nöbetinde İlk Yardım, Erişim Tarihi: 21.06.2012,www.epinet.org.au/downloads/.../79073%20TURKISH%20web.pdf

Esen, M.F., Veritabanlarında Bilgi Keşfi: Veri Madenciliği ve Bir Sağlık Uygulaması, Yüksek Lisans Tezi, İstanbul Üni., Sosyal Bil. Ens., 2009

ETD (Epilepsi ve Toplum Derneği)a, Epilepsi Nedir?,Erişim Tarihi:20.06.2012, http://www.epilepsi.org.tr/epilepsi/epilepsi.asp

ETD (Epilepsi ve Toplum Derneği)b, Epilepsinin Tarihi, Erişim Tarihi:21.06.2012, http://www.epilepsi.org.tr/epilepsi/epilepsi_tarih.asp

ETD (Epilepsi ve Toplum Derneği)c, Elektroensefalografi (EEG), Erişim Tarihi:01.07.2012, http://www.epilepsi.org.tr/epilepsi/eeg.asp

Farboudi , S., Tıp Bilişiminde İstatistiksel Veri Madenciliği, Yüksek Lisans Tezi, Hacettepe Üni.,Fen Bil. Ens., 2009

Fayyad, U.M.; Piatesky-Shapiro, G.; Smyth, P.; Uthurusamy, R., 1994, "Advances in data mining and Knowledge Discovery", AAAI Pres, USA.

Fidan, S., Dalga Kılavuzunda Yayılan Elektromanyetik Dalganın Dalgacık Dönüşümü ile Modellenmesi, Yüksek Lisans Tezi, Gazi Üni.,Fen Bil. Ens., 2006

Fisch BJ. Spehlmann'ın EEG el kitabı, Çeviri editörü: Şahiner T. İstanbul: Turgut Yayıncılık, 1998

Güler, D., Übeyli, E.D., Güler, E., A Mixture of Experts Network Structure for EEG Signals Classification, Proc. of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp.2707–2710, 2005.

Güler, D., Übeyli, E.D., Adaptive Neuro Fuzzy Inference Sysytem for Classification of EEG Signals using Wavelet Coefficients, Journal of Neuroscience Methods, Vol. 148, 113–121, 2005.

Hand, D.J., Data Mining: Statistics and More?, The American Statistician, 52:112118, 1998.

Incaklinik (İstanbul Nörocerrahi ve Ağrı Merkezi) Epilepsi , Erişim Tarihi:20.06.2012, http://www.incaklinik.com/epilepsi.html

Irmak, S., Veri Madenciliği Yöntemleri ile Sağlık Sektörü Veritabanlarında Bilgi Keşfi: Tanımlayıcı ve Kestirimci Model Uygulamaları, Doktora Tezi, Akdeniz Üni.,Sosyal Bil. Ens., 2009

Kayaalp, K., Asenkron Motorlarda Veri Madenciliği ile Hata Tespiti, Yüksek Lisans Tezi, Süleyman Demirel Üni.,Fen Bil. Ens., 2007

Kitler, R., Wang, W., The Emerging Role of Data Mining, Solid State Technology, 42 (11): 45, 1998.

Nationaal Epilepsie Fonds, Epilepsi (Sara) Hastalığı, Erişim Tarihi: 21.06.2012. http://www.epilepsie.nl/attachments/epilepsieinhetkortturks.pdf

Nisbet, R.A., 2004, How to Choose a Data Mining Suite, DM Direct Special Report.

Papoulis, A., Signal Analysis,Mcgraw-Hill International Editions,New York, 1984

Polikar, R., The Engineer's Ultimate Guide to Wavelet Analysis The Wavelet Tutorial,1999, Http://Engineering.Rowan.Edu/~Polikar/Wavelets/Wtpart3.Html

RESTIVO, K., The Drill on Data Mining, Computer Dealer News, 15 (14): 29-30, 1999.

Safiye BİLGİN Prof. Dr., Elektroensefalografi (EEG), Erişim Tarihi: 01.07.2012. http://safiyebilgin.com/eeg.htm

Srinivasan, V., Eswaran, C., Sriraam, N., Artificial Neural Network Based Epileptic Detection using Time-Domain and Frequency-Domain Features, Journal of Medical Systems, Vol. 29, No. 6, 2005.

Subaşı, A., Automatic Recognition of Alertness Level from EEG by using Neural Network and Wavelet Coefficients, Expert System with Applications, Vol.28, pp.701–711, 2005.

Subaşı, A., EEG Signal Classification using Wavelet Feature Extraction and a Mixture of Expert Model. Expert Systems with Applications, 32, 1084–1093, 2007.

Sweldens, W. 1994. Wavelet Sampling Technics. Http://W3.Nai.Net/~Alokaeng/Wavelet/Sampling.Pdf

Şentürk, Z .K., Veri Madenciliği ile Kanser Tanısı, Yüksek Lisans Tezi, İstanbul Üni., Fen Bil. Ens., 2011

Umut., İ., 2011, Dijital Sinyal İşleme Yazılımı Geliştirilmesi ve Geliştirilen Bu Yazılım ile Elektroensefalografi Kayıtları Kullanılarak Tıkayıcı Uyku Apnesi Bulunan Bireylerin Apnesiz Bireylerden Ayırt Edilmesi, Doktora Tezi, Trakya Üniversitesi Fen Bilimler Enstitüsü, 2011.

Uzun, Y., Tıbbi Veriler Üzerinde Makine Öğrenme Algoritmaları ve Bulanık Mantık ile Kurallar Öğrenme, Yüksek Lisans Tezi, Selçuk Üni., Fen Bil. Ens., 2005

Toprak, B., Çağlar, M.F., Merdan, M., Dalgacık Dönüsümü ve Bulanık Mantık Sistemine Dayalı Uyarlanır Ağ ile Epilepsi Hastalığı Teshisi, SDÜ 15. Yıl Mühendislik Mimarlık Sempozyumu, sf. 9-17, Isparta, 14-16 Kasım 2007.

Torrence, C., Compo, P.,C., A "Practical Guide To Wavelet Analysis. Bulletin Of American Meteorological Society", 79(L): 61 – 78,1998

Wikipedia a, Epilepsi, Erişim Tarihi:21.06.2012, http://tr.wikipedia.org/wiki/Epilepsi

Wikipediab, Seizure trigger, Erişim Tarihi: 21.06.2012, http://en.wikipedia.org/wiki/Seizure_trigger

Yurtsever, U., 2002, Veri Madenciliği ve Uygulaması, Yüksek Lisans Tezi, Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, Sakarya

.

EKLER

EK-1 Öznitelik Tablosu.

HNo	Eğ	Delta1	Teta1	Alfa1	Beta1	Delta2	Teta2	Alfa2	Beta2	Delta3	Teta3	Alfa3	Beta3	Delta4	Teta4	Alfa4	Beta4	Delta5	Teta5	Alfa5	Beta5	Delta6	Teta6	Alfa6	Beta6	Delta7	Teta7	Alfa7	Beta7	Delta8	Teta8	Alfa8	Beta8	Delta9	Teta9	Alfa9	Beta9	Delta10	Teta10	Alfa10	Beta10	Delta11	Teta11	Alfa11	Beta11
1	Yes	36,57	16,1	10,9	36,4	39,23	15,4	9,27	36,11	16,77	13,7	11,4	38,2	20,76	17,53	15,6	46,1	33,83	14,5	11,7	39,9	31,93	19,3	14,6	34,12	27,62	21,2	16,5	34,7	22,06	20,4	17,5	40,1	34,49	20,1	13,6	31,9	32,282	21,85	15,74	30,13	33,683	14,61	11,35	40,35
2	Yes	40,04	15	8,32	35,6	42,19	15,1	8,44	34,25	25,17	17,1	11,6	46,11	24,57	17,46	12,1	45,8	38,1	14,2	8,76	39	36,13	17,7	11,7	34,49	35,5	21,9	13,7	28,9	33,14	21,5	13,5	39,3	35,82	22,3	15,19	42,3	27,956	21,49	14,74	35,82	36,683	14,24	10,59	42,3
3	Yes	38,68	17,9	11,9	31,6	30,73	18,2	15,5	35,54	21,25	18,1	17,3	43,35	22,77	16,1	24,1	37,1	36,73	18,6	11,6	35,1	29,01	19,12	20,5	31,36	25,9	19,1	25,9	29,2	28,54	21,8	23,1	28,6	36,03	20,3	15,1	26,5	33,219	21,59	20,67	24,32	30,499	17,21	12,67	39,63
4	Yes	48,86	20,1	8,1	22,9	48,51	19,8	9,67	22,04	40,16	20,5	11,2	26,13	36,46	18,6	11,7	33,2	44,05	21,1	8,91	25,9	32,17	20,11	14,4	33,32	31,18	21	15,3	32,5	28,41	18,9	13,9	38,7	31,73	20,6	14,8	32,9	27,083	20,22	16,71	35,99	43,695	21,67	11,21	23,43
5	Yes	46,42	24,3	8,44	20,8	36,77	26,2	9,63	27,38	34,21	22,6	11,7	31,4	32,16	28,7	13,1	26	38,65	24,7	8,79	27,9	28,99	26,42	11,4	33,14	30,64	27,7	11,8	29,9	33	28,6	12,7	25,7	37,61	29,6	9,41	23,4	34,788	29,11	11,58	24,52	32,467	21,37	6,538	39,58
6	Yes	47,01	23,2	8,04	21,8	47,41	22	8,68	21,89	46,33	21,9	9,29	22,29	44,7	21,07	10,4	23,9	47,71	22,2	7,65	22,5	41,79	25,22	8,57	24,42	43,39	24,1	9,71	22,8	43,43	24	9,59	23	41,48	28,2	6,89	21,4	39,291	26,72	9,462	24,53	44,489	22,23	8,054	25,23
7	Yes	27,95	13,4	10,6	48	23,8	15,6	14,5	46,15	22,38	14,7	18,5	44,46	25,29	15,32	25,7	33,7	26,76	14	9,99	49,3	20,67	14,03	12,9	52,4	30,8	14,3	17,4	37,5	25,82	15	26,1	33,1	27,53	20	16	36,4	30,627	15,76	19,14	34,27	30,407	15,85	10,89	42,85
8	Yes	32,52	21,9	19	26,5	33,9	22,6	18,1	25,35	28,38	22,4	24,5	24,72	28,91	21,17	25,3	24,6	29,14	20,8	17,3	32,8	23,7	19,62	25,2	31,47	24,39	20,2	26,6	28,8	24,44	21,8	29,1	24,6	19,22	19,5	26,2	35,1	24,476	18,24	25,56	34,72	27,237	18,11	16,05	38,6
9	Yes	35,01	24,8	14,8	25,4	24,67	28,5	19,3	27,58	22,99	20,2	27,7	29,05	21,18	18,25	25,6	35	35,49	21,8	13,2	29,5	21,96	16,68	21,1	40,23	20,26	16,1	27,5	36,1	20,58	17,4	30,3	31,7	23,1	17	21,4	38,5	22,65	16,62	25,4	35,33	34,697	18,01	14,33	32,96
10	Yes	35,13	15,9	9,52	39,4	39,14	16,6	11,9	30,38	18,3	18,2	20,8	42,69	18,65	19,11	21,3	41	39,14	17,4	10	33,5	21,03	20,38	19,5	39,09	20,85	20,8	21,6	37	19,16	22,9	22,9	35	21,43	23,8	18,2	36,5	18,461	22,71	22,47	36,35	32,461	16,5	11,49	39,55
11	Yes	33,06	29,3	13,8	23,8	24,88	26,1	17,1	31,56	19,61	27,8	21,2	31,41	21,23	28,62	17,9	32,3	32,34	28,8	12,4	26,5	23,07	26,94	20,5	29,48	20,89	27,3	23,5	28,3	20,2	29,9	22,1	27,8	27,72	31	16,6	24,7	23,988	28,63	21,35	26,02	33,289	29,81	12,26	24,64
12	Yes	43,89	21,3	9,79	25,1	29,26	20,8	15,3	34,66	23,06	24,9	21,6	30,49	17,65	21,48	31,3	29,7	39,36	20,7	9,36	30,6	21,64	21,94	15,8	40,65	17,32	24,3	23,9	34,5	15,91	22,5	31,3	30,3	24,86	20,2	12,7	42,7	16,125	22,4	22,8	38,68	44,514	20,35	8,006	27,13
13	Yes	29,3	17,6	13,1	40,1	22,78	14	14,7	49,55	18,27	12,8	14,6	54,38	21,06	17,1	19,5	42,4	27,21	17	13,1	42,7	19,58	16,25	18,2	45,93	22,58	17,6	18,1	41,7	22,78	18,4	19,4	39,4	23,01	20	15,7	41,2	24,837	17,5	16,2	41,66	32,293	20,32	13,36	34,08
14	Yes	41,42	22,5	11,2	24,9	34,86	26,2	12	26,56	32,21	28,5	11,8	27,5	30,57	30,43	12,9	26,1	40,14	22,6	12	25,2	34,57	26,07	13,6	25,78	32,09	27,8	14,1	26	32,99	27,9	13,3	25,9	34,09	27,5	13	25,4	28,849	30,6	14,63	25,92	42,522	22,51	11,39	23,59
15	Yes	32,37	12,1	7,55	47,9	19,72	10	10,7	59,49	20,63	11,1	10,5	57,75	28,52	15,33	10,2	45,9	29,34	11	7,58	52,1	31,12	14,23	11,1	43,54	28,45	18,1	14,4	39,1	32,25	19,8	13,4	33,8	41,59	16,6	11,8	30	32,182	19,8	14,66	33,36	31,063	10,48	7,493	51,02
16	Yes	27,58	19,9	12,1	40,4	24,94	21,3	11,3	42,49	24,92	20,9	11,3	43,5	40,18	23,2	10,8	25,9	25,45	16,5	12,7	45,4	22,96	22,49	15,1	39,52	35,33	20,1	13,3	31,3	51,21	16	10,4	22,3	21,26	25,1	16	37,6	25,189	23,53	16,21	35,08	26,727	17,19	16,4	39,68
17	Yes	41,61	27,4	10,4	20,6	34,17	27,4	13,2	25,22	36,77	30,4	11,3	21,51	41,38	27,88	12,7	18,1	43,6	25,1	13,3	18	36,12	25,72	16,5	17,68	30,78	25,5	23,5	20,2	35,7	29,9	16,7	17,7	38,69	30,4	12,2	18,8	40,797	29,87	12,25	17,08	38,399	26,71	17,58	17,21
18	Yes	25,85	25,9	16,5	31,8	15,59	20,9	23,8	39,73	13,38	15,9	30,8	39,92	14,78	16,45	33,3	35,5	23,44	20,5	17,9	38,1	15,22	17,96	23,1	43,73	14,32	15,8	33,5	36,4	12,88	13,5	41,9	31,7	22,74	21,7	20,2	35,4	21,684	16,54	27,64	33,93	29,173	18,94	17,93	33,96
19	Yes	30,2	16,6	10,3	42,9	26,6	17,2	12,1	44,1	18,57	16,5	15	49,91	17,45	17,8	19,6	45,2	27,29	17,2	9,55	45,6	17,82	16,8	14,1	51,29	15,3	15,5	16	53,2	15,12	16,3	19,6	49	19,34	19,1	12,6	48,5	17,664	16,88	13,25	52	26,582	19,48	10,75	48,19
20	Yes	16,37	16,3	14,5	52,8	14,61	16,7	21,5	47,22	13,94	17,5	28	40,57	15,13	15,73	27	42,1	16,18	14,5	14,9	54,4	12,28	14,05	19,6	54,08	12,86	14	25,4	47,7	13,08	16,2	31,9	38,8	11,94	13,9	18,4	55,8	10,004	13,28	23,97	51,95	16,5	14,57	16,31	52,08
21	Yes	41,83	19,5	11,6	27,1	36,92	23,1	15,3	24,72	33,52	20,6	20,2	25,68	33,82	20,63	24,8	20,7	40,49	18,1	10,5	31	38,84	17,28	14,1	29,76	34,85	18,9	19,1	27,2	32,73	20,2	24,8	22,3	38,07	17,3	15,8	28,8	32,495	18,48	20,09	28,19	35,398	18,38	12,15	34,07
22	Yes	47,64	17,5	7,97	26,9	41,41	19,1	8,75	30,74	25,28	19,1	10,8	44,78	24,58	15,4	11,2	40,8	41,62	15,6	7,55	35,2	26,52	18,4	8,88	46,2	25,37	19,7	10,2	44,7	31,63	23,1	12,6	32,6	39,77	26,8	11,3	22,1	36,175	27,28	12,55	24	51,722	20,74	8,996	18,55
23	Yes	31,09	23,2	12	33,7	34,38	25,9	13,3	26,36	28,04	28,4	18,4	25,1	31,06	23,81	20,4	24,7	42,1	20,8	10,5	26,6	47,04	22,31	11,1	19,59	40,64	21,8	15,2	22,4	27,17	26,6	19,8	26,5	29,33	27,3	16,9	26,5	26,918	26,01	20,63	26,43	38,172	22,37	11,03	28,42
24	Yes	35,42	18,3	10,2	36,1	22,57	18,7	15,6	43,12	19,58	15,9	16,3	49,18	19,51	18,18	23,6	38,7	31,29	18,5	9,28	40,9	20,95	21,41	20,7	36,59	20,53	21,8	19,3	38,3	17,22	20,3	26,6	35,9	25,83	23,7	15,9	34,5	20,712	24,55	18,38	35,36	33,585	20,91	11,16	34,34
25	Yes	33,35	26,7	11,4	28,5	33,2	22,4	11,6	32,85	28,84	22,9	10,3	37,96	19,3	28,06	19,7	33	33,77	27,4	9,79	29,1	29,24	25,45	11,5	33,83	21,25	32,2	18,6	28	23,45	35,6	17,3	23,6	27,05	28,4	13,6	31	23,168	31,58	18,33	26,92	25,608	20,12	10,13	44,15
26	No	22,85	15	15,8	46,3	17,71	13,1	15,9	53,33	16,39	13,3	22,1	40,21	15,34	18,48	36,3	30	23,53	18,7	15,8	42	16,93	20,49	27,1	35,53	19,74	19,1	28,9	32,2	15,55	18,8	38,2	27,5	17,87	22,3	26,6	33,3	22,02	19,59	30,77	27,63	26,115	19,48	16,62	37,78
27	No	23,61	12	13,1	51,3	22,57	13,4	17,4	46,61	16,57	15,5	31,7	36,26	15,22	16,79	34	34	26,46	14,3	13,3	45,9	19,38	15,68	25,7	39,21	17,89	16,1	31,7	34,3	16,59	16,9	37,7	28,8	24,99	19	22,8	33,1	18,278	16,68	32,01	33,03	30,024	16,58	15,36	38,03
28	No	33,02	12,4	20,8	33,8	32,82	22,5	13,6	31,06	15,54	21,9	21,8	40,8	25,72	16,84	27,8	29,6	32,21	24,2	15	28,6	15,12	16,56	30,6	37,74	23,19	16,7	28,2	31,9	22,09	16,6	32,5	28,8	23,06	29,6	20	27,3	28,165	20,82	22,23	28,78	30,612	19,56	13,42	36,01
29	No	41,81	16,6	10,1	31,4	18,51	11,1	13,7	56,71	17,51	11,4	18,8	52,29	20,35	12,7	28,2	38,8	40,23	16,7	9,05	34	23,31	15,9	20,4	40,35	23,9	15,5	28,9	31,7	17,83	12,7	31	33,5	27,8	18,7	22,6	30,9	15,482	12,62	40,21	31,75	41,603	18,54	10,13	29,73
30	No	39,35	23,9	12	24,8	28,23	26,5	19,1	26,18	21,48	20,2	31,2	27,07	19,68	20,66	28,4	31,3	34,28	26,8	13,5	25,4	28,24	22,01	18,1	31,62	18,64	20,9	25,7	34,7	19,1	20,4	30	30,6	34,58	24,2	14,3	27	19,083	22,28	26,47	32,22	38,956	23,48	11,9	25,66
31	No	50,21	16,6	9,4	23,8	39,75	18,2	11,4	30,67	22,74	17,5	18,9	40,34	17,35	14,57	26,8	41,3	48,91	17,5	8,88	24,7	29,07	20,38	14,3	36,24	25,13	17,1	17,2	40,5	16,59	14,9	26,2	42,3	44,84	17,9	10,4	26,9	43,686	19,12	11,29	25,89	56,932	15,86	7,792	19,42
32	No	39,74	19,1	13,3	27,9	24,61	17,2	19,1	39,11	15,09	15,8	30,8	38,28	17,02	15,04	28,6	39,3	39,65	18,8	10,6	30,9	23,15	17,19	22,7	37,02	16,68	16,8	26,9	39,6	15,24	15,2	33,1	36,5	28,81	21,1	19,7	30,4	41,288	17,36	16,67	24,68	42,758	22,66	12,27	22,31
33	No	34,25	17,1	10,6	38	31,56	16,7	12,4	39,36	36,05	19,6	12,3	32,1	22,6	16,61	21	39,8	32,43	17,1	11,4	39	23,82	18,9	20,1	37,17	19,68	17,5	20,5	42,4	17,72	18,1	25,4	38,8	24,6	22,3	18,3	34,8	21,733	21,53	22,46	34,3	27,349	15,74	10,1	46,81
34	No	33,84	21	12,5	32,6	33,86	23,1	12,8	32,28	24,04	17,7	14,3	49,9	21	19,88	16,9	42,3	30,61	19,4	13,1	36,9	19,7	16,52	18,1	45,72	21,51	17,4	16,3	44,7	21,17	18,9	18,1	41,9	21,49	21,2	15,9	41,3	22,402	19,31	16,2	42,1	30,356	21,11	12,88	35,66
35	No	40,77	17,8	10,4	31	31,22	16,8	12,3	39,68	23,57	16,4	15,3	44,73	20,73	15,25	18	46	39,38	17,9	10,3	32,4	26,3	19	14,1	40,65	23,71	17,9	15,1	43,3	20,27	15,6	17,6	46,5	26,94	20,3	14	38,7	21,733	18,8	18,28	16,22	42,77	38,865	18,3	10,72	32,11
36	No	42,83	15,8	15,3	26	36,14	12,5	17,7	33,65	18,29	12,5	36,7	32,56	31,57	13,17	29,2	26	37,52	17,4	13,5	31,6	22,18	14,14	31,9	31,81	16,41	12,9	41,8	28,8	26,52	12,5	34	25	21,99	15,5	28,5	34	14,43	12,92	44,17	28,48	44,489	17,74	11,04	26,73
37	No	22,37	18	14,6	45,1	20,74	17,7	16,3	45,25	16,65	19,2	21,4	42,73	16,91	19,1	24,9	39,1	20,72	16,2	13,6	49,5	17,57	18,28	18,3	45,86	22,29	18,7	18,8	40,2	21,18	17,2	23,3	38,4	18,28	19,6	19	43,1	20,346	18,96	20,46	40,23	24,58	18,53	14,7	42,69
38	No	37,47	14,6	13,5	34,4	18,93	15	23,8	37,33	17,25	15,4	30,2	37,07	15,48	13,99	34,2	36,3	35,34	14,6	12,2	37,9	15,41	12,6	33,8	38,15	15,46	12,7	34,3	37,5	14,38	13,1	37,5	35	23,31	16	26,7	34	15,575	13,16	34,97	36,49	34,641	16,09	14,13	35,14
39	No	36,06	18,2	10,3	35,4	32,07	18,6	13,1	36,23	25,19	17	15,7	42,15	22,16	16,87	20,9	40,1	32,38	16,8	10,4	40,5	26,55	20,62	14,8	38	24,23	19,3	19,2	37,3	20,97	17,7	20	41,3	28,88	23,8	14,3	33	25,065	20,88	19,71	34,35	36,265	19,94	11,02	32,78
40	No	32,31	17,3	11,6	38,8	31,68	15,6	11,4	41,29	17,35	13,3	13,5	55,85	21,01	14,72	13,6	50,7	29,49	16,4	10,8	43,3	23,05	18,48	13,2	45,27	21,3	16,9	14,4	47,5	19,21	15,3	13,4	52,1	35,32	18,9	11,7	34,1	31,482	17,13	12,61	38,82	29,479	17,34	11,37	41,82
41	No	40,72	22,2	8,53	28,5	37,54	18,9	7,88	35,72	23,49	19,9	8,61	46,01	26,81	19,65	9,48	44,1	42,2	18,8	7,83	31,2	37,06	27,36	9,86	25,72	35,57	31,1	11,1	22,2	37,03	28,8	11,2	23,1	34,65	32,5	11,8	21	33,121	31,68	13,13	22,06	37,502	22,67	9,085	30,74
42	No	44,91	20,6	11,6	22,9	46,29	19,7	13,1	20,89	29,77	21,9	17,2	31,13	28,15	21,97	17,3	32,6	40,64	20,6	11,3	27,5	24,49	23,57	19	32,95	20,95	23,9	20,6	34,6	16,78	27,5	23,1	32,6	26,48	22,6	15,6	35,4	19,288	24,31	21,19	35,21	36,229	22,24	12,84	28,67
43	No	32,64	16,8	14,7	35,9	22,95	18	20,3	38,79	16,22	18,5	30,3	34,96	11,47	16,9	30,7	33	27,4	16,3	15,1	41,2	14,6	18,18	28,1	39,13	11,7	17,5	35,8	35	13,14	18,2	37,8	30,9	19,15	21,2	23,1	36,6	14,668	19,15	34,47	31,68	28,293	18,45	15,82	37,44
44	No	40,48	23,5	15	21	25,79	20,9	25	28,27	21,44	22,4	28,7	27,45	15,57	22,33	25,2	26,9	39,49	22,3	15,2	22,9	20,8	18,05	28,6	32,52	17,68	20,3	33,1	28,9	16,49	21,3	36,3	25,9	22,13	18,1	22,4	37,4	20,362	18,51	28,96	32,17	35,161	21,48	19,49	23,87
45	No	34,21	16	12,8	36,8	29,27	17	14	39,74	23,19	18,1	21,6	37,07	24,1	18,1	25,3	32,4	33,15	17,9	12,8	36,1	24,75	17,25	16,3	41,66	23,29	17,3	20,7	38,7	24,6	17,9	26,1	31,3	25,93	17,1	13,8	43,1	22,72	17,83	18,76	40,7	34,048	18,21	12,7	35,05
46	No	19,62	10,3	9,19	60,9	23,73	13,1	10,6	52,55	17,74	12,6	14,5	55,12	16	12,69	17,6	53,7	16,19	8,54	8,65	66,6	14,41	10,98	10,1	64,52	14,84	11,9	12,5	60,8	14,19	11,5	17,9	56,4	16,51	12,5	10,7	60,2	15,637	11,68	11,98	60,71	16,732	10,64	9,04	64,59
47	No	25,86	14,6	11,7	47,8	27,83	16,2	13,5	42,41	25,38	15,4	18,4	40,81	27,14	16,36	18,7	37,8	25,29	15,5	11,4	47,9	21,79	15,11	13,1	50,04	25,99	15,5	17,3	41,2	22,67	15,6	23,5	38,3	22,54	16,8	12,1	48,4	24,013	16,72	15,28	43,99	25,166	15,74	12,06	47,04
48	No	30,73	28	10,3	11	35,7	16,8	12,1	33,35	26,53	17,7	11,4	44,39	18,81	14,49	12,9	53,8	49,25	27	10,5	13,2	37,17	21,13	13,1	28,58	24,72	17,1	11,3	46,8	25,56	14,1	11,7	46,6	42,54	23,7	12,5	21,3	27,734	17,79	13,51	40,99	49,926	26,58	10,92	12,57
49	No	42,91	16,7	10,7	29,7	26,99	15,9	15,7	41,36	21,18	16,3	21,9	40,65	14,8	14,48	32,6	38,1	42	20,1	11,7	26,2	30,78	15,73	11,3	22,23	43,25	20,6	13,2	22,9	24,52	15,1	27,6	32,8	27,77	21,3	16,9	34	23,889	17,41	25,44	33,26	37,328	18,06	11,38	33,23
50	No	34,46	22,8	15,4	27,3	21,65	20,6	22	35,76	18,92	18	27,2	35,94	15,94	17,21	25,1	41,8	34,09	21,1	12,8	32	18,26	17,61	16,9	47,25	15,21	16,3	20,9	47,5	14,26	15,5	28,8	41,5	21,51	19,3	16,1	48	17,289	17,94	19,67	44,9	33,12	21,46	13,29	32,13

Delta12	Teta12	Alfa12	Beta12	Delta13	Teta13	Alfa13	Beta13	Delta14	Teta14	Alfa14	Beta14	Delta15	Teta15	Alfa15	Beta15	Delta16	Teta16	Alfa16	Beta16	Delta17	Teta17	Alfa17	Beta17	Delta18	Teta18	Alfa18	Beta18
30,672	19,72	14,17	35,43	29,003	21,27	16,04	33,69	23,343	20,58	18,86	37,218	28,694	14,4	11,16	45,75	26,889	14,48	10,73	47,9	16,993	13,95	12,07	56,99	27,628	21,4	16,56	34,406
39,81	19	10,61	30,58	36,78	19,77	14,18	29,27	47,984	20,72	10,55	20,744	41,958	13,39	6,809	37,84	45,153	17,82	8,768	28,26	29,363	14,6	9,871	46,16	42,3	22,97	8,979	25,756
35,356	22,31	14,86	27,48	37,525	20,63	18,84	23,01	25,801	27,35	20,61	26,242	44,643	19,13	9,856	26,37	48,351	20,46	9,262	21,93	26,555	21,88	15,27	36,29	24,677	23,44	23,85	28,028
18,683	20,47	20,36	40,48	22,879	21,94	21,66	33,51	23,833	20,42	17,94	37,805	47,701	22,75	9,351	20,2	33,54	20,11	14,06	32,29	29,294	21,01	15,38	34,31	26,87	19,66	16,09	37,379
38,073	27,88	9,115	24,93	36,956	27,27	11,73	24,05	36,939	28,01	11	24,052	40,511	19,76	8,224	31,51	37,717	26,01	10,59	25,69	35,476	25,11	11,21	28,2	35,673	27,14	11,48	25,708
42,239	26,51	8,646	22,61	44,793	24,02	8,523	22,66	45,76	20,48	8,772	24,986	44,977	22,9	7,697	24,43	47,374	19,82	7,814	24,99	46,246	18,89	7,262	27,6	52,394	17,41	7,337	22,856
22,038	15,42	15,27	47,27	19,222	15,63	22,05	43,1	24,063	14,75	27,26	33,934	31,773	15,99	11,8	40,43	24,227	15,73	15,37	44,68	20,05	14,93	22,84	42,18	25,927	15,45	23,74	34,881
18,195	17,11	26,82	37,87	19,7	18,59	27,03	34,68	20,202	18,54	32,56	28,7	28,285	20,29	19,53	31,89	29,681	18,23	18,22	33,87	20,774	18,23	27,51	33,48	21,219	19,21	29,71	29,859
20,246	15,7	21,44	42,61	21,788	16,91	25,2	36,11	19,17	15,99	30,3	34,54	29,896	19,73	18,11	32,27	25,913	16,82	21,86	35,41	18,398	17,01	31,13	33,46	18,886	15,78	24,89	40,447
21,922	22,4	18,05	37,63	17,82	22,48	22,5	37,2	16,332	20,7	25,54	37,434	37,293	15,86	9,908	36,94	27,128	17,87	14,64	40,36	14,373	16,45	19,29	49,89	16,015	18,01	23,68	42,299
24,823	29,76	18,19	27,23	23,264	28,33	21,28	27,13	20,273	29,47	23,09	27,169	34,244	27,76	13,31	24,68	27,419	24,76	18,06	29,77	24,639	27,67	19,82	27,87	23,529	26,83	23,54	26,1
23,074	20,47	15,1	41,36	17,501	21,62	22,39	38,49	14,566	21,68	32,08	31,669	49,443	21,64	8,611	20,31	23,703	16,69	12,68	46,93	17,645	19,08	22,63	40,64	18,137	23,17	26,12	32,571
20,23	16,87	17,4	45,49	21,856	15,72	16,66	45,77	20,286	14,73	20,57	44,407	29,389	18,34	15,34	36,93	21,842	14,01	15,32	48,83	19	15,69	17,56	47,75	22,725	16,61	21,44	39,225
31,225	26,89	14,23	27,66	27,246	27,88	14,51	30,36	29,017	33,77	12,97	24,235	44,299	20,48	9,701	25,52	42,052	19,55	9,633	28,76	26,578	30,33	12,48	30,61	29,956	34,3	13,09	22,662
25,408	12,43	10,07	52,09	37,321	12,98	8,528	41,17	53,712	14,79	6,917	24,583	34,261	12,82	7,317	45,6	23,496	8,42	8,347	59,74	23,36	9,98	8,573	58,09	45,065	17	8,11	29,824
21,502	21,85	17,23	39,42	26,085	22,83	18,03	33,05	34,433	24,15	15,72	25,691	25,272	23,37	15,15	36,21	34,561	23,76	13	28,68	35,116	25,87	13,81	25,21	38,314	25,21	12,84	23,637
40,707	31,68	12,33	15,28	42,204	26,9	15,48	15,42	31,499	27,86	21,15	19,491	49,554	29,07	7,835	13,54	46,31	30,6	8,859	14,23	40,591	31,63	13,46	14,33	39,24	28,15	16,38	16,234
15,89	16,48	27,84	39,79	18,972	14,87	30,54	35,62	14,795	12,38	42,09	30,734	27,547	17,15	22,21	33,09	17,941	13,27	27,79	41	10,832	10,74	40,16	38,27	12,981	12,79	35,71	38,514
19,405	18,97	12,45	49,17	16,032	17,2	14,31	52,45	16,757	16,4	18,22	48,616	28,034	18,76	11,78	41,43	27,389	16,47	11,76	44,38	16,481	15,7	17,93	49,89	18,94	18,53	17,25	45,281
10,422	12,9	21,29	55,39	10,838	13,91	30,03	45,22	9,7199	16,57	35,31	38,404	17,411	16,25	15,91	50,43	9,9358	12,16	19,89	58,01	10,703	14,52	28,32	46,46	12,426	16,52	30,79	40,261
29,476	17,54	19,64	33,34	25,551	16,3	27,5	30,64	21,396	14,98	36,01	27,616	36,802	16,72	13,57	32,92	26,898	15,53	20,72	36,85	21,036	14,33	25,92	38,71	23,09	15,39	34,22	27,301
34,93	26,38	12,15	26,54	30,934	25,1	12,57	31,39	28,739	28,41	15,8	27,051	50,736	16,95	7,493	24,82	35,36	21,92	10,63	32,09	27,477	25,25	14,3	32,97	24,011	20,08	12,95	42,958
33,842	28,46	14,86	22,83	27,369	25,41	21,4	25,82	30,886	23,47	21,79	23,856	41,187	28,03	11,66	19,12	33,781	27,16	16,37	22,68	32,163	28	17,44	22,4	39,623	26,16	16,36	17,862
22,645	23,19	18	36,16	20,848	22,88	20,31	35,96	19,315	20,23	27,31	33,141	38,022	20,49	11,19	30,3	26,724	21,43	17,29	34,56	22,851	19,65	23,54	33,97	20,988	20,91	23,02	35,08
20,038	23,92	15,18	40,87	21,866	31,08	19,66	27,39	21,432	35,28	20,47	22,814	31,71	17,4	8,504	42,38	31,753	19,79	11,35	37,11	19,499	32,86	16,78	30,86	19,938	34,2	19,76	26,103
20,595	21,19	26,31	31,91	18,128	18,98	33,15	29,74	16,056	18,09	40,36	25,493	25,375	19,05	16,69	38,88	18,951	16,75	17,77	46,52	16,852	15,11	25,13	42,92	18,253	19,21	36,34	26,201
23,303	17,69	21,98	37,02	16,936	14,44	32,75	35,88	17,865	16,1	37,78	28,256	26,685	14,26	15,16	43,9	21,597	12,77	20,73	44,89	17,798	15,78	33,65	32,77	16,411	15,99	32,09	35,508
38,863	21,35	13,58	26,21	18,912	19,23	23,71	38,15	31,831	19,42	21,24	27,508	30,327	17,76	17,91	34,01	32,132	21,48	12,6	33,8	14,968	15,86	31,15	38,02	19,401	14,62	23,93	42,053
25,239	15,87	26,81	32,08	14,01	10,11	41,88	33,99	14,945	10,19	32,22	42,644	42,809	17,68	11,34	28,17	20,204	11,89	16,3	51,6	12,951	9,463	30,3	47,28	13,808	9,206	36,55	40,439
30,102	22,83	19,74	27,33	18,921	21,14	30,83	29,11	18,298	19,71	35,42	26,572	39,912	26,37	13,09	20,63	30,083	26,52	18,19	25,21	24,149	22,67	30,7	22,49	19,781	23,9	27,75	28,573
39,394	19,48	12,49	28,63	27,903	13,88	20,64	37,57	31,909	11,67	28,12	28,295	57,979	14,33	7,866	19,82	31,714	17,2	13,82	37,27	20,051	15,31	24,44	40,2	20,076	13,42	31,17	35,33
25,747	19,18	20,59	34,48	17,548	18,88	24,87	38,7	18,042	17,24	26,11	38,601	43,08	19,81	12,84	24,27	23,934	18,05	17,19	40,82	16,843	15,9	26,65	40,61	17,067	16,69	26,19	40,055
23,685	17,25	16,72	42,34	23,126	19,43	19,66	37,78	20,353	18,47	23,28	37,888	30,374	17,25	11,09	41,29	21,27	13,98	13,15	51,6	13,258	10,57	12,27	63,9	18,239	15,19	17,87	48,7
21,165	19,28	16,27	43,28	22,895	18,65	16,84	41,62	22,19	18,18	18,3	41,337	29,397	18,75	13,47	38,39	24,112	16,6	14,18	45,11	22,124	16,31	14,45	47,11	20,559	18,5	16,87	44,069
22,726	18,19	14,98	44,11	20,936	17,4	15,31	46,35	26,657	14,41	16,17	42,758	42,894	17,39	9,576	30,14	30,451	16,48	12,28	40,78	22,4	15,68	14,34	47,58	25,702	14,8	16,82	42,679
21,127	14,34	30,8	33,73	21,702	13,47	37,28	27,54	29,673	13,32	32,43	24,582	44,004	16,73	15,25	24,01	26,098	14,36	18,8	40,75	19,266	12,66	30,17	37,91	27,514	13,39	32,29	26,804
17,649	18,98	19,42	43,96	19,194	18,38	21,54	40,89	17,766	17,52	26,15	38,565	23,259	19,7	15,91	41,13	23,287	19,52	17,4	39,79	18,139	19,54	23,74	38,59	15,686	20,09	25,13	39,098
18,997	14,43	30,64	35,94	14,037	12,67	33,49	39,8	17,423	13,99	33,04	35,546	36,92	15,87	14,96	32,25	23,783	15,74	25,83	34,65	17,729	14,17	32,27	35,83	20,216	14,95	29,5	35,334
29,487	22,86	13,68	33,98	24,851	21,13	18,09	35,93	23,038	19,41	18,84	38,711	39,975	20	10,62	29,4	33,417	17,85	11,41	37,32	24,312	15,76	15,37	44,56	27,606	20,53	17,21	34,648
22,485	17,74	12,83	46,94	22,887	15,83	13,27	48,01	15,334	12,92	13,3	58,44	27,862	16,65	11,83	43,67	22,61	13,66	10,9	52,83	15,93	12,05	12,86	59,16	14,38	12,49	14,08	59,052
29,148	33,84	13,79	23,22	32,111	30,1	15,77	22,02	34,634	27,77	13,24	24,351	33,022	24,71	8,586	33,68	26,073	22,67	10,6	40,66	20,494	23,51	12,64	43,36	23,075	27,45	11,46	38,017
24,873	23,36	18,27	33,5	19,962	25,51	22,51	32,02	20,213	26,31	20,61	32,871	41,271	20,39	13,1	25,24	34,52	20,85	16,6	28,04	21,78	22	18,69	37,53	18,963	20,56	17,37	43,104
16,408	19,47	26,6	37,53	12,05	17,45	38,3	32,19	12,394	17,19	39,28	31,137	28,944	18,24	16,77	36,05	21,447	19,35	23,75	35,45	16,186	18,34	33,85	31,62	13,802	18,21	36,32	31,676
15,646	15,47	30,5	38,38	17,757	19,94	32,11	30,19	16,974	20,64	36,38	26,006	39,079	22,83	15,23	22,87	23,613	18,76	24,96	32,66	18,467	20,61	29,84	31,08	15,522	20,67	34,46	29,345
28,587	17,3	14,27	39,84	23,545	17,14	19,51	39,8	23,274	16,54	22,14	38,048	32,75	17,58	13	36,67	33,239	16,26	13,11	37,38	25,361	17,52	17,26	39,86	25,126	18,21	19,89	36,775
18,312	12,6	10,56	58,53	14,594	10,8	11,55	63,06	17,969	12,65	14,59	54,792	20,556	11,23	9,805	58,41	21,41	12,89	11,17	54,53	16,559	11,94	14,22	57,28	16,976	13,02	13,54	56,462
21,987	16,07	12,39	49,56	26,908	16,57	17,05	39,47	20,676	14,84	26,03	38,452	23,363	14,41	11,91	50,32	22,917	15,13	13,07	48,88	25,801	14,94	19,49	39,77	26,763	15,06	18,26	39,912
44,803	25,19	11,32	18,68	26,574	19,51	14,46	39,46	19,389	14,68	16,5	49,43	49,547	27,3	11,29	11,86	40,609	21,5	9,214	28,68	23,113	17,25	12,42	47,22	20,747	14,96	19,36	44,935
35,832	16,74	12,84	34,58	37,329	19,94	14,25	28,49	32,18	15,33	24,55	27,939	40,748	17,44	11,88	29,93	23,769	13,86	15,92	46,45	17,2	12,79	22,14	47,87	14,554	15,37	30,26	39,82
20,207	18,76	16,19	44,85	13,631	15,76	19,96	50,65	13,425	15,06	29,78	41,737	32,158	22,88	17	27,97	22,722	21,19	19,08	37,01	18,396	17,34	26,41	37,85	16,667	17,72	25,19	40,42

EK-2. Weka Elde Edilen Sonuçları

1.bayes.BayesNet Algoritmasının Sonuçları

LogScore Bayes: -270.4831865780908
LogScore BDeu: -287.6621909905319
LogScore MDL: -298.08058069290297
LogScore ENTROPY: -257.0043391359074
LogScore AIC: -278.00433913590746

Time taken to build model: 0.05 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	35	70 %
Incorrectly Classified Instances	15	30 %
Kappa statistic	0.4	
Mean absolute error	0.3193	
Root mean squared error	0.5173	
Relative absolute error	63.5936 %	
Root relative squared error	103.0013 %	
Total Number of Instances	50	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.6	0.2	0.75	0.6	0.667	0.695	Yes
	0.8	0.4	0.667	0.8	0.727	0.695	No
Weighted Avg.	0.7	0.3	0.708	0.7	0.697	0.695	

=== Confusion Matrix ===

a	b	<-- classified as
15	10	a = Yes
20		b = No

2. functions.RBFNetwork Algoritmasının Sonuçları

pCluster_0_0 -0.9489
pCluster_0_1 0.9489
Intercept 0.3021
Odds Ratios...

	Class
Variable	Yes
==========	==========
pCluster_0_0	0.3872
pCluster_0_1	2.5829

Time taken to build model: 0.05 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	34	68 %
Incorrectly Classified Instances	16	32 %
Kappa statistic	0.36	
Mean absolute error	0.441	
Root mean squared error	0.4726	
Relative absolute error	87.8347 %	
Root relative squared error	94.0906 %	
Total Number of Instances	50	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.48	0.12	0.8	0.48	0.6	0.592	Yes
	0.88	0.52	0.629	0.88	0.733	0.592	No
Weighted Avg.	0.68	0.32	0.714	0.68	0.667	0.592	

=== Confusion Matrix ===

a	b	<-- classified as
12	13	a = Yes
3	22	b = No

3.meta.Bagging Algoritmasının Sonuçları

Time taken to build model: 0.19 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	34	68 %
Incorrectly Classified Instances	16	32 %
Kappa statistic	0.36	
Mean absolute error	0.4526	
Root mean squared error	0.4884	
Relative absolute error	90.1367 %	
Root relative squared error	97.2479 %	
Total Number of Instances	50	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.64	0.28	0.696	0.64	0.667	0.636	Yes
	0.72	0.36	0.667	0.72	0.692	0.636	No
Weighted Avg.	0.68	0.32	0.681	0.68	0.679	0.636	

=== Confusion Matrix ===

a	b	<-- classified as
16	9	a = Yes
7	18	b = No

4.meta.MultiBoostAB Algoritmasının Sonuçları

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	34	68	%
Incorrectly Classified Instances	16	32	%
Kappa statistic	0.36		
Mean absolute error	0.3587		
Root mean squared error	0.5562		
Relative absolute error	71.4347 %		
Root relative squared error	110.7536 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.6	0.24	0.714	0.6	0.652	0.625	Yes
	0.76	0.4	0.655	0.76	0.704	0.625	No
Weighted Avg.	0.68	0.32	0.685	0.68	0.678	0.625	

=== Confusion Matrix ===

a	b	<-- classified as
15	10	a = Yes
6	19	b = No

5. bayes.NaiveBayesMultinomialUpdateable Algoritmasının Sonuçları

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	33	66	%
Incorrectly Classified Instances	17	34	%
Kappa statistic	0.32		
Mean absolute error	0.3507		
Root mean squared error	0.5835		
Relative absolute error	69.8492 %		
Root relative squared error	116.1892 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.56	0.24	0.7	0.56	0.622	0.686	Yes
	0.76	0.44	0.633	0.76	0.691	0.694	No
Weighted Avg.	0.66	0.34	0.667	0.66	0.657	0.69	

=== Confusion Matrix ===

a	b	<-- classified as
14	11	a = Yes
6	19	b = No

6.meta.LogitBoost Algoritmasının Sonuçları

Time taken to build model: 0.05 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	33	66 %
Incorrectly Classified Instances	17	34 %
Kappa statistic	0.32	
Mean absolute error	0.4123	
Root mean squared error	0.5326	
Relative absolute error	82.1095 %	
Root relative squared error	106.0408 %	
Total Number of Instances	50	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.64	0.32	0.667	0.64	0.653	0.634	Yes
	0.68	0.36	0.654	0.68	0.667	0.634	No
Weighted Avg.	0.66	0.34	0.66	0.66	0.66	0.634	

=== Confusion Matrix ===

a	b	<-- classified as
16	9	a = Yes
8	17	b = No

7. bayes.BayesianLogisticRegression Algoritmasının Sonuçları

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	32	64	%
Incorrectly Classified Instances	18	36	%
Kappa statistic	0.28		
Mean absolute error	0.36		
Root mean squared error	0.6		
Relative absolute error	71.6949 %		
Root relative squared error	119.4658 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.56	0.28	0.667	0.56	0.609	0.64	Yes
	0.72	0.44	0.621	0.72	0.667	0.64	No
Weighted Avg.	0.64	0.36	0.644	0.64	0.638	0.64	

=== Confusion Matrix ===

a	b	<-- classified as
14	11	a = Yes
7	18	b = No

8.bayes.ComplementNaiveBayes Algoritmasının Sonuçları

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	32	64	%
Incorrectly Classified Instances	18	36	%
Kappa statistic	0.28		
Mean absolute error	0.36		
Root mean squared error	0.6		
Relative absolute error	71.6949 %		
Root relative squared error	119.4658 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.56	0.28	0.667	0.56	0.609	0.64	Yes
	0.72	0.44	0.621	0.72	0.667	0.64	No
Weighted Avg.	0.64	0.36	0.644	0.64	0.638	0.64	

=== Confusion Matrix ===

```
 a  b   <-- classified as
14 11 |  a = Yes
 7    18 |  b = No
```

9.bayes.NaiveBayes Algoritmasının Sonuçları

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	32	64	%
Incorrectly Classified Instances	18	36	%
Kappa statistic	0.28		
Mean absolute error	0.3587		
Root mean squared error	0.5934		
Relative absolute error	71.4422 %		
Root relative squared error	118.149 %		
Total Number of Instances	50		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.56	0.28	0.667	0.56	0.609	0.688	Yes
	0.72	0.44	0.621	0.72	0.667	0.685	No
Weighted Avg.	0.64	0.36	0.644	0.64	0.638	0.686	

=== Confusion Matrix ===

```
 a  b   <-- classified as
14 11 |  a = Yes
 7 18 |  b = No
```

10.bayes.NaiveBayesMultinomial Algoritmasının Sonuçları

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	32	64	%
Incorrectly Classified Instances	18	36	%
Kappa statistic	0.28		
Mean absolute error	0.3508		
Root mean squared error	0.5843		
Relative absolute error	69.8612 %		

Root relative squared error 116.3337 %
Total Number of Instances 50

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.56	0.28	0.667	0.56	0.609	0.687	Yes
	0.72	0.44	0.621	0.72	0.667	0.694	No
Weighted Avg.	0.64	0.36	0.644	0.64	0.638	0.691	

=== Confusion Matrix ===

a	b	<-- classified as
14	11	a = Yes
7	18	b = No

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	Ümit Can KUMDERELİ
Uyruğu	**:**	T.C.
Doğum Yeri ve Tarihi	**:**	04/02/1987
Telefon	**:**	
e-mail	**:**	kumdereli.umitcan@gmail.com

EĞİTİM

Derece		**Adı**	**Yılı**
Lise	:	**Tekirdağ Anadolu Lisesi**	**2001-2005**
Lisans	:	**Sakarya Üniversitesi,Bilgisayar Mühendisliği**	**2006-2010**
Yüksek Lisans	:	**Trakya Üniversitesi,Bilgisayar Mühendisliği**	**2010-2012**

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2007-2008	**Sakarya İl Milli Eğitim Müdürlüğü, Bilgi İşlem**	**Memur**
2008-2009	**Adapazarı İlçe Milli Eğitim Müdürlüğü, Bilgi İşlem**	**Memur**
2009-2010	**Sakarya Üniversitesi, Bilgi İşlem Dairesi Başkanlığı**	**VHKİ**
2010-2012	**Hacettepe Üniversitesi, Polatlı Teknik Bilimler MYO**	**Öğretim Görevlisi**
2012-	**Trakya Üniversitesi, Bilgisayar Mühendisliği Bölümü**	**Araştırma Görevlisi**

UZMANLIK ALANI

Veri Madenciliği, Makine Öğrenmesi Bilgisayar Yazılımı,

YABANCI DİLLER

İngilizce İleri Düzey, Almanca Başlangıç Düzeyi