

T.C.
SÜLEYMAN DEMİREL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MİKROEKRANLI HATLARIN VERİ MADENCİLİĞİ
YÖNTEMLERİ İLE MODELLENMESİ

Sedat ATEŞ

Danışman: Yrd.Doç.Dr.Yavuz CENGİZ

YÜKSEK LİSANS TEZİ
ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ ANABİLİM DALI
ISPARTA - 2012

TEZ ONAYI

Sedat ATEŞ tarafından hazırlanan “**Mikroekranlı Hatların Veri Madenciliği Yöntemleri İle Modellenmesi**” adlı tez çalışması aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Anabilim Dalı’nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman :Yrd.Doç. Dr. Yavuz CENGİZ

Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Mühendisliği A.B.D.

Jüri Üyeleri :

Yrd. Doç. Dr. Övünç POLAT

Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Mühendisliği A.B.D.

Doç.Dr. Adnan KAYA

Süleyman Demirel Üniversitesi Elektronik ve Haberleşme Mühendisliği A.B.D.

Prof. Dr. Mehmet Cengiz KAYACAN

Enstitü Müdürü

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

İÇİNDEKİLER

	Sayfa
TEŞEKKÜR	viii
ŞEKİLLER DİZİNİ	ix
ÇİZELGELER DİZİNİ	x
SİMGELER DİZİNİ	1
1. GİRİŞ	3
2.KAYNAK ÖZETLERİ	5
3.MATERYAL VE YÖNTEM	9
3.1.Materyal.....	9
3.1.1.Havasız Ortam Dikdörtgen, Trapez, V,Eliptik Şekilli Mikroekranlı Hatlar	9
3.1.1.1. Dikdörtgen Ve Trapez Şekilli Mikroekranlı Hatlar	9
3.1.1.1.1, Havasız Ortam Dikdörtgen ve Trapez Şekilli Mikroekranlı Hatların yapıları	9
3.1.1.1.2 Havasız Ortam Dikdörtgen ve Trapez Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı.....	11
3.1.1.1.3. Veri Madenciliğinin Dikdörtgen ve Trapez Şekilli Mikroekranlı hatta uygulanması	13
3.1.1.2. V-Şekilli İletken Sonlu Mikroekranlı Hatlar	15
3.1.1.2.1 V-Şekilli İletken Sonlu Mikroekranlı Hatların Yapıları.....	15
3.1.1.2.2 Havasız Ortam V-Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı	16
3.1.1.2.3. Veri Madenciliğinin V- Şekilli Mikroekranlı hatta uygulanması	18
3.1.1.3. Eliptik Şekilli İletken Sonlu Mikroekranlı Hatlar	19
3.1.1.3.1 Eliptik Şekilli İletken Sonlu Mikroekranlı Hatların Yapıları	19
3.1.1.3.2 Havasız Ortam Eliptik Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı	20
3.2. Yöntem	22
3.2.1. Veri Madenciliği	22
3.2.1.1 Tanımı.....	22
3.2.1.2. Gereksinimleri.....	23
3.2.1.3. Amaçları	24

3.2.1.4. Karşılaşılan Problemler	24
3.2.1.5. Kullanım Alanları.....	26
3.2.1.6. Etkilendiği Unsurlar	28
3.2.1.7. Modelleri.....	29
3.2.1.7.1. Sınıflama ve Regresyon Modelleri.....	29
3.2.1.7.2. Kümeleme Modelleri.....	31
3.2.1.7.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler	31
3.2.1.8. Yöntemleri	31
3.2.1.8.2. Tahminleme	32
3.2.1.8.2.1. Tahmin Kavramı.....	32
3.2.1.8.2.2. Tahmin Yöntemleri.....	33
3.2.1.8.2.3. Tahmin Modellerinin Doğruluklarının Ölçülmesi	34
3.2.1.8.3.1. Basit Doğrusal Regresyon.....	36
3.2.1.8.3.2. Çoklu Doğrusal Regresyon	37
3.2.1.8.4. Zaman Serileri Analizi.....	38
3.2.1.8.4.1. Zaman Serilerinin Temel Özellikleri.....	39
3.2.1.8.4.2. Stokastik (Tasadüfi) Süreç Olması.....	40
3.2.1.8.4.3. Zaman Serilerinin Sınıflandırılması	40
3.2.1.8.4.4. Analiz Amaçları ve Uygulama Alanları	41
3.2.1.8.4.5. Zaman Serilerinde İleriye Dönük Tahmin Yöntemleri.....	41
3.2.1.8.5. Yapay Sinir Ağları.....	44
3.2.1.8.5.1. Yapay sinir ağları matematiksel modeli	45
3.2.1.8.5.2. Aktivasyon fonksiyonları.....	46
3.2.1.8.5.3. İleri beslemeli yapay sinir ağları	47
3.2.1.8.5.4. Geri beslemeli yapay sinir ağları.....	48
4.ARAŞTIRMA BULGULARI VE TARTIŞMA.....	50
4.1 Trapez ve Dikdörtgen Şekilli Mikroekranlı Hatta Veri madenciliği uygulaması.....	50
4.2. V-şekilli Mikroekranlı Hatta Veri Madenciliği Uygulaması.....	57
4.3. Eliptik Şekilli Mikroekranlı Hatta Veri Madenciliği Uygulaması.....	64
5. SONUÇ	70
KAYNAKLAR.....	72

ÖZGEÇMİŞ.....	74
---------------	----

ÖZET

Yüksek Lisans Tezi

MİKROEKRANLI HATLARIN VERİ MADENCİLİĞİ YÖNTEMLERİ İLE MODELLENMESİ

Sedat ATEŞ

Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü
Elektronik Ve Haberleşme Anabilim Dalı

Danışman: Yrd. Doç Dr. Yavuz CENGİZ

Veri madenciliği yöntemi ile havasız ortamda trapez, dikdörtgen, V ve Eliptik şekilli mikro ekranlı hatların karakteristik empedansları hesaplanmıştır. Literatürde konu üzerinde analitik, Bilgisayar destekli tasarım (CAD), yapay sinir ağları (YSA), adaptif bulanık yapay sinir ağları (ANFIS) vb yöntemlerle yapılmış çalışmalara rastlamak mümkündür: Bu çalışmada ilk kez mikroekranlı hatların empedansları veri madenciliği yöntemi ile hesaplanmıştır. Bu hesaplama için günümüzde veri madenciliğinde kullanılan programlardan (YALE, R, WEKA, SPSS, Clementine vb) bir tanesi olan Weka; seri çalışması birçok algoritmayı bünyesinde barındırması, sınıflandırma (classify), kümeleme (cluster), birleştirme (associate), görsellik (visualize), data transfer, vb özelliklerinden dolayı tercih edilmiştir. analitik olarak hesaplanmış bulunan Hat karakteristik empedans değerleri veri madenciliği yöntemine uygulanmıştır. IBK, Kstar, Additive Regresyon, Bagging, Decision Table, M5 Rules, Trees M5P REPTree Multilayer perceptron vb veri madenciliği yöntemlerinin performansları test edilmiştir. Yapılan çalışmalar neticesinde trapez şekilli mikroekranlı hatta M5P yöntemi, v şekilli mikroekranlı hatta Bagging yöntemi, eliptik şekilli mikroekranlı hatta Multilayer Perceptron yöntemi en iyi performansı sergilemiştir. Bu esnada çalışma iki ana grupta toplanmıştır. İnterpolasyon ve ekstrapolasyon. Bu tez çalışması ile veri madenciliği yöntemlerinin mikroekranlı hat modellemesinde kullanılabileceği gösterilmiştir.

Anahtar Kelimeler Mikroekranlı hatlar, karakteristik empedans, veri madenciliği, weka

2012, 84 sayfa

ABSTRACT

M.Sc. Thesis

DATA MINING METHOD FOR MODELLING MICROSHIELD LINE

Sedat ATEŞ

**Süleyman Demirel University
Graduate School of Applied and Natural Sciences
Department of Electronic and Communication Engineering**

Supervisor: Yrd. Doç Dr. Yavuz CENGİZ

The characteristic impedance of trapezoidal, rectangular, V, elliptic shaped microshield line calculated by data mining method. In the literature this calculation was made by numerically, CAD (computer aided design), and ANFİS (artificial neural network, adaptive neuro fuzzy interference system). In this study the characteristic impedance of microshield line calculated by data mining method first-time. For this calculation, one of the data mining method (for instance YALE, R, WEKA, SPSS, Clementine vb) weka is chosen because of working speedy, having various algorithms, classify, cluster, associate, visualize, easy data transfer attributes. Microshield line characteristic impedance, that calculated analytically, is applied to data mining method. The performance of IBK, Kstar, Additive Regression, Bagging, Decision Table, M5 Rules, Trees M5P REPTree Multilayer perceptron etc. data mining algorithms are tested. Consequence of these tests the best results are M5P algorithm for trapezoidal; bagging for v-shaped; Multilayer perceptron for elliptical shaped microshield line. There are two main study groups Interpolation and extrapolation. By this thesis it is shown that data mining methods can be used for modelling microshield lines.

Keywords:Microshield Lines, Characteristic Impedances, Data Mining, Weka

2012, 84 pages

TEŞEKKÜR

Bu araştırma için beni yönlendiren, karşılaştığım zorlukları bilgi ve tecrübesi ile aşmamda çok büyük emeği geçen bu uğurda gece gündüz elinden gelen her türlü fedakârlığı esirgemeyen değerli Danışman Hocam Yrd. Doç. Dr. Yavuz CENGİZ 'e teşekkürlerimi sunarım. Veri madenciliği konusunda yardımcı olan değerli hocam Uğur Ecir KÜÇÜKSİLLE'ye ve diğer emeği geçen hocalarımıza teşekkür ederim.

Tezimin her aşamasında bana karşı hiçbir fedakarlıktan kaçınmayan aileme sonsuz sevgi ve saygılarımı sunarım.

Sedat ATEŞ
ISPARTA, 2012

ŞEKİLLER DİZİNİ

Şekil 3.1. Havasız ortam dikdörtgen şekilli mikroekranlı hatlarının geometrisi.....	10
Şekil 3.2. Havasız ortam trapez şekilli (+) açılı mikroekranlı hatlarının geometrisi .	10
Şekil 3.3. Havasız ortam trapez şekilli (-) açılı mikroekranlı hatlarının geometrisi ..	10
Şekil 3.4 Trapez şekilli mikroekranlı hattın blok şeması.....	11
Şekil 3.5. CSVBVPW ın yatay kesiti (Türkmen,vb2006).	15
Şekil 3.6. CBCPW ın yatay kesiti(Türkmen,vb 2006).....	15
Şekil 3.7. V- şekilli mikroekranlı hattın blok şeması	16
Şekil 3.8. ESML Nil yatay kesiti(Türkmen, 2008).....	19
Şekil 3.9 CSML nin yatay kesiti(Türkmen, 2008).....	19
Şekil 3.10. Eliptik şekilli mikroekranlı hattın blok şeması	20
Şekil 3. 11. yapay sinir ağı matematiksel açılımı	45
Şekil 3.12. Basamak fonksiyonu ve matematiksel ifadesi	46
Şekil 3.13. Eşik fonksiyonu ve matematiksel ifadesi.....	46
Şekil 3 14. Sigmoid fonksiyonu ve matematiksel ifadesi	47
Şekil 4.1 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo.....	53
Şekil 4.2 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo	54
Şekil 4.3 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo	55
Şekil 4.4.hedef Zo ile M5P algoritmasının bulunduğu Zo.....	57
Şekil 4.5 d/h ve b/h in farklı değerlerine göre hedef Zo ve Bagging algoritması çıkış Zo	60
Şekil 4.6 a/h ve β nin farklı değerlerine göre hedef Zo ve Bagging Zo.....	61
Şekil 4.7.hedef Zo ile Bagging algoritmasının bulunduğu Zo	63
Şekil 4.8 b ve w nin farklı değerlerine göre hedef Zo ve MLP algoritması.....	66
Şekil 4.9. h_2 ve ϵ_r nin farklı değerlerine göre hedef Zo ve MLP algoritması çıkış Zo	67
Şekil 4.10.hedef Zo ile MLP algoritmasının bulunduğu Zo	69

ÇİZELGELER DİZİNİ

Çizelge 4.1. TSML de %66 eğitim, %34 test verileri ile yapılan uygulama sonucu elde edilen korelasyon ve hata.....	50
Çizelge 4.2. TSML de ikinci uygulama. %80 eğitim, %20 test verileri için hesaplanan korelasyon ve hata	51
Çizelge 4.3. Trapez şekilli hat M5P algoritması için ekstrapolasyon test sonuçları .	56
Çizelge 4.4. Trapez şekilli hat M5P algoritması için interpolasyon test sonuçları.....	56
Çizelge 4.5. V-Şekilli Mikroekranlı hat birinci uygulama. % 66 eğitim, %34 test verileri için hesaplanan korelasyon ve hata.....	58
Çizelge 4.6. V-şekilli mikroekranlı hat için ikinci uygulama %80 eğitim, %20 test verileri için korelasyon ve hata.....	59
Çizelge 4.7. V- şekilli mikroekranlı hat Bagging algoritması için ekstrapolasyon test sonuçları	62
Çizelge 4.8. Eliptik V- şekilli mikroekranlı hat Bagging algoritması için interpolasyon test sonuçları	62
Çizelge 4.9. Eliptik Şekilli Mikroekranlı hat birinci uygulama. % 66 eğitim, %34 test verileri için hesaplanan korelasyon ve hata.....	62
Çizelge 4.10. ESML de ikinci uygulama. %80 eğitim, %20 test veri oranı için korelasyon ve hata	65
Çizelge 4.11 Eliptik şekilli mikroekranlı hat MLP algoritması için ekstrapolasyon test sonuçları.....	68
Çizelge 4.12. Eliptik şekilli mikroekranlı hat MLP algoritması için interpolasyon test sonuçları	68

SİMGELER DİZİNİ

MİC	Mikrodalga entegre devre
MMİC	Monolitik mikrodalga entegre devre
KDT	Konform dönüşüm tekniği
DMH	Dairesel mikroekranlı hatların
AS-RSML	Havasız ortam dikdörtgen şekilli mikroekranlı hat
AS-TSML	Havasız ortam trapez şekilli mikroekranlı hat
PMM	Nokta karşılaştırma metodu
TEM	Transfers elektro-magnetik propagasyon modu
CMT	Uyumlu haritalama tekniği
CAD	Bilgisayar destekli dizayn
CD	Kompakt disk
CRM	Müşteri ilişkileri yönetimi
MAE	Ortalama mutlak hata
RMSE	Ortalama karesel hata
OLTP	Operasyonel veri sistemleri
ERP	Kurumsal kaynak planlama
MİH	Mikrodalga iletim hatları
CPW	Eşdüzlemlı dalga kılavuzları
CPS	Eşdüzlemlı şerit hatlar
IBk	instance-based algoritması
OLAP	Çevrimiçi Analitik İşleme
XML	Extensible Markup Language
VSCBCPW	V-şekilli iletken sonlu eşdüzlemsel dalga kılavuzu
CBCPW	İletken sonlu eşdüzlemsel dalga kılavuzunun
ESML	Eliptik şekilli mikroekranlı hat
CSML	Dairesel şekilli mikroekranlı hat
TSML	Trapez şekilli mikroekranlı hat
VSML	V-şekilli mikroekranlı hat
ARIMA	Otoregresif entegre hareketli ortalama

AIC	Akaike Information Criterion
SIC	Schwarz Information Criterion

1. GİRİŞ

Giderek küçülen dünyamızda artık mekân inanılmaz bir değer kazanmıştır. İnsanlar bir yandan teknolojinin yeni imkânlarından da vazgeçmek istemezken bir yandan da giderek artan dünya nüfusunu ve onun gereksinimlerini bu daracık dünyamıza sığdırmak istemektedir. Bu da bizi ister istemez aynı teknolojiyi çok daha küçük alanlara sığdırmak mecburiyet bırakmıştır. Günümüzde mikrodalga teknolojisi de bundan nasibini almış ve daha küçük alanda çalışacak mikrodalga elemanlara gereksinim artmıştır. Bu konuda mikroşerit hatlar kullanılmış ve hala kullanılmaktadır. Fakat birçok aktif mikrodalga elemanlarla paralel bağlantıda sıkıntılara sebep olduğu, daha küçük alanları gereksinim duyduğu, daha iyi yayılım parametreleri sağladığı için şerit ve mikroşerit hatlar yerine mikroekranlı hatların kullanımı tercih edilmektedir.

Bu hatların kullanılması için bu devrelerin analiz edilmesi, bu analizde bize doğru sonucu veren ve mümkün olduğunca hızlı çalışan modellerin elde edilmesi gerekmektedir. Bunun için son yıllarda yaygınlaşan soft çözümlene modelleri mikrodalga ve elektromagnetik devre çözümlenmelerinde kullanılmaya başlamıştır. Bu soft modellerle birlikte birçok çözüm algoritmalarını da bünyesinde barındıran veri madenciliği yöntemi tezimizde sunulmuştur.

Tezin amacı mikrodalga iletim hatlarını açıklamakta yeni bir yaklaşım ortaya koymak, böylece çözüm adına daha geniş perspektif oluşturulmasına katkı sağlamaktır. Bu şekilde işlenen verilerin daha hızlı ve verimli bir şekilde iletiminin ve kullanımının sağlanmasıdır.

Tezin birinci bölümünde kısa bir giriş yapılarak amaç ve yapılanma anlatıldı. İkinci kısımda hem mikroekranlı hat hem de veri madenciliği konularında kaynaklar ve önceden yapılan çalışmalar ilk günden buyana anlatılarak, bu çalışmalardaki yöntem avantaj ve dezavantajları bahsedilmiştir. Üçüncü bölümde mikroekranlı hatların çeşitleri ve yapılarından bunları etkileyen değişkenlerden, elde etmek istediğimiz sonuç değişken ve değerlerinden ve bunun matematiksel ifadelerinden

bahsedilmiştir. Bu bölümün devamında çözüm için yeni bir yaklaşım olan veri madenciliği yöntemi anlatıldı. Dördüncü bölümde bu yöntemle hesaplama yapıldı. Devamında sonuçlar ve yorum bulunmaktadır.

2.KAYNAK ÖZETLERİ

Çağlar boyunca bilgi sahibi olmak insanlar için büyük önem taşımıştır. Öyle ki bunun için çok uzun ve meşakkatli yollar çekmek zorunda kalanlar hiç de az değildir. Bunun için Güvercin yetiştirmek gibi çeşitli yollar bulunsa da kayıpların olması ve anlık bilgi elde edilememesi gibi dezavantajları bulunmaktadır. William Watson 1749 yılında elektrik akımının iletken tel üzerinde hareket ettiğini ortaya atmıştır. Bu ilk iletim hattı olarak kabul edilmektedir. Bundan yola çıkarak telgraf yapılabileceği düşünülmüş ancak yapılamamıştır.

1850 yılına gelindiğinde ilk denizaltı telgraf sistemi birkaç başarısız denemenin ardından kurulmuştur. Dış kısmına koruyucu amaçlı yalıtkan bir tabaka yapılmıştır. Ayrıca kopmaması için ek katmanlar konmuştur. Okyanus ötesi mesafelere bilgi transfer etmek için bu malzemelerin matematiksel gösterimleri bulunmalıydı. Bunu Stokes ve Thomson başarmıştır. Fakat bu gösterimler yeterli değildi. Dış kısımdaki yalıtkan maddenin de bir iletkenliği vardır. Ayrıca indüktansın da hesaba katılması gerekmektedir. Bunu da Heaviside başarmıştır.

II. Dünya savaşından sonra mikrodalga teknolojisinde ilerlemelerin başlangıcı olmuştur. Bu dönemde radar elde etmek için deneysel çalışmalar yapılmıştır. TEM koaksiyel kablo ve dalga kılavuzları o dönemde kullanılmıştır. Yalıtkan üzerine konan iletken şerit ile mikrodalga iletim hat meydana getirilmiştir. Bu da mikrodalga entegre devrelerin temel yapısıdır. Bu yapı mikrodalga iletimi yapmakla birlikte empedans uyumlaştırıcı, filtre vb birçok aktif eleman olarak da çalışabilir. Bu konuda yapılabilecek daha birçok farklı yapı ve işlev bulunmaktadır.

En içe iletken levha, onu etrafına dielektrik malzeme en dış iki yüzeye de iletken levha koyan Barrett ve Barnes 1951 yılında şerit hatları ortaya atmışlardır. Bu yapısı sayesinde dış iletken kısımlar kesilmeden de ara bağlantı, giriş çıkış port yerleştirmek mümkün olmuştur. Merkez iletken genişliği ile karakteristik empedansın ayarlanabildiği de 1956 yılında Peter tarafından bulunmuştur.

Grieg ve Engelmann (1952) ise bir başka tip iletim hattı olarak, mikroşerit hat olarak isimlendirilen yapıyı tasarlamışlardır. Bu yıllarda mikroşerit hatlar üzerine yoğun araştırmalar yapılmış ve tasarımlarda mikroşerit hatlar yaygın bir kullanım alanı bulmuşlardır. Mikrodalga devre elemanlarının silah, roket ve uydu uygulamalarında kullanılabilmesi için gerekli olan boyut küçültmelerine olan ihtiyacın artması, 60'lı yıllarda mikroşerit hatlara olan ilgiyi artırmıştır. Düşük kayıplı, yüksek dielektrik sabitli malzemeler ve metalik film yoğunlaştırılması ile ilgili olarak ortaya çıkan teknolojik gelişmeler, mikroşerit hatların kullanımını daha da yaygınlaştırmıştır. Ancak şerit ve mikroşerit hatların dielektrik tabanın karşı tarafına yerleştirilmiş olan toprak düzlemleri, birçok aktif mikrodalga devre elemanları için gerekli olan sönt bağlantılara erişimi engellemektedir. Karakteristik parametrelerinin, dielektrik tabanın kalınlığına direkt olarak bağlı olması şerit ve mikroşerit hatlarda yüksek dielektrik sabitli ve düşük kayıplı malzemelerin kullanımını da zorlaştırmaktadır. Bu durum boyut küçültülme çalışmaları için bir dezavantajdır. Bütün bu olumsuz gelişmeler, toprak düzlemi de dâhil olmak üzere bütün iletim elemanları dielektrik tabanın üst yüzeyinde yer alan iletim hatlarının geliştirilmesine neden olmuştur(Türkmen, 2009).

İletim hattı yapısına uygun olan bir mikrodalga entegre devre de yarık hatlardır. 1960 lı yıllarda ilk kez tasarlanan bu yapı sayesinde ara bağlantı deliklerine ihtiyaç kalmamıştır. Fakat bu hatlar iyi bir TEM dalga taşıyıcı değildir. Bu bunun için yüksek değerli dielektrik malzeme kullanmak gerekir. Fakat bu da mikrodalga entegre anten devrelerinde sıkıntıya neden olmaktadır

Şerit, mikroşerit ve yarık hatların yanında mikrodalga entegre devrelerde değişik geometrik yapılar da kullanılmıştır. Bunlar eşdüzlemlerli dalga kılavuzları, eşdüzlemlerli şerit hatlar ve mikroekranlı hatlardır. Mikrodalga entegre devre ve monolitik mikrodalga entegre devreler ile eşdüzlemlerli iletim hatları arası bağlantı kolaydır. Çünkü aynı düzlemi paylaşmaktadırlar. Eşdüzlemlerli dalga kılavuzları ve eşdüzlemlerli şerit hatlar ise mikroşerit ve şerit hatlar hatlara nazaran kolay imalat, kolay bağlantı, az kayıp gibi üstünlüklere sahiptir.

Hattımızın tasarımı açısından iki parametre hesaba katılmalıdır. Bunlar karakteristik empedans ve etkin dielektrik sabittir. Bunlar da iki durum için düşünölmelidir. Tam dalga analiz daha karmaşık hesapları gerektirtmesine rağmen daha iyi frekans tepkisi bulmaktadır. Quasistatik analiz ise daha basit hesaplanabilmesine rağmen hattın dispersif yapısını tam yansıtamamaktadır.

Yakın tarihimize geldiğimizde ise mikroelemanlı hatların empedans hesaplaması üzerine çok sayıda çalışmaya rastlamak mümkündür. Bunlardan ilki 1992 yılında Nihad ve arkadaşları tarafından yazılan mikroekranlı hat için empedans hesabı isimli makalesidir. Makalesinde mikroekranlı hat karakteristik empedansı nokta eşleştirme metodu (PMM) ve uygun haritalama tekniğı (CMM) ile elde edilmiştir. Ayrıca sonlu uzunlukta toprak düzlemlerin empedansa etkisi gösterilmiştir.

Hemen arkasından Yuan Naichag ve arkadaşları tarafından 1994 yılında V, eliptik, dairesel, şekilli mikroekranlı düzlemsel dalga kılavuzlarının tam TEM yayılım, sıfır dispersiyon, Alt ve üst sınır bölgelerinin geometrik ortalaması, grafik yaklaşımı ve uyumlu haritalama metodu ile karakteristik empedansları hesaplanmıştır.

Bu tarihe kadar hep simetrik hatlar kullanılmıştır. İlk kez 1995 yılında asimetrik hatları Kwok-Keung tarafından V şekil için hesaplamıştır. Aynı yıl ve aynı yazarlar tarafından dikdörtgen ve trapez şekilli hatların durgun durum hesabını da yapılmıştır.

2003 yılına gelindiğinde Mai Lu ve Paul j. Leonard tarafından v şekilli mikroekranlı hattın alan çizgilerinin sonlu elemanlar metodu ile hesabı anlatılmıştır. Alan çizgilerinin metalik strip ve dielektrik malzemeye bağılılığı gösterilmiştir. Bundan bir yıl sonra ise aynı yazarlar tarafından dikdörtgen şekilli mikroekranlı hatların kesim dalga boyu karakteristiğı ortaya konmuştur.

Bugüne kadar literatürde mikroekranlı hatların, veri madenciliğı yöntemi ile çalışma yapılmamıştır. Fakat veri madenciliğinde kullanılan bazı algoritmalar ile yapılan çalışmalar olmuştur. Bunlar genellikle soft hesaplama yöntemleridir. Karmaşık hesaplamalar yerine bilgisayarların da yardımıyla datalar ile sonuca gidilmektedir.

Hem kolay hem de gerçeğe çok yakın sonuçlar elde edilmiştir. 2005 yılından sonraki yıllarda artık soft hesaplama yöntemlerine yönelim artmıştır.

2006 yılında ise M.Türkmen ve arkadaşları tarafından Eliptik Ve Dairesel Şekilli Mikroekranlı Hatların Sinirsel Modelleri ortaya konmuştur. Bunun için 6 adet öğrenme algoritması uygulanmış en iyi sonucu Bayes (BR) sağlamıştır. Aynı yazarlar tarafından 2009 yılında , “Boşlukta Trapez Ve Dikdörtgen Şekilli Mikroekranlı Hatların Karakteristik Empedanslarının Bulanık Adaptif Ağ Metoduyla Hesaplanması” ve ”Dairesel Mikroekranlı Hatların Quasi-Statik Analizinin Yapay Sinir Ağları ile Gerçekleştirilmesi” ortaya konmuştur.

3.MATERYAL VE YÖNTEM

Mikroekranlı hatlar dielektrik gövdenin etrafı iletken ince malzeme ile kaplanarak ve üst kısmında iki açık bölge bırakılarak elde edilir. Açık bölge de ince dielektrik malzeme ile örtülerek elde edilen yapı yüksek frekans iletiminde kullanılmaktadır. Bu amaçla literatürde mikroşerit hatlar kullanılmıştır. Mikroşerit hatlara nazaran boyut küçültmesi, sönt bağlantı, girişim ve yayılım kaybının küçük olması vb avantajları sebebiyle mikroekranlı hatlar tercih edilmektedir.

3.1.Materyal

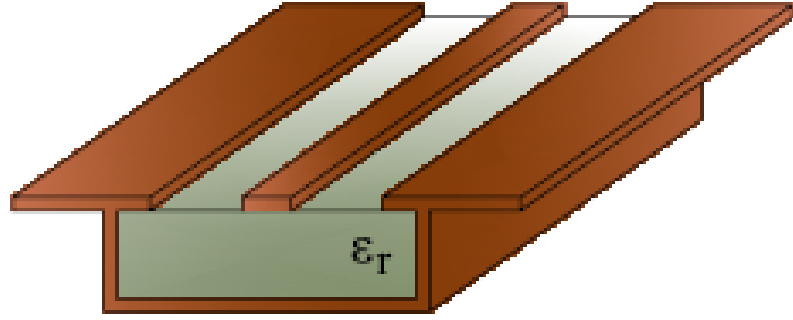
3.1.1.Havasız Ortam Dikdörtgen, Trapez, V, Eliptik Şekilli Mikroekranlı Hatlar

Bu çalışmada dört tip mikroekranlı hat yapısı üç başlık altında incelenecektir. Dikdörtgen şekilli mikroekranlı hat trapez şekilli mikroekranlı hattın $\beta=90$ değerindeki durumu olduğu için aynı başlık altında incelenecektir.

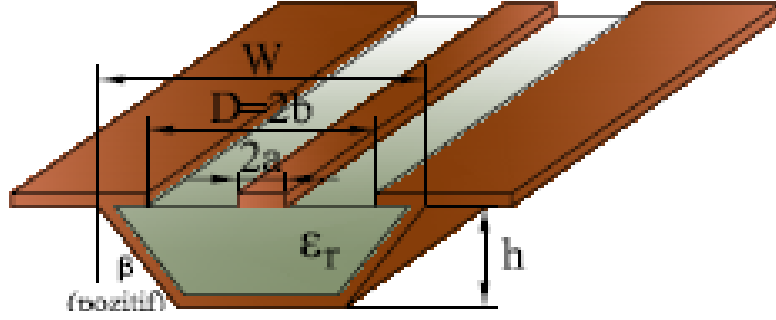
3.1.1.1. Dikdörtgen Ve Trapez Şekilli Mikroekranlı Hatlar

3.1.1.1.1, Havasız Ortam Dikdörtgen ve Trapez Şekilli Mikroekranlı Hatların yapıları

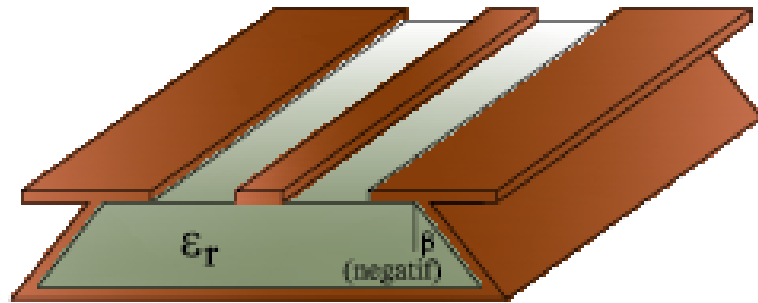
Şekil 3.1 Havasız ortam trapez şekilli mikroekranlı hat ve havasız ortam dikdörtgen şekilli mikroekranlı hat yatay kesitini göstermektedir. Bu şekilde $2a = D$ merkez dilim genişliğidir(Türkmen vb, 2009)



Şekil 3.1. Havasız ortam dikdörtgen şekilli mikroekranlı hatlarının geometrisi(Türkmen vb, 2009)



Şekil 3.2. Havasız ortam trapez şekilli (+) açılı mikroekranlı hatlarının geometrisi (Türkmen vb, 2009)



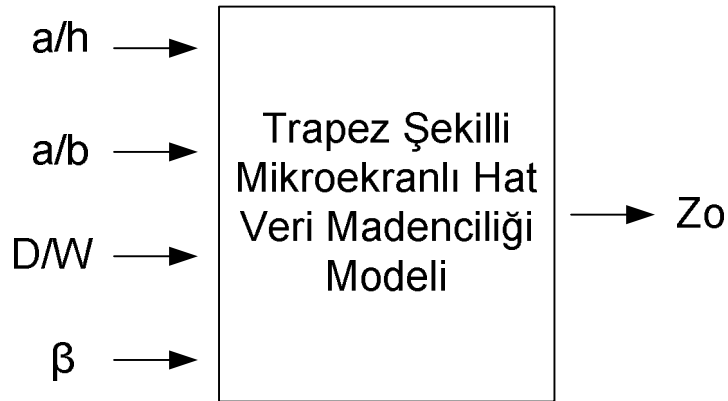
Şekil 3.3. Havasız ortam trapez şekilli (-) açılı mikroekranlı hatlarının geometrisi (Türkmen vb, 2009)

Mikroekranlı iletim hatlarının geometrileri şekil 3.1. de yer almaktadır. Bu şekilde, 2 a merkez iletken genişliğini, $D = 2b$ ise merkez iletken genişliği ile yarı genişlikleri toplamını, h dielektrik taban kalınlığını, r ise taban malzemesinin bağıl dielektrik

sabitini göstermektedir. Bu şekilde beta açısı pozitif ve negatif olabilen bir açı değeridir ve hava-süspansiyonlu mikroekranlı iletim hattı yapısı beta'ya bağlı olarak pozitif ve negatif yamuk mikroekranlı iletim hatlarına dönüşmektedir. Hava süspansiyonlu dikdörtgen kesitli mikroekranlı iletim hattı yapısı, hava-süspansiyonlu yamuk mikroekranlı iletim hattı yapısının özel bir halidir (beta = 0) (Türkmen, 2009)

3.1.1.1.2 Havasız Ortam Dikdörtgen ve Trapez Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı

Bu hesaplamayı yapabilmek için şekil 3.1. şekil 3.2. şekil 3.3. de geometrik yapısı verilen şekillerdeki a/h , a/b , D/W , β uzunlukları giriş olarak kabul edilirken Z_0 ise çıkış olarak kabul edilmektedir.



Şekil 3.4 Trapez şekilli mikroekranlı hattın blok şeması

Aşağıdaki formülü kullanarak: iletim hattı karakteristik empedansı hesaplanabilir.

$$Z_0 = \frac{1}{c \cdot \sqrt{C_T(\epsilon_r) \cdot C_T(1)}} \quad (3.1.)$$

ki burada

c : serbest uzayda ışık hızıdır.

$C(1)$: hava ortamındaki dielektrik sabite göre kapasitans değeridir.

$C_T(\epsilon_r)$:iletim hattının birim uzunluğu için kapasitans değeridir. $C_T(\epsilon_r)$ denklem 4.5 ışığında şöyle yazılabilir (Türkmen, 2009)

$$C_T(\epsilon_r) = 2\epsilon_0 \cdot \left[\frac{K(k_1)}{K(k_1')} + \epsilon_r \cdot \frac{K(k_2)}{K(k_2')} \right] \quad (3.2.)$$

ki burada

$K(k_i)$:ki modülünün birinci tür kapalı eliptik integralidir.

$K(k'_i)$:k'i modülünün birinci tür kapalı eliptik integralidir.

k'_i ise k_i nin bütünleyici modülüdür: (Türkmen, 2009)

$$k'_i = (1 - k_i^2)^{1/2} \quad (3.3.)$$

k_i ler mikroekranlı hattın (denklem 6-10) geometrik boyutlarının terimleri olarak ifade edilebilir:

$$k_1 = a/b \quad (3.4.)$$

ve

$$k_2 = t_a/t_b \quad (3.5.)$$

parametreler t_a ve t_b aşağıdaki denklemlerden iteratif nümerik metot kullanılarak elde edilebilir:

$$\frac{2a}{W} \cdot \Delta = \int_0^{\arcsin(t_a/t_c)} \frac{d\theta}{\cos^{1-2p} \theta (1 - t_c^2 \sin^2 \theta)^p} \quad (3.6.)$$

$$\frac{2b}{W} \cdot \Delta = \int_0^{\arcsin(t_b/t_c)} \frac{d\theta}{\cos^{1-2p} \theta (1 - t_c^2 \sin^2 \theta)^p} \quad (3.7.)$$

$$\frac{2h}{W} \cdot \Delta = \cos \beta \int_0^{\arccos h(1/t_c)} \frac{d\theta}{\sinh^{1-2p} \theta (1-t_c^2 \cdot \cosh^2 \theta)^p} \quad (3.8.)$$

$$\Delta = \int_0^{\pi/2} \frac{d\theta}{\cos^{1-2p} \theta (1-t_c^2 \sin^2 \theta)^p} \quad (3.9.)$$

$$p=1/2-\beta/\pi \quad (3.10.)$$

t_c değeri iteratif nümerik metot kullanılarak elde edilebilir. t_c , denklem 5c şartıyla 0 ile 1 arasında değişen değerlerdir. Yine denklem 5c ile t_a ve t_b değerleri tatmin edici seviyede hesaplanabilmektedir (Türkmen, 2009)

Literatürde, farklı yapılarda birçok Mikroekranlı iletim hattı birçok araştırmacı tarafından incelenmiştir. Mikroekranlı iletim hattı, boşluk bölgesi integral metodu kullanılarak incelenmiştir. Sonra Mikroekranlı iletim hattının karakteristik empedansı nokta karşılaştırma metodu (PMM) ve uyumlu haritalama tekniği ile ortaya konmuştur. Durağan analizde, V şekilli mikroekranlı hattın TEM propagasyon parametrelerini belirlemek için moment metodu kullanılmıştır. V, eliptik, dairesel şekilli karakteristik empedansı, 0 dispersiyon ve tam TEM propagasyon varsayılarak, uyumlu haritalama tekniği metoduyla bulunmuştur. Havasız ortamda Mikroekranlı iletim hatların karakteristik empedanslarının nümerik çözümleri CMT metoduyla gösterilir. Bilgisayar destekli dizayn CAD temelli ifadeler Dikdörtgen şekilli mikroekranlı hat analizi için de önerilir. Dahası dikdörtgen şekilli mikrohattın hat parametre duyarlılık analizi, tam olmayan duvar etkisinin giderilmesi(denklem 6,7,8,9) işlemiyle yapılmıştır (Türkmen, 2009)

3.1.1.1.3. Veri Madenciliğinin Dikdörtgen ve Trapez Şekilli Mikroekranlı hatta uygulanması

Yukarıdaki 1-10 nolu denklemlerde, belirtildiği şekilde yerleştirilerek aşağıdaki 4 adet girdi ve 1 adet çıktı değerleri hesaplanmıştır.

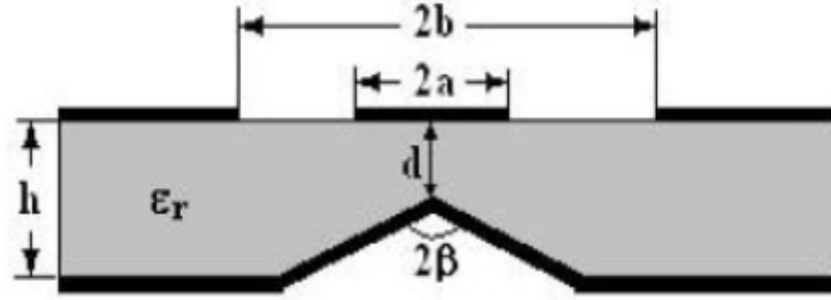
Girdi ve ıktılar Őekildeki gibi bir izelge oluŐturulmuŐtur ve denklemlerle sadece girdiler (a/h, a/b, D/W, beta ve sonucu (Zo) elde edilmiŐtir.. Elde ettiĐimiz bazı deĐerleri data olarak kullanarak hesaplaması daha kolay olan veri madenciliĐi ile özümü inceleyeceĐiz. Buradaki ok fazla sayıda dataları, alıŐmamızın sayfalarının alması mümkün olmamaktadır.

Datalar elde edildikten sonra Bu dataları kullanılabilir hale getirilmelidir. Weka programı kullanılacaĐı için onun kullanacaĐı dosya formatları tercih edilmeli .arff, .csv, .dat, .libsvm, .data, .names, .arff.gz, .bsi, .xrff, .xrff.gz, uzantılı dosyaları kullanabilmektedir. Kolay elde edilebilmesi için .csv uzantılı dosya tercih edilmiŐtir. Elde edilen datalar Excel dosyasında farklı kaydet komutuyla. csv seilerek kaydedilmiŐtir. Bylece wekanın algılayabileceĐi csv dosyası elde edilmiŐtir. Weka programında Explorer aılarak iinde yeni dosya aılmıŐ ve csv dosyasının konumuna gidilerek datalar wekaya okutulmuŐtur.

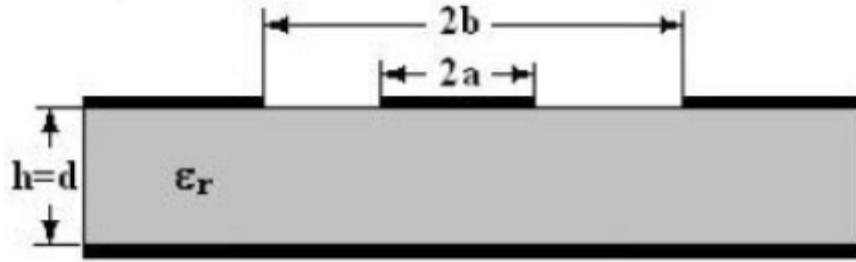
Daha sonra classfy menüsünden classifier kısmından choose butonu tıklanarak, burada classifierın altında bulunan algoritma gruplarından istenilen algoritmalar seilerek program yürütölmüŐtür. Seilen algoritmalar deĐerlendirme performans ölçütleri aısından kıyaslanarak uygun algoritma bulunmuŐtur. Bulgular bir sonraki bölümde incelenmiŐtir.

3.1.1.2. V-Şekilli İletken Sonlu Mikroekranlı Hatlar

3.1.1.2.1 V-Şekilli İletken Sonlu Mikroekranlı Hatların Yapıları



Şekil 3.5. CSVBVPW'nin yatay kesiti (Türkmen, vb2006).

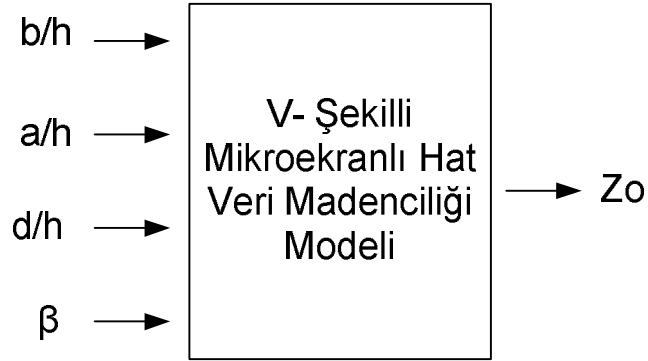


Şekil 3.6. CBCPW'nin yatay kesiti (Türkmen, vb 2006).

Şekil 3.5 ve şekil 3.6 V-şekilli iletken sonlu eşdüzlemsel dalga kılavuzu (VSCBCPW) ve iletken sonlu eşdüzlemsel dalga kılavuzunun (CBCPW) yatay kesitini göstermektedir. VSCBCPW, 2b genişliğindeki iki adet üst toprak düzlemleri ve onların arasındaki 2a genişliğinde iletken yüzeyi içerir. Bunlar h kalınlığındaki dielektrik sabiti ϵ_r olan alt tabaka malzeme üzerine yerleştirilmiştir. İki toprak katmanları arası açı 2β dir. Şekil 3.5 ve şekil 3.6 de görüldüğü gibi CBCPW, VSCBCPW'in $\beta = 90^\circ$ ve $d/h = 1$ değerleri için bir özel durumudur (Türkmen,2006).

3.1.1.2.2 Havasız Ortam V-Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı

Bu hesaplamayı yapabilmek için şekil 3.5., şekil 3.6. de geometrik yapısı verilen şekillerdeki b/h , a/h , d/h , β uzunlukları giriş olarak kabul edilirken Z_0 ise çıkış olarak kabul edilmektedir.



Şekil 3.7. V- şekilli mikroekranlı hattın blok şeması

Merkez iletkenin ve iki üst toprak düzlemin mükemmel magnetik duvar gibi davrandığı varsayılır. Bu, merkez iletkenin iki üst toprak düzlemleri yoluyla hiçbir elektrik alan çizgilerinin sızmadığını kabulü anlamına gelir. Birim uzunluk için toplam kapasitans değeri (Türkmen,2006):

$$C_T(\epsilon_r) = 2\epsilon_0 \frac{K(k_1)}{K(k_1')} + \epsilon_r \frac{K(k_2)}{K(k_2')} \quad (3.11.)$$

ki burada $K(k_i)$ ve $K(k_i')$ ler, k_i ve k_i' modülleri ile birlikte birinci tip karmaşık eliptik integrallerdir. k_i' nin modülüdür. $k_i'^2$ ise k_i nin bütünleyici modülüdür:

$$k_i'^2 = (1 - k_i^2)^{1/2} \quad (3.12.)$$

k_i ler VSCBCPW in geometrik boyutları cinsinden ifade edilirse.

$$k_1 = a/b \quad (3.13.)$$

$$k_2 = \frac{\sqrt{1 + t_c / t_b}}{\sqrt{1 + t_c / t_a}} \quad (3.14.)$$

Denklem 14 de verilen ara parametreler t_a , t_b , t_c ler aşağıdaki formüller ile iteratif nümerik metot kullanarak çözülebilir (Türkmen,2006).

$$\frac{a}{h} = \frac{2}{\pi} \int_0^{\arcsin \sqrt{t_a/t_c}} \left(\frac{t_c \cosh^2 \theta}{1 + t_c \sin^2 \theta} \right)^p d\theta \quad (3.15.)$$

$$\frac{b}{h} = \frac{2}{\pi} \int_0^{\arcsin \sqrt{t_b/t_c}} \left(\frac{t_c \cosh^2 \theta}{1 + t_c \sinh^2 \theta} \right)^p d\theta \quad (3.16.)$$

$$\frac{d}{h} = \frac{2}{\pi} \int_0^{\pi/2} \left(\frac{t_c \cosh^2 \theta}{1 + t_c \sin^2 \theta} \right)^p d\theta \quad (3.17.)$$

Bunlarla birlikte;

$$p = 1/2 - \beta/\pi \quad (3.18.)$$

t_a , t_b , t_c değerleri 0 ile 1 arasında değişir. Etkin dielektrik sabit ve karakteristik empedans 11-18 nolu denklemlerin yanında aşağıdaki denklemlerle hesaplanır (Türkmen,2006).

$$\epsilon_{eff} = \frac{C_T(\epsilon r)}{C_T(1)} \quad (3.19.)$$

$$Z_0 = \frac{1}{c_v \sqrt{C_T \epsilon(r) \cdot C_T(1)}} \quad (3.20.)$$

Ki burada c_v boşluktaki ışık hızı ve $C_T(1)$ de dielektrik materyalin hava ile değiştirildiği durum için kapasitans değeridir (Türkmen,2006).

3.1.1.2.3. Veri Madenciliğinin V- Şekli Mikroekranlı hatta uygulanması

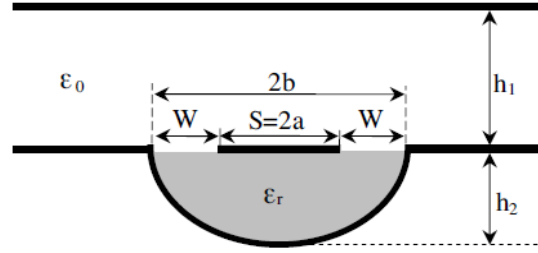
Çizelge 3.3 de yapıldığı gibi veri madenciliği programına uygun dataları aşağıdaki Çizelgeye uygun şekilde elde edildi. Dört adet girdimiz ve bir adet çıktımız vardır.

Verilen girdiler ve çıktı için 360 ar adet veri elde edildi. Fakat sayfamıza sığamayacak kadar fazla olduğu için dataları yazamıyoruz. Çizelgedeki datalar elde edildikten sonra Bu dataları kullanılabilir hale getirilmiştir. Weka programı kullanabilmesi için onun kullandığı dosya formatları tercih edilmiştir. .arff, .csv, .dat, .libsvm, .data, .names, .arff.gz, .bsi, .xrff, .xrff.gz, uzantılı dosyaları kullanabilmektedir. Kolay elde edilebilmesi için .csv uzantılı dosya tercih edilmiştir. Elde edilen datalar Excel dosyasında farklı kaydet komutuyla .csv seçilerek kaydedilmiştir. Böylece wekanın algılayabileceği csv dosyası elde edilmiştir. Weka programında Explorer açılarak içinde yeni dosya açıldı ve csv dosyasının konumuna gidilerek datalar wekaya okutulmuştur.

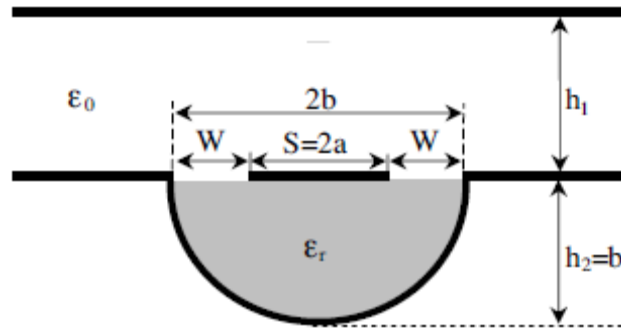
Daha sonra classfy menüsünden classfier kısmından choose butonu tıklanmıştır. Burada classfierın altında algoritma grupları bulunmaktadır. Bu gruplardan istenilen algoritmalar seçilmiş ve program yürütülmüştür. Seçilen algoritmalar model değerlendirme performans ölçütleri açısından kıyaslanarak uygun algoritma bulunmuştur.

3.1.1.3. Eliptik Şekilli İletken Sonlu Mikroekranlı Hatlar

3.1.1.3.1 Eliptik Şekilli İletken Sonlu Mikroekranlı Hatların Yapıları



Şekil 3.8 ESML yatay kesiti(Türkmen, 2008).



Şekil 3.9 CSML yatay kesiti(Türkmen, 2008).

Eliptik şekilli mikroekranlı hat (ESML) ve dairesel şekilli mikroekranlı hat(CSML) nin yatay kesiti ve üzerindeki iletken örtü şekil 3.8 şekil 3.9de gösterilmiştir. ESML, $2b$ genişliğindeki iki adet üst toprak düzlemleri ve onların arasındaki $2a$ genişliğinde iletken yüzeyi içerir. Bunlar h kalınlığındaki dielektrik sabiti ϵ_r olan alt tabaka malzeme üzerine yerleştirilmiştir. İletken örtü arası mesafe h_1 dir. Şekil 3.8. şekil 3.9. de görüldüğü gibi CSML, ESML nin $h_2 = b$ değerleri için bir özel durumdur. Merkez iletkenin ve iki üst toprak düzlemin mükemmel magnetik duvar gibi davrandığı varsayılır (Türkmen, 2008).

3.1.1.3.2 Havasız Ortam Eliptik Şekilli Mikroekranlı Hatların Karakteristik Empedansı Hesabı



Şekil 3.10. Eliptik şekilli mikroekranlı hattın blok şeması

Bu hesaplamayı yapabilmek için şekil 3.8., şekil 3.9. de geometrik yapısı verilen şekillerdeki b , w , h_1 , h_2 uzunlukları giriş olarak kabul edilirken Z_0 ise çıkış olarak kabul edilmektedir.

Bu, merkez iletkenin iki üst toprak düzlemleri yoluyla hiçbir elektrik alan çizgilerinin sızmadığı kabul edilir. Birim uzunluk için toplam kapasitans değeri (Türkmen,2008):

$$C_T(\epsilon_r) = C_1 + C_2 \quad (3.21.)$$

ki

$$C_1 = 2\epsilon_0 \frac{K(k_1)}{K(k_1')} \quad (3.22.)$$

ki burada $K(k_i)$ ve $K(k_i')$ ler, k_i ve k_i' modülleri ile birlikte birinci tip kompleks eliptik integrallerdir. k_i' ise k_i nin bütünleyici modülüdür:

$$k_1' = (1 - k_1^2)^{1/2} \quad (3.23.)$$

k_1 ler ESML ın geometrik boyutları cinsinden ifade edilirse

$$k_1 = \frac{\tanh\left(\frac{\pi.S}{4h_1}\right)}{\tanh\left(\frac{\pi.(S+2W)}{4h_1}\right)} \quad (3.24.)$$

$$C_2 = \pi.\varepsilon_0.\varepsilon_r \left[\ln\left(\frac{b+h_2}{a}\right) + \ln(1-B') \right]^{-1} \quad (3.25.)$$

$$B' = \frac{(\sqrt{b^2 - h_2^2} - a)^2 (b - h_2)}{2h_2.a} \quad (3.26.)$$

ESML ın etkin dielektrik sabiti ve karakteristik empedans değerleri

$$\varepsilon_{\text{eff}} = C_T(\varepsilon_r)/C_T(1) \quad (3.27.)$$

$$Z_0 = \frac{1}{c.\sqrt{\varepsilon_{\text{eff}}}.C_T(1)} \quad (3.28.)$$

Ki burada c_v boşluktaki ışık hızı ve $C_T(1)$ de dielektrik materyalin hava ile değiştirildiği durum için kapasitans değeridir(Türkmen,2008).

Yukarıdaki 3.21.-3.28. nolu denklemlerde, belirtildiği şekilde yerleştirilerek aşağıdaki 4 adet girdi ve 1 adet çıktı değerleri hesaplanması gerekmektedir.

Denklemlerle sadece girdileri (b, w, h1, h2) ve sonucu (Zo) elde edilmiştir. Elde ettiğimiz bazı değerleri data olarak kullanarak hesaplaması daha kolay olan veri madenciliği ile çözümü incelenmiştir.

Yukarıdaki çizelgede verilen girdiler ve çıktı için 1200'er adet veri elde edildi. Fakat sayfamıza sığamayacak kadar fazla olduğu için dataları yazamıyoruz.

Datalar elde edildikten sonra bu datalar kullanılabilir hale getirilmiştir. Weka programı kullanıldığı için onun kullandığı dosya formatları tercih edilmiştir. Weka; .arff, .csv, .dat, .libsvm, .data, .names, .arff.gz, .bsi, .xrff, .xrff.gz, uzantılı dosyaları kullanabilmektedir. Kolay elde edilebilmesi için .csv uzantılı dosya tercih edilmiştir. Elde edilen datalar Excel dosyasında farklı kaydet komutuyla .csv seçilerek kaydedilmiştir. Böylece wekanın algılayabileceği csv dosyası elde edilmiştir. Weka programında Explorer açılarak içinde yeni dosya açılmış ve csv dosyasının konumuna gidilerek datalar wekaya okutulmuştur.

Daha sonra classify menüsünden classifier kısmından choose butonu tıklanır. Burada classifierin altında algoritma grupları bulunmaktadır. Bu gruplardan istenilen algoritmalar seçilerek program yürütülmüştür. Seçilen algoritmalar performans değerlendirme ölçütleri açısından kıyaslanarak uygun algoritma bulunmuştur.

3.2. Yöntem

Bu tezde mikroekranlı hatların modellenmesi için veri madenciliği yöntemi kullanılmasına karar verildi.

3.2.1. Veri Madenciliği

3.2.1.1 Tanımı

1950'li yıllardan sonra bilgiye olan ihtiyaç hızla artmaya başlamıştır. Bilgisayarlar bilginin saklanması ve istendiği zaman erişilebilmesinde kullanılmıştır. 1960'larda basit dosya sistemleri, 1970'lerde veri tabanları, 1980'lerde büyük boyutlardaki nesne tabanlı veritabanları geliştirilmiştir. 1990'lı yıllarda veri ambarı ve VTBK (veritabanlarında bilgi keşfi) terimleri ortaya atılmıştır

1995 yılında düzenlenen “Knowledge Discovery in Databases” konferansının bildiri kitabı sunusunda, enformasyon teknolojilerinin oluşturduğu veri dağları aşağıdaki cümleler ile vurgulanmaktadır (Akpınar, 2000). “Dünyadaki enformasyon miktarının her 20 ayda bir ikiye katlandığı tahmin edilmektedir. Bu ham veri seli ile ne yapabiliriz? İnsan gözleri bunun ancak çok küçük bir kısmını görebilecektir. Bilgisayarlar bilgeliğin pınarı olmayı vaat etmekte, ancak veri sellerine neden olmaktadır.

Özellikle internet bağlantılarının ortaya çıkması ve yaygınlaşmasıyla veri miktarlarında akıl almaz ölçülerde artış meydana gelmiştir. Her ne kadar raporlama ve sorgulama araçları olsa da bu kadar veri arasındaki gizli ilişkilerin ortaya konması veri madenciliği yöntemleri ile gerçekleşmiştir. Veri madenciliği konusunda çok fazla çalışma ve kaynak bulunmaktadır. Bu kaynaklar tarafından çok sayıda tanımlama yapılmıştır. Veri madenciliği çok sayıda bilgi yığınları içerisinde belirli bir amaç doğrultusunda, bilgisayar ve çeşitli çözümlenme yöntemleri kullanılarak anlamlı bilgi elde edilmesidir.

Bu tanımlamalardan bir tanesi de Piatetsky-Shapiro ya aittir. Shapiro: “verilerden daha önceden bilinmeyen, gizli ve muhtemelen faydalı bilginin monoton olmayan bir süreçte çıkartılması işlemi” olduğunu söylemiştir. Larose da veri madenciliğini veriler arasındaki anlamlı ve yararlı ilişkileri tespit ederek veriyi yeni bir şekilde özetlemek şeklinde ifade etmektedir. Bunun yanında koyuncu, ‘Veri kümesi içerisinde keşfedilmemiş örüntüleri bulmayı hedefleyen teknikler koleksiyonunu betimleyen veri madenciliği, geçmiş faaliyetlerin analizini temel alarak gelecekteki davranışların tahminine yönelik karar verme modelleri oluşturmayı amaçlamaktadır.’ demektedir.

3.2.1.2. Gereksinimleri

Veri madenciliğinin temel gereksinimleri şunlardır (Asilkan Ö. 2008):

- Bu konuda uzman kişi(ler),

- Yeterli büyüklükte veri,
- Veritabanı ve/veya veri ambarı,
- Yüksek performanslı bilgisayarlar,
- Veri madenciliği yazılımları,
- Etkin algoritmalar.

Olarak kısaca özetlenebilir.

3.2.1.3. Amaçları

Veri madenciliğinin dört temel amacı bulunmaktadır (Fayyad vd., 1996):

- 1) Sınıflandırma: Verinin daha önce tanımlanmış belirli kategoriler içerisine yerleştirilmesi.
- 2) Tahminleme: Bilinen özelliklerden yola çıkılarak, bilinmeyenlerin ortaya çıkarılması.
- 3) Kümeleme: Benzerliklere dayalı olarak nesnelerin doğal gruplarının araştırılması.
- 4) Örüntü tanıma: İşlemsel veri içerisinde yer alan yararlı örüntülerin ortaya çıkarılması.

3.2.1.4. Karşılaşılan Problemler

Her yöntemde olduğu gibi bu yöntemde de girdiler ve çıktılar olmaktadır. Bu yöntemin girdileri el altında hazır bulundurduğumuz verilerimizdir. Bu verilerimiz işlemlerden geçirilerek işe yarar bilgiler haline getirilip veri ambarlarına aktarılırlar.

Büyük boyutlu veritabanı: Veritabanlarının hacmi iki boyutlu olarak genişlemektedir:

Yatay Boyut: nesnelerin özellik sayılarındaki artış,

Dikey Boyut: nesnelerin kayıt sayısındaki artış.

Örneklerin alındığı evrensel uzayın geniş olması, bu uzaydaki tüm elemanların kapsanmasını sağlamaktadır. Bu da çözüm yöntemlerinin kaçırdığı alan kalmamasını sağladığı için bir avantajdır. Fakat bunun da bir bedeli vardır. Bu esnada veri sayısı artmakta ve dolayısıyla bu verilerin işlenme süreleri de o oranda artmaktadır. Bu

problemi gidermek için YSA gibi yaklaşımlarla evrensel uzayı tarayarak bütünü yansıtabilecek şekilde veri sayısı azaltılabilir. Ya da bunu uygulayacak kişi tarafından gereksiz görülen yatay ve dikey nesnelere azaltılacak

Gürültülü veri: verilerin işleme sokulmadan önce bir şekilde verilerde meydana gelen istem dışı değişikliklere gürültü, bu haldeki veriye ise gürültülü veri diyoruz. Eğer buna sebep olan faktörleri ve etkilenmiş veriler tanımlanabilirse bir nebze de olsa gürültü azaltılabilir.

Artık veri: Verilerimizin bir kısmı ise aynı yöntemlerle elde edilmiş, gürültüye uğramamış olsalar bile problem çözümüne katkı sağlayacak nitelikte değildirler. Bunlar hem yer kaplamakta hem de süreci yavaşlatmaktadır. Ayrıca sayıları çoğaldığında çözüme ulaşmaya engel olmaya başlamaktadırlar. Bu verilerin ortak özellikleri belli algoritmalarca tespit edilip artık verilerin uzaklaştırılması sağlanabilir.

Dinamik veri: İçindeki verilerin sık sık değiştiği ortamlar da bulunmaktadır. Aktif olarak çalışan işyerleri buna örnektir. Burada değişen verilerin veri madenciliği sistemine sürekli entegre edilmesi gerekmektedir. Bu da işlemleri biraz daha arttırmakta ve yavaşlatmaktadır. Fakat gerçek hayatta bu veriler hakimdir ve gerçeğe yakın sonuçlar bulmak için bu entegrasyon uygulanmalıdır.

Eksik veri: örneklerin alındığı kümedeki verilerin bir kısmının olmaması durumudur. Bu ya yatay boyutta özellikler satırında bir eksiklik olması ya da düşey boyutta verilerin eksik olması halinde gerçekleşmektedir. Bir şekilde verilerin alınmadığı durumlarda oluşmaktadır.

Boş değerler: tanımlanmış standart niteliklerin dışında olan veriler için kullanılan bir kavramdır. Uygulanacak algoritmada veri sayıları belirlendiği şekilde olması gerektiği için boş veriler de kullanılabilir fakat mümkünse boş verileri ihmal etmek ya da geçerli verilerle değiştirmek gerekmektedir.

3.2.1.5. Kullanım Alanları

Günümüzde veri madenciliği astronomi, biyoloji, finans, pazarlama, sigorta, tıp ve birçok başka dalda uygulanmaktadır. Son 20 yıldır Amerika Birleşik Devletleri'nde çeşitli veri madenciliği algoritmalarının gizli dinlemeden, vergi kaçakçılıklarının ortaya çıkartılmasına kadar çeşitli uygulamalarda kullanıldığı bilinmektedir (Dilly, 1999). Bununla birlikte günümüzde veri madenciliği yöntemleri özellikle işletmelerin çeşitli alanlarında başarı ile kullanılmaktadır. Bu uygulamaların başlıcaları ilgili alanlara göre aşağıda gösterilmiştir (Akpınar vb, 2000):

- Pazarlama

Müşteri gruplandırması

Müşterilerin satın alma örüntülerinin belirlenmesi

Posta kampanyalarında cevap verme oranının artırılması

Mevcut müşterilerin elde tutulması için pazarlama stratejilerinin geliştirilmesi

Pazar sepeti analizleri

Çapraz satış analizleri

Müşteri değerlendirme

Müşteri ilişkileri yönetimi

Satış tahminleri

- Bankacılık

Farklı finansal göstergeler arasındaki gizli korelasyonların bulunması

Kredi kartı dolandırıcılıklarının tespiti

Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi

Kredi taleplerinin değerlendirilmesi

Risk analizleri

- Sigortacılık

Yeni poliçe talep edebilecek potansiyel müşterilerin tahmin edilmesi

Sigorta dolandırıcılıklarının tespiti

Riskli müşterilerin belirlenmesi

- Perakendecilik

Satış noktası veri analizleri

Alışveriş sepeti analizleri

- Borsa

Hisse senedi fiyat tahmini

Genel piyasa analizleri

Alım-satım stratejilerinin optimizasyonu

- Sağlık

Ürün geliştirme

Teşhis ve tedavi

Test sonuçlarının ya da ilaçların hastalar üzerindeki etkilerinin tahmini

- Eğitim

Öğrenci davranışlarının tahmin edilmesi

Öğrencilerin performanslarına göre branşlara ayrılması

Öğrencilerin ders seçme eğilimlerinin belirlenmesi

- Turizm

Bir bölgeye gelen turistlerin o bölgeyi tercih sebeplerinin belirlenerek ilişkili reklam kampanyası düzenlenmesi

Gelecek sezonlardaki turist sayısının tahmini ve buna yönelik kapasite planlaması

- İnternet

Kullanıcıların profilini belirleyip ziyaret ettiği sayfalarda onların ilgilerine uygun ürün reklamlarının gösterilmesi

Maillere gelen yanıt oranının arttırılması

Veri madenciliğinin bankacılık, perakendecilik, sigortacılık, borsa vb gibi alanlarda uygulanabildiğinden bahsedilmiştir. Bunlar da gösteriyor ki ticaret sahasında veri madenciliğinin çok geniş uygulama alanları bulunmaktadır. Ayrıca tüm verilerin zaten kayıt altında olması hatta anlık olarak elde edilebilmesi ve dijital ortamda olduğu için diğer dijital uygulamalara direk bağlanabilmesi yönlerinden veri madenciliği açısından çok büyük imkân ve avantajlara sahiptir.

Veri madenciliği genellikle aşağıdaki karakteristiklere sahip problemlerin çözümünde tercih edilmektedir (Han ve Kamber, 2001):

- Büyük miktarlarda veriye erişilebildiği zaman
- Veri çok sayıda değişkene sahipse
- Veri karmaşık, çok değişkenli ve doğrusal olmayan yapıdaysa
- Çıktıları ya da davranışı tahmin etmek gerekiyorsa
- Henüz anlaşılmayan birliktelik ve ilişkileri bulmak gerekiyorsa

3.2.1.6. Etkilendiği Unsurlar

Veri madenciliği 5 temel harici unsurdan etkilenmektedir (Asilkan Ö. 2008):

1) Veri: Son yirmi yılda hızla artan sayısal veri sayısı veri madenciliğindeki gelişmeleri hızlandırmıştır. Veri sayısının çok hızlı oranda artmasına rağmen, bu verilerle uğraşan bilim adamı sayısının önemli seviyede artmaması nedeniyle verileri analiz etme yöntemleri geliştirilmektedir.

2) Donanım: Veri madenciliği, büyük veri kümeleri üzerinde yoğun işlemler yapmayı gerektirir. Gelişen bellek ve işlem hızı kapasitesi sayesinde, birkaç yıl önce madencilik yapılamayan veriler üzerinde çalışmak mümkün hale gelmiştir.

3) Bilgisayar Ağları: Hızı ve kapasitesi hızla artan bilgisayar ağları ve internet, dağıtık verileri analiz etmeyi ve farklı algoritmaları kullanmayı mümkün hale getirmektedir.

4) Bilimsel Hesaplamalar: Günümüzde bilim adamları, simülasyonu teori ve deneyden sonra bilimin üçüncü yolu olarak görmektedirler. Veri madenciliği ve bilgi keşfi, bu üç metodu birbirine bağlamada önemli rol almaktadır.

5) Ticari Eğilimler: Bilgisayar ve internet teknolojilerinin hızlı ilerlemesi sonucu günümüzde ticaret hızı ve hacmi büyük boyutlara ulaşmış, ticaretin önemli bir bölümü elektronik ortamda gerçekleştirilir olmuştur. E-ticaret; minimum maliyet ve en az insan gücünü kullanarak maksimum karlılığı hedeflemektedir. Bu ticari eğilimler ve hedefler, veri madenciliğinin kullanımını neredeyse zorunlu hale getirmiştir. Zira, veri madenciliği sayesinde müşteri işlemlerinin doğurduğu fırsat ve riskler kolaylıkla ve açıkça görülebilmektedir.

3.2.1.7. Modelleri

Veri madenciliğinde çeşitli çözüm yöntemleri kullanılmaktadır. bu modeller iki temel kısma ayrılabilir. Birincisi tahmin edici modellerdir. Bunlara örnek sınıflama ve regresyon modelleridir. İkincisi tanımlama modelleridir. Bunlara örnek kümeleme, birliktelik kuralları, ardışık zamanlı örüntülerdir. Koyuncugil'e göre Yazında tahmin edici modeller “denetimsiz modeller”, tanımlayıcı modeller ise “denetimli modeller” olarak da ifade edilmektedir. Bu iki model arası fark kesin sınırlarla ayrılmış değildir.

Tahmin edici modellerde elimizde belli miktarda veri bulunur Bu veriler ile yine belli miktarda bilinmeyen ve bulunması istenen verilerin elde edilmesi amaçlanır. Bu amaçla eldeki veriler tahmin edici veri madencilik modellerine tabi tutularak istenen verilerin tahmini değerleri elde edilir. Örneğin elektrik hammaddesine bağımlı olarak çalışan bir işyerinin gelecek yıllarda yapacağı yatırımlara yön vermek için gelecekte elektrik fiyatlarının ne olacağını geçmiş verilere dayanarak, tahmin edici veri madencilik modelleriyle yaklaşık olarak hesaplamasıdır.

Tanımlayıcı modellerde ise hazır bulunan verilerin ne anlama geldiği, bu veriler arasında nasıl bir ilişki olduğunun bulunması amaçlanır. Örneğin, vücut bakım ürünleri satışı yapan bir işyeri satış rakamlarını arttırabilmek için müşterilerinin hangi ürünleri aldığı, nasıl kullandığı, vb gibi veriler yardımıyla bu ürünler arası ilişkileri tahmin eden tanımlayıcı veri madenciliği modellerini uygulayabilir.

3.2.1.7.1. Sınıflama ve Regresyon Modelleri

Sınıflama ve regresyon modelleri çok sık karşılaşılan veri madencilik modelleridir. Bunlar tahmin edici modellerdir. Sınıflama ve regresyon modelleri arasında çeşitli farklılıklar bulunmaktadır. Bunlardan en önemlisi sınıflandırılabilir bir değerinin olmasıdır. Sınıflandırma kendi özelliklerine benzer yapıdaki verilerin bir araya

toplanmasıdır. Fakat bu benzer yapı en başta kararlaştırılmıştır. Bir araya toplanan veriler, belli bir sınıf yapısı oluştururlar. Bu sınıftaki veriler tarafından eğitilen sistem bilinmeyen verileri bu sınıflardan birine dâhil eder. Böylece bilinmeyen veriler hakkında tahmini bilgi sahibi olunmaktadır.

Regresyon modelinde ise verinin diğer verilerle arasında bir ilişki olduğu kabul edilir. Bu ilişki yardımı ile elde etmek istediğimiz veriler hakkında tahmini bilgi sahibi olunmaktadır.

Veri madenciliğinde kullanılan modeller hız, yakın tahmin, hatalı veriye direnç, veri sayısının sonucu bozmaması, anlaşılma seviyesi gibi kriterlerle kıyaslanabilir.

Han ve Kamber (2001), sınıflandırma ve regresyon yöntemlerinde aşağıdaki karşılaştırma kriterlerini önermektedir:

- Tahmin doğruluğu: Modelin yeni ya da daha önceden bilinmeyen bir verinin sınıf etiketinin doğru olarak tahmin edilme yeteneğidir.
- Hız: Modelin geliştirilmesindeki hesaplama maliyetine karşılık gelir.
- Sağlamlık: Modelin, gürültülü ya da kayıp veri verildiğinde doğru tahmin yapma yeteneğidir.
- Ölçeklendirilebilirlik: Büyük miktarlarda veri verildiğinde, modeli etkin bir şekilde kurma yeteneğidir.
- Yorumlanabilirlik: Model tarafından sağlanan anlaşılma seviyesidir.

Sınıflandırma ve regresyon modellerinde kullanılan başlıca yöntemler aşağıdaki gibi sıralanabilir (Akpınar, 2005):

- Karar Ağaçları (Decision Trees)
- Yapay Sinir Ağları (Artificial Neural Networks)
- Genetik Algoritmalar (Genetic Algorithms)
- K-En Yakın Komsu (K-Nearest Neighbor)
- Bellek Temelli Nedenleme (Memory Based Reasoning)
- Naive-Bayes

- Çoklu Regresyon (Multiple Regression)
- Lojistik Regresyon (Logistic Regression)
- Faktör ve Ayırma Analizleri

3.2.1.7.2. Kümeleme Modelleri

Farklı yapılarıdaki verileri, benzer yapıda gruplara ayırma işlemidir. Kümeleme modelinde başlangıçta belirlenecek bir hedef değişkene ihtiyaç yoktur.

3.2.1.7.3. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

Bu konuyu bir örnekle açıklamak istersek. Çay alan insanların yanında neler aldığı araştırıldı. Bu bilgi için çay alan tüm müşterilerin tüm verileri alınarak ve birliktelik kuralları metodu ile yanında ne aldığıın hesabı yapılarak %30 oranında şeker satın aldıkları elde edilmiştir. Bu veri ışığında satış stratejisi geliştirilmiştir.

Benzer şekilde bugün gerçekleşen bir olayın diğer olayı da tetikleme ve bunun de veri madenciliği ardışık zamanlı örüntüler ile modellenmesi sonucu hangi olayın % kaç ihtimalle olacağıın tahmin edilmesi konuya bir örnektir.

3.2.1.8. Yöntemleri

Veri madenciliği çözümlene yapabilmek yöntemleri kullanmaktadır. Günümüzde bununla ilgili birçok yöntem vardır. Ve hâlâ yenileri bulunmaktadır. Uzun yıllardan buyana bu yana kullanılan K, kümeleme, Regresyon, en yakın komşuluk gibi yöntemlerin yanı sıra makine öğrenimi, yapay zekâ gibi, birliktelik kuralları, karar ağaçları gibi yeni keşfedilmiş yöntem ve algoritmalar veri madenciliğinde kullanılmaktadır. Bunların yanında zaman serileri ve birden çok algoritmalı bünyesinde barındıran hibrit yöntemler de bulunmaktadır.

Kdnuggets' in Mart 2007' de çevrimiçi gerçekleştirdiği bir anket sonucuna göre en yaygın kullanılan veri madenciliği yöntemleri şunlardır (Anonim):

- Karar Ağaçları, % 62,6
- Regresyon, % 51,2
- Kümeleme, % 50,2
- Tanımlayıcı İstatistik, % 46,3
- Görselleştirme, % 32,5
- Birlikte Kuralları, % 26,1
- Yapay Sinir Ağları, % 17,2
- Zaman Serisi Analizi, % 17,2

Veri madenciliğinde bulunan algoritmalarda bir sınırlama ya da mecburi standart bulunmamaktadır. Çözüme ulaştıran ve uygulanabilir olan her yöntem kullanılabilir fakat bu konuda hazır paket programlar vardır. Bu programlardan bize uygun olanı seçip kullanmak, algoritmaları tek tek elle uygulamak ya da yeni paket program üretmek gerekmektedir. Tabii ki problemimize uygun paket programı seçmek en doğrusudur.

Çalışmanın kapsamını aşmamak için burada tüm veri madenciliği yöntemlerinin açıklanmasına gerek duyulmamıştır. Bunun yerine, özellikle tahminlemeye yönelik en çok tercih edilen ve çalışmanın uygulamasında da kullanılan Regresyon, YSA ve Zaman Serileri yöntemleri ilerleyen kısımlarda ayrıntılı olarak açıklanmıştır.

3.2.1.8.2. Tahminleme

3.2.1.8.2.1. Tahmin Kavramı

Tahmin kavramı için bilinmeyen hakkında çıkarımlarda bulunmak denilebilir. Tahmin işleminde bilinen verileri inceleyerek, bilinmeyen veriler hakkında çıkarımlarda bulunmak esastır. Tahminin kelime anlamı için çoğunlukla iki farklı kullanımı bulunmaktadır. Birincisi verilerin gelecekteki değerleri hakkındadır. Bu anlamda tahmin için forecasting kelimesinin karşılığı olan anlam kastedilmektedir.

İkincisi ise verilerin ortaya çıkmış fakat bilinmediği durumlar için geçerlidir. Bu anlamda tahmin için İngilizcede estimation kelimesinin karşılığı olan anlam kastedilmektedir. Prediction kelimesi ise hem ortaya çıkmış durumlar için hem de gelecekte ortaya çıkacak durumlar için kullanılmaktadır.

Tahmin için nitel olarak uzmanların tecrübelerine, yargılarına ve görüşlerine ya da nicel olarak geçmiş verilerin genel yapısına dayanan matematik modellere başvurulabilir. Neredeyse bütün tahminler, tahmin edilecek nesnenin başka bir değişken (örneğin zaman) ya da bir değişken grubu (bir bölgedeki genel ekonomik koşullar gibi) tarafından etkilendiği varsayımına dayanır. Tahmin yöntemlerinin birçoğu on dokuzuncu yüzyılda geliştirilmiş olmasına rağmen, bu alana damgasını vuran gelişmeler bilgisayarların yaygınlaşması ile başlamıştır. Günümüzde, tahmin yöntemlerindeki kendini tekrar eden yordamlardan dolayı bilgisayarların kullanılması bir şart haline gelmiştir. Bu amaçla 90'lı yıllarda çeşitli tahmin yöntemlerini uygulamak amacıyla özel yazılım paketleri geliştirilmiştir (Deniz, 2006).

3.2.1.8.2.2. Tahmin Yöntemleri

Öncelikle yapılacak tahmin için bir yöntem ihtiyacı vardır. Bu yöntem için de bazı kriterler göz önüne alınmalıdır. Verilerin özellikleri, doğruluğu, gerekli zaman, içerdiği zaman, maliyet, kolaylık vb. kriterlerdir. Tahmin yöntemleri, tahmin edilecek süre ve tahmin çıktı sayısına göre sınıflandırılabilir. Sürelerine göre kısa, orta, uzun süreli tahminler olarak sınıflandırılabilir. Çıktı sayısına göre ise tek çıktılı ve çok çıktılı olarak sınıflandırılabilir. Genel olarak ise nicel ve nitel diye iki temel sınıfa ayrılır.

1.) Nitel Tahmin Yöntemleri: Bu tür tahmin yönteminde somut veriler tam olarak elde edilememektedir. Bunun çeşitli sebepleri vardır. Veri elde etmenin zorluğu (teknik, maliyet vb zorluklar) anlamlı verinin hiç elde edilememesi, elde edilen verilerin elde edilemeyen diğer verilerce bozulması, vb sebepler bulunmaktadır. Demir'e göre Nitel tahmin yöntemleri; esas olarak dört grupta

toplanabilir: Delphi Tekniđi, Pazar Arařtırmaları, Uzman Grup Görüřleri ve Satıř Gücü Karması. Bu yöntemler kiřisel görüř vb gibi soyut kaynaklara dayanmaktadır. Bu yüzden dođruluđunun kontrolü ve verilerin geliştirilmesi imkânına sahip olunamamaktadır. Ayrıca kiřisel eğilim ve önyargı da sonucu etkileyebilmektedir.

2) Nicel Tahmin Yöntemleri: Somut verileri kullanan tahmin yöntemidir. Çeřitli yöntemlerle elde edilen sayısal veriler ile bir model kurulur. Bu modele göre gelecek ya da bilinmeyen hakkında tahminde bulunulabilir. Veriler barkot sistemi, gözlem ve ölçüm sonuçları, çevrimiçi işlemler vb şeklinde elde edilir. Bu tahmin yöntemlerini iki grupta inceleyebiliriz. Bunlardan birincisi ilişkiyel yöntemler ki buna regresyon da denmektedir. İkincisi zaman serileridir. İliřkiyel yöntemlerde çıkıř verisinin giriř verileriyle olan ilişkiyi belirlenir. Buna göre bir model oluşturulur. Bu yöntemin zayıf yönleri ise model kurma zorluđu ve iyi bir sonuç için çok sayıda (örnek uzayın her bölgesinden veri alarak) veriye ihtiyaç duymasıdır. Zaman serilerinde ise geçmişe ait veriler düzenli ve sıralı bir veri diziliřine sahiptirler. Bu veriler analiz edilerek geleceđe ait tahminlerde bulunulur. Bu yöntem sabit şartlar altında daha iyi sonuçlar vermektedir. Bir diđer veri madenciliđi yöntemi ise yapay sinir ađları yöntemidir. Yapay sinir ađları yöntemi hem ilişkiyel hem de zaman serileri tahmin yöntemlerine uygulanabilir. Bu üç yöntem ileri bölümlerde daha ayrıntılı açıklanmıřtır.

3.2.1.8.2.3. Tahmin Modellerinin Doğruluklarının Ölçülmesi

Bir modelin dođruluđu gerçekte elde edilmiř ölçümlenmiř deđer ile modelin bulduđu deđer arasındaki farkın azlıđıdır. B iki deđer arası farka da tahmin hatası denir. Bu hataları ve dođruluđu ölçmek için çeřitli hesaplama yöntemleri vardır. Bu hesaplama yöntemlerinde verilerin bir kısmı model oluřturmaya bir kısmı oluřturulan bu modelin dođruluđunu test etmeye ayrılmaktadır. İlk önce oluřturulan model çalıştırılarak tahmini sonuçlar elde edilir. Daha sonra test için ayrılan veriler ile bu sonuçlar kıyaslanır. Kıyaslamada gerçek deđerler ile tahmin edilen deđerlerin farkları alınarak çeřitli hesaplamalar yapılarak hata ölçütleri elde edilir. Bunlar:

1) Korelasyon: Girdi ile çıktı arasında ne kadar ilişki olduđunu gösterir. Sembölü R^2 dir. Bu deđer sıfıra yaklařtıka verimizin sonuca bir etkisinin olmadıđı

anlaşılmaktadır. Bu anlamda bu veri bir işe yaramamaktadır. Bu verinin hesaba katılmasına gerek kalmamıştır. Sonuç bire yaklaştıkça verinin sonuçla tutarlı bir ilişkisinin olduğu anlaşılmaktadır. Bu iyi bir sonuçtur.

$$R^2 = 1 - \frac{\sum_{m=1}^n (t_m - y_m)^2}{\sum_{m=1}^n (t_m - t'_m)^2} \quad (3.29)$$

2) Ortalama mutlak hata: Yukarıda bahsedildiği gibi gerçek değerler ile hesaplanan değerlerin farkı alındıktan sonra mutlak değerleri ve sonuçların ortalaması bulunarak elde edilir.

$$MAE = \frac{\sum_{m=1}^n |h_m|}{n} \quad (3.30)$$

3) Ortalama mutlak yüzde hata: Ortalama mutlak hatadan farklı olarak hata yüzde olarak hesaplanır. Bunun avantajı hataların veri büyüklüklerinden bağımsız olarak görülebmesidir.

$$MAPE = \frac{\sum_{m=1}^n \frac{|h_m|}{y_m}}{n} 100(\%) \quad (3.31)$$

4) Karesel mutlak hata: bulunan her bir hata miktarının kareleri alınır. Ortalaması alınır ve karekökü alınarak elde edilir.

$$RMSE = \sqrt{\frac{\sum_{m=1}^n h_m^2}{n}} \quad (3.32)$$

Denklem 3.29-3.32 arası formüllerde;

m: toplam operatörü

n : tahmin edilen veri sayısını,

h_m :hata değerini,

t_m : hesaplanan tahmin değerini,

y_m : gerçek veri değerini

Bu tahminleri yapabilmek için çok uzun işlemler gerekmektedir. Bunları ancak bilgisayar destekli olarak gerçekleştirilebilir. Aynı şekilde test yöntemleri de dijital ortamda gerçekleştirilmektedir.

3.2.1.8.3. Regresyon Analizi

Girdi olarak kullanılan veri(ler) ile çıktı olarak seçtiğimiz veriler arasındaki matematiksel ilişkiyi keşfedip bunu bir matematiksel denkleme dönüştüren veri madenciliği modelidir. Bu modelde her girdi değerinin çıktı değerine olan katkısı hesaplanır. Bu katkılar hesaba katılarak çıkış değişkeni tahmin edilmektedir. Üç özelliğe göre sınıflandırma yapılabilir.

1) Verilerine göre:

Bu sınıflandırma şekline göre ana kütle verileriyle ve örnek verilerle analiz olmak üzere iki kısma ayrılmaktadır

2) Fonksiyonuna göre:

Bu sınıflandırma şekline göre doğrusal ve doğrusal olmayan olmak üzere iki kısma ayrılır.

3) Girdi sayılarına göre:

Bu sınıflandırma şekline göre basit ve çoklu regresyon olmak üzere iki kısma ayrılır. Basit regresyonda bir girdi varken çoklu regresyonda çok sayıda girdi vardır.

3.2.1.8.3.1. Basit Doğrusal Regresyon

Bağımsız değişken olarak adlandırılan bir girdi ile bağımlı değişken olarak adlandırılan bir çıktının matematiksel ilişkisinden oluşan modeldir. Matematiksel denklemi aşağıdaki gibidir.

$$Ç = a_0 + a_1 G + h \quad (3.33.)$$

Bu denklemdeki;

G: Bağımsız değişkeni,

Ç: Bağımlı değişkeni,

a_0 : Doğrunun $X=0$ olduğunda y eksenini kestiği noktayı,

a_1 : Doğrunun eğimini, diğer adıyla regresyon katsayısını,

h: Hata değeridir. a_1 , girdi değerindeki değişimin çıktıya nasıl yansıdığını göstermektedir

3.2.1.8.3.2. Çoklu Doğrusal Regresyon

Basit Doğrusal regresyondan farkı girdi sayısının birden fazla olmasıdır. Her bir girdi, çıktı üzerinde belli bir etkiye sahiptir. Bu etkiler tespit edilerek denklem 3.34 deki katsayılar elde edilmektedir. Bu şekilde tüm denklem de elde edilerek çıktının matematiksel ifadesi ortaya konmuş olur.

$$\text{Ç} = a_0 + a_1 G_1 + a_2 G_2 + \dots + a_n G_n + e \quad (3.34.)$$

Yukarıdaki a_0, \dots, a_n katsayıların hesaplanmasında en küçük kareler yöntemi kullanılmıştır. Bu şekilde elde edilen modelin doğruluğunu 3 yöntem kullanılmaktadır.

R:denklem 3.29.da matematiksel ifadelerinden birinin ortaya konduğu korelasyonu ifade eder. Bu ifadenin yüksekliği girdi verilerinin çıktıdaki değişimi o ölçüde açıkladığını gösterir.

R₂: bu ifade ise R ifadesindeki gibi çıktı değişkeninin ne ölçüde açıklandığını hesaplar fakat bu işlemi yüzde olarak yapar.

Düzeltilmiş R₂: bazı girdi verileri çıktı verilerini neredeyse hiç etkilememektedir fakat veri listesinde adı görülmektedir. Bu durumda R₂ ifadesi olduğundan daha iyi

bir değere sahipmiş gibi görülmektedir. Bu durumda kullanılmaması gereken veriler de hesaba katılarak yeni bir hesaplama yöntemi ortaya konmuştur. Buna düzeltilmiş R2 denir. R2 değerine göre biraz daha küçük bir değere sahip olan düzeltilmiş R2 daha gerçekçi korelasyon değerleri elde etmektedir.

F Testi: Regresyon modelinin anlamlı olup olmadığını incelemek için ANOVA (Analysis of Variance) testi uygulanmaktadır. ANOVA testi sonucunda ortaya çıkan F değerine karşılık gelen anlamlılık seviyesi, oluşturulan modelin uygun olup olmadığının kararında yardımcı olmaktadır. F test değerinin anlamlı olması (yani 0,05'ten küçük olması) söz konusu modelin bağımlı değişkeni açıklamada önemli katkı sağladığına işaret etmektedir (Asilkan Ö,2008).

3.2.1.8.4. Zaman Serileri Analizi

Belli sayıda değişken için belirli zaman aralıklarında yapılan ölçümlerini veri olarak kabul eden ve buna göre model ve çıktı elde eden veri madenciliği yöntemidir. Genellikle geleceğin tahmini yapılmaktadır. Zaman serilerini etkileyen ekonomik, demografik, psikolojik, sosyal vb pek çok faktör vardır. Bu faktörlerin etkisiyle dalgalanmalar olmaktadır. Bu bozucu etkileri aşağıdaki dört grupta toplanabilmektedir.

- Trend (Uzun dönemli eğilim) (T)
- Mevsimsel dalgalanma (S)
- Konjonktürel (Devirsel) dalgalanmalar (C)
- Tesadüfi dalgalanmalar (I)

Üstte anılan unsurlara sahip bir zaman serisi genel olarak alttaki denklemlerle gösterilebilir: (Gürtan, 1977):

$$Y = T + C + S + I \quad (3.35.)$$

Yazında “rastsal terim” ya da “rastsal gürültü” olarak da ifade edilen tesadüfi dalgalanmalar, I sembolü dışında ϵ sembolü ile de gösterilmektedir.

- Trend; bir zaman serisinin uzun dönemde belirli bir yöne doğru gösterdiği gelişmedir. Trendler doğrusal veya eğrisel olabilir, önemli olan istikrarlı olmalarıdır.
 - Mevsimsel dalgalanma; Serinin bazı dönemlerin etkisiyle kısa zaman aralıklarında kendini tekrarlama eğilimi göstermesidir. Bu dönemler, bir yıldan uzun süreli olamaz. Örneğin tarım ürünlerinin fiyatlarının ya da turist sayılarının mevsime bağlı olarak değişmesi ve her yılın aynı mevsimlerinde benzer iniş-çıkışlar göstermesi gibi.
 - Konjonktürel dalgalanmalar; Serinin 3-5 yıllık orta zaman aralıklarında kendini tekrarlama eğilimi göstermesidir. Örneğin; üretim, satışlar, gelirler gibi iktisadi olaylarda önce bir süre gelişme görülür, bu yükselmenin maksimuma çıktığı anda bir iktisadi kriz ortaya çıkar ve düşüş başlar, bu düşüşten bir müddet sonra tekrar yukarı doğru hareketlenme başlar ve aynı olaylar benzer şekilde sürekli tekrarlanır. Bu hareketler mevsimsel dalgalanmalara benzer şekilde devri olarak tekrar etseler de, devreler daha uzun, süreler de belirsizdir.
 - Tesadüfi dalgalanmalar; Serinin, üstte sayılan dalgalanmaların dışında, belirli bir sistematik yapıya sahip olmaksızın rastgele değişim gösteren kısmını ifade etmektedir.
- “Oto regresif süreç” olarak adlandırılan kargaşa, savaş, deprem, sel, grev gibi bu hareketlerin ne zaman nasıl tekrarlanacağı belirsizdir.

3.2.1.8.4.1. Zaman Serilerinin Temel Özellikleri

İç Bağımlılık: Zaman serilerinde veriler içerisinde temel değişiklik zaman verilerindedir. Zamandaki değişime göre çıktıdaki değişim bulunmaya çalışılır. Zamanda birbirine yakın ölçümler birbirine yakın sonuçlar doğururken süre arttıkça sonuçlar birbirinden uzaklaşmaktadır. Girdi verilerinin bağımsız olduğu kabulü ile çalışan modeller bu seriler için geçerliliğini veya başarısını yitirmektedir. Burada veriler arasındaki bağlantıları hesaba katan modeller iyi sonuçlar verecektir.

3.2.1.8.4.2. Stokastik (Tesadüfi) Süreç Olması

Matematiksel ifadeler ile tam olarak bulunabilen çıktılar için deterministik ifadesi kullanılmaktadır. Gelecekle ilgili ifadelerde ise kesin sonuç bulmak mümkün değildir çünkü çıktıyı etkileyen tüm girdilerin tahmin edilmesi mümkün olamamaktadır. Ayrıca tüm girdiler tespit edilse bile hepsinin doğru verilerinin tam olarak elde edilebilmesi pek mümkün değildir. Onun yerine ‘çan eğrisi’ örneğinde olduğu gibi olasılık dağılım fonksiyonu ile çıktı tahmin edilmeye çalışılmaktadır. Bu yüzden bu bir stokastik süreçtir.

3.2.1.8.4.3. Zaman Serilerinin Sınıflandırılması

Gürsakaal’a göre Zaman serilerini; değişken sayısına göre; “tek veya çok değişkenli”, elde edilme biçimine göre; “sürekli veya kesikli”, ortalamadan gösterdiği sapmalara göre; “durağan veya durağan olmayan”, devirsel hareketlenmelerine göre; “mevsimsel veya mevsimsel olmayan” şeklinde sınıflandırmak mümkündür. Tek veya çok değişkenli zaman serileri: Zaman içindeki hareketi izlenen olay sayısı bir olan seriler “tek değişkenli”, birden fazla olan seriler ise “çok değişkenli” olarak adlandırılır. Sürekli ve Kesikli Zaman Serileri: Şayet incelenen olaya ait gözlem değerleri zaman içinde sürekli olarak elde ediliyorsa, bu serilere “sürekli zaman serisi” denilmektedir. Bu tür seriler genellikle zaman içinde eşit olmayan aralıklarla elde edilen gözlem değerlerinden oluşur. Şayet gözlemler sadece belirli zaman aralıklarıyla yapılıyorsa, bu tip serilere “kesikli zaman serileri” denir. Kesikli zaman serileri genellikle eşit ve birbirini takip eden zaman aralıklarıyla yapılan gözlem değerlerinden oluşur. Özellikle ekonomik büyüklükleri gösteren seriler, belirli zaman aralıklarında ölçüldüğü için, kesikli zaman serileri sınıflaması altında incelenmektedir.

3.2.1.8.4.4. Analiz Amaçları ve Uygulama Alanları

Zaman serilerinin birçok amaçla kullanılma imkânı vardır. Genel olarak tahmin, kontrol, açıklama, belirleme amaçlarında kullanılmaktadır. bu amaçla birçok alanda (faiz miktarları, fiyatlar, otomobil kullanım, kaza oranları vb.) çok geniş alanda kullanılmaktadır. Bu seriler dört aşamada uygulanmaktadır.

- 1) Tanımlama: bu aşamada veriler zaman eksenine yerleştirilerek özellikleri ortaya konur. Homojen varyans, normal dağılım sağlanır. Mevsimsel etkiler azaltılır.
- 2) Açıklama: iyi bir model kurulur.
- 3) Kurulan modele göre tahmin yapılır.
- 4) Tahmin sonuçları kontrol edilir. Yetersiz ise farklı modeller ile tekrarlama yapılır.

3.2.1.8.4.5. Zaman Serilerinde İleriye Dönük Tahmin Yöntemleri

Zaman serilerinde ileriye dönük tahmin amaçlı kullanılan yöntemler, serinin tek ya da çok değişkenli olmasına göre farklılık gösterebilmektedir. Çok değişkenli zaman serilerinde genelde dinamik regresyon modelleri ve çok değişkenli zaman serileri analizi gibi yöntemler kullanılmaktadır. Çok değişkenli zaman serileri analizi, tahmin sistemiyle ilgili her şeyin bilindiğini varsayımını dikkate almaktadır, bu amaçla çeşitli simülasyon ya da istatistik yöntemleriyle bilinmeyen değişken değerlerinin yapay olarak üretilmesi yoluna gidilir. Bu şekilde hesaplanan verileri kullanarak yapılan tahminler gerçekçi sonuçlar vermekten uzaklaşacağı için güvenilirlikleri düşük olmaktadır. Bu nedenle uygulamada çok değişkenli zaman serilerinin tahmini çok tercih edilmemekte; bunun yerine yaygın olarak tek değişkenli zaman serileri kullanılmaktadır. Bu çalışmada da tek değişkenli zaman serileri kullanıldığı için, ilerleyen kısımlarda sadece tek değişkenli zaman serilerine yönelik tahmin yöntemlerine yer verilmiştir. Tek değişkenli zaman serileriyle ilgili tahmin yöntemleri, mevcut olan zaman serisi unsurlarının gelecek dönemde de aynı kalacağı varsayımına dayanmaktadır. Bu varsayım nedeniyle geçmiş dönem gözlem değerleri kullanılarak gelecek dönem tahmin değerlerinin elde edilmesi amaçlanır (Asilkan, 2008).

Geleceğe ait verilerin tahmin edilirken birçok yöntemden hangisinin daha iyi sonuç vereceği, girdi veri miktarına, dönem genişliğine, zaman serisi unsurlarına, tipine ve amaca göre değişmektedir. Zaman serileri tahmin yöntemleri:

1) Trend analizi: çok eski bir yöntemdir. Serpilme diyagramındaki verilere uygun bir denklem elde edilir. Bu denklem ışığında geleceğe ait tahminler yapılabilmektedir. Bu yöntemin anlaşılması, hesaplanması, kolay ve maliyeti düşüktür. Bunların yanında tek girdi ile çalışması, çıktının doğruluğunun kontrol edilememesi, veri güncelleme problemleri, çok eski verileri de istemesi, mevsimselliği planlama, gibi olumsuz yönleri de bulunmaktadır.

2) Mekanik tahmin: Bu tür tahminlerde yöntem bir önceki zaman periyodundaki değer geleceğin tahmininde kullanılmasıdır. Bu haliyle en kaba yöntemdir. Diğer yöntemlerle yapılmış tahminlerin doğruluğunu ölçmede de kullanılabilir fakat çok başarılı bir tahmin aracı değildir.

3) Hareketli Ortalamalar: Zaman serisindeki veriler büyüklüklerine göre belli gruplara ayrılır. Her grup için ayrı ayrı ortalama değer hesaplanır. Bulunan değer yeni gelecek veri tahmini olmaktadır. Yeni verilerin (örneğin yeni dönemin girmesi ile) gruplara eklenmesi durumunda direkt olarak ortalamaları alınarak güncelleme gerçekleşir ve yeni veriler elde edilir. Verilerdeki bu hareket sebebiyle modele hareketli ortalamalar denmektedir. Elde edilen verilerdeki rastlantısallığın yüksek ve veriler arası korelasyonun düşük olduğu durumlarda tercih edilebilir. Dezavantajları ise çok sayıda ölçüm vb verisine ihtiyaç olması, bazı durumlarda adaptasyon zayıflığı, geçmiş hareketsiz dönemleri hesaba katamaması, mevsimsel değişiklikleri hesaba katamamasıdır. Hareketli ortalamalar yöntemi basit, ağırlıklı, çift hareketli ortalama olmak üzere üç çeşittir.

4) Üstel Düzleştirme: Bu yöntemde girdi olarak kabul edilen verilerin ortalamaları alınarak çıktı elde edilmektedir. Hareketli ortalamalar yönteminden farkı ise zaman olarak yakın verilerin ağırlıklarının uzak verilere göre üstel olarak artmasıdır. 1958 yılında holt tarafından bulunan bu yöntem mevsim ve trend değerleri içermemekteydi. Ardından Brown yöntemi uygulanabilir kılımtır. Winter ise

mevsimselliği de yönteme eklemiştir. Bu yöntem ucuz, hızlı, güncellenebilir ve mevsimseldir.

Aşağıda yazında en çok adı geçen belli başlı üstel düzleştirme yöntemleri görülmektedir(Asilkan Ö, 2008).

- a) Brown'un Basit Üstel Düzleştirme Yöntemi
- b) Brown'un Tek Parametrelili Doğrusal Üstel Düzleştirme Yöntemi
- c) Holt'un Çift Parametrelili Doğrusal Üstel Düzleştirme Yöntemi
- d) Winter'ın Mevsimsel Üstel Düzleştirme Yöntemi
- e) Chow'un Uyarlanabilir Kontrol Yöntemi
- f) Harrison'un Harmonik Düzleştirme Yöntemi
- g) Uyarlanabilir Tepki Oranlı Basit Üstel Düzleştirme Yöntemi
- h) Brown'un Tek Parametrelili Uyarlanabilir Yöntemi

5) Uyarlayıcı Arındırma: Makridakis ve Whellwright tarafından geliştirilmiştir. 1970 li yıllarda ortaya çıkmıştır. Geçmişte ölçümlenen girdi değerlerinin toplamı alınarak oluşturulan modelle çıktı değerleri tahmin edilir. Bu yönüyle hareketli ortalamalar ve üstel yöntemlere benzemektedir. Fakat sonuçların doğruluğu iki yöntemden de daha iyidir. Kullanıcının müdahalesi çok azdır. İşlemler daha hızlıdır. Girdi verileri ile bir model oluşturularak çıktı elde edildikten sonra girdilerdeki değişikliklerden dolayı yeni bir model oluşturmadan kolay bir şekilde yeni çıktı bulunabilmektedir.

6) Box-Jenkins Yöntemleri: Olayların stokastik olduğu ve rastsal karakterde gerçekleştiği varsayımı ile yöntem oluşturulmaktadır. Bu yöntem iç bağımlılığı iyi bir şekilde hesaba katmaktadır. Bu yüzden doğrusal stokastik model de denmektedir. Bu model mevsimsel ve durağan olmayan Zaman serilerinde kolaylıkla kullanılabilir. Mevsimsel zaman serisi ve durağan olmayan zaman serisi zaman ve çalışma açılarındaki uygun modellerdir.

George Box ve Gwilym Jenkins "Time Series Analysis: Forecasting and Control" adlı kitabında ARİMA modellemesi anlatılmaktadır. Durağan olan ya da olduğu kabul edilen serilerde kullanılmaktadır. ARİMA modeli, modelin seçimi, parametrelerin bulunması, modelin uygunluğuna karar verilmesi olarak üç bölümde gerçekleşir. Modelin notasyonu ARIMA(p,d,q) şeklindedir. p otoregresiflik

derecesini, q hareketli ortalama derecesini, d ise durağanlaştırılabilme derecesini temsil etmektedir. Karar verebilmek için kriterlerde en iyi sonuç veren model seçilir. En iyi sonuç veren modelin belirlenmesinde zaman serisi otokorelasyon katsayısı gecikme uzunlukları ile standart $AR(p)$, $MA(q)$ ve $ARMA(p,q)$ modellerinin kıyaslama yöntemi kullanılır. Parametrelerin bulunmasında iki temel seçenek vardır. Maksimum olasılıklar ve nonlinear en küçük kareler yöntemi. Yöntem belirlemek çalışmacı kişilere kalmaktadır. Bu yüzden tam objektif olunamamaktadır.. Fakat bu durum Akaike Information Criterion (AIC) veya Schwarz Information Criterion (SIC) gibi kriterlerle giderilmeye çalışılmaktadır.

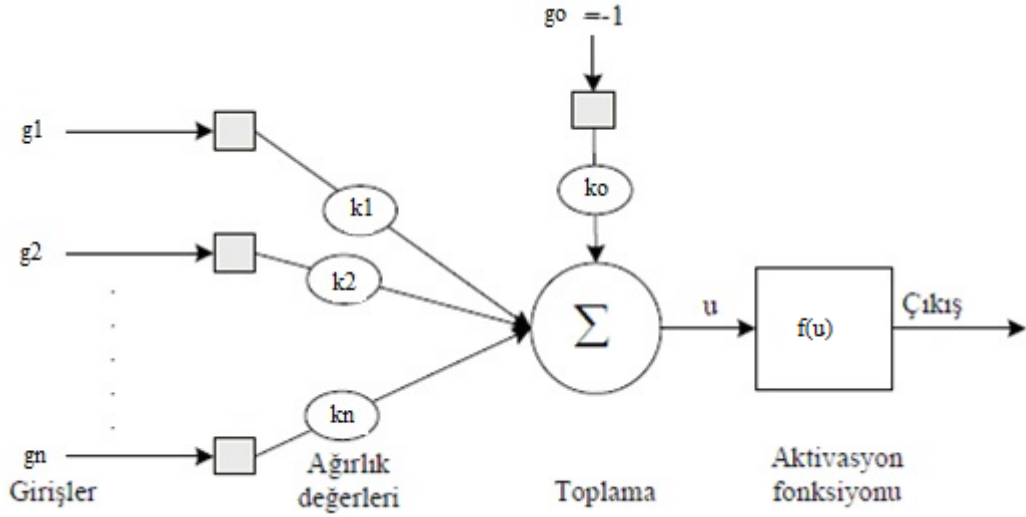
Box-Jenkins yöntemleri durağan olup olmamasına göre ikiye ayrılır. Durağan olmayan yöntemler (ARIMA) de mevsimsel olup olmamasına göre ikiye ayrılır.

3.2.1.8.5. Yapay Sinir Ağları

İnsan çok büyük miktarlardaki bilgileri çok kısa zamanda ve yeterince doğru bir şekilde algılamakta ve çözümlenmektedir. Bunu nöron adı verilen sinir hücreleri ile başarmaktadır. Hücrelerin yapısı çekirdeği olduğu gibi veri alışında görevli dentritleri ve veri üretilmesinde görevli akson tepeciği, akson, sonlandırıcı düğümlerden oluşmaktadır. Bir insanda ortalama 10 milyar beyin hücresi bulunmaktadır. Paralel çalışan bir sistemdir. Bu sistem örnek alınarak yapay sinir ağları modeli ortaya konmuştur. Yapay sinir ağlarının klasik yöntemlerden farkı eğitilmesi, öğrenmesi ve istenen görevi yerine getirebilmesidir. Yapay sinir ağlarının temel özellikleri Nonlinear sistemlere uygulanabilmesi, genelleme yapılabilmesi sınıflandırma özelliği, hızlı çalışma vb dir. Keşfinden bu yana Hopfield ağları, Kohonen ağları, adaptif rezonans teorisi ve ileri beslemeli yapay sinir ağları gibi yöntemler ortaya atılmıştır.

3.2.1.8.5.1. Yapay sinir ağı matematiksel modeli

Yapay sinir hücresi en küçük işlem birimidir. çok sayıda dentrit ile çok sayıda veri alırken akson ile tek çıkış üreten bir sistem olarak gösterilebilir. Girdiler kümesi ile veriler çevreden alınır. Sinir hücresinde işlenir. Buradaki işleme aktivasyon fonksiyonuna tabi tutulmaktadır. Bağlantılar vasıtasıyla veriler diğer hücrelere aktarılır. Bu esnada belli bir ağırlık ile çarpılmaktadır.



Şekil 3. 11. yapay sinir ağı matematiksel açılımı

Yukarıda blok diyagram gösterimi verilen yapay sinir ağı matematiksel ifadeleri ise:

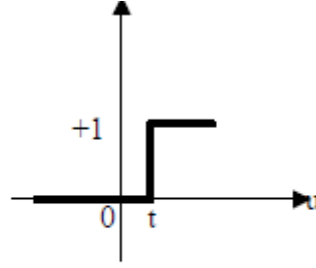
$$u(t) = k_1(t)g_1(t) + k_2(t)g_2(t) + \dots + k_n(t)g_n(t) - k_0 \quad 3.36.$$

$$y(t) = f(u(t)) = f\left(\sum_{i=1}^n (k_i g_i(t) - k_0)\right) \quad 3.37.$$

g_1, g_2, \dots, g_n girdileri k_1, k_2, \dots, k_n katsayıları ile çarpılarak toplanır ve $f(u)$ fonksiyonuna tâbi tutulur. Burada f aktivasyon fonksiyonudur. Sürekli azalmayan ve türevi alınabilen bir fonksiyon olmalıdır.

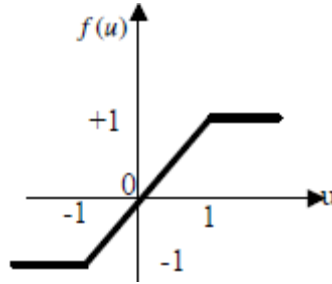
3.2.1.8.5.2. Aktivasyon fonksiyonları

Şekil 3.11.de gösterilen $f(u(t))$ fonksiyonuna, aktivasyon fonksiyonu denir. $f(u)$ olarak da anılmaktadır. Aktivasyon fonksiyonlarına örnekler aşağıda sunulmuştur:



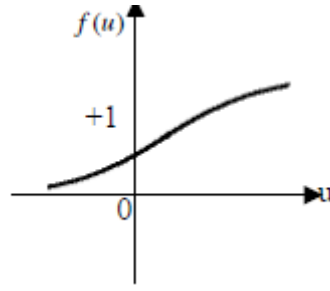
$$\dots\dots\dots f(u) = \begin{cases} 1, u \geq t \\ 0, u < t \end{cases} \dots\dots\dots 3.38$$

Şekil 3.12. Basamak fonksiyonu ve matematiksel ifadesi



$$f(u) = \begin{cases} 1 \dots\dots u \geq 1 \\ u \dots -1 < u < 1 \\ -1 \dots\dots u < -1 \end{cases} 3.39$$

Şekil 3.13. Eşik fonksiyonu ve matematiksel ifadesi



$$f(u) = \frac{1}{1 + e^{-au}} \quad 3.40$$

Şekil 3 14. Sigmoid fonksiyonu ve matematiksel ifadesi

Daha önce karşılaşılan problemlerin çözümünde yapay sinir ağları daha başarılıdır. Çünkü ağırlık ve katsayılar daha iyi seçilebilmektedir. Fakat daha önce karşılaşılmamış problem için farklı ağırlık ve katsayılar denenerek ya da aktivasyon fonksiyonları değiştirilerek daha iyi sonuçlar aranabilir. Gerçek sinir hücrelerine en çok benzeyen aktivasyon fonksiyonu sigmoid aktivasyon fonksiyonudur. En çok tercih edilen fonksiyonlardan biridir.

Giriş verileri ile çıkış verileri arasında doğrusallık olmayan ağlar için çok katmanlı yapay sinir ağları uygundur. Çok katmanlı ağlar bir giriş katmanı bir çıkış katmanı ve aradaki gizli katman(lar)dan oluşur. Bir ağ yapısında verinin iki farklı akış yönü vardır. Bu özelliklerine sınıflandırma yapılacak olursa, ileri beslemeli yapay sinir ağları ve geri beslemeli yapay sinir ağları olarak ikiye ayrılır:

3.2.1.8.5.3. İleri beslemeli yapay sinir ağları

ilk katmandan girip gizli katmanlarda devam edip çıkış katmanına doğru ilerleme yönüne sahipse tek yönde ilerleme olmaktadır. Bu ağ yapısına ileri beslemeli yapay sinir ağı adı verilmiştir. Genel anlamda yapay sinir ağlarında olduğu gibi ileri beslemeli yapay sinir ağında da girişteki veriler doğrudan ilk düğümlere aktarılır. Bu düğümde katsayılar ile çarpılarak toplanır. Aktivasyon fonksiyonundan geçirilerek, bir sonraki katmana iletilir. Aynı katmandaki veriler arasında bağlantı yoktur. Bu şekilde çıktı katmanına kadar devam edilir. Daha sonra yeni katsayılar belirlenerek

ağın öğrenme işlemi gerçekleştirilir. Bu yapıyla lineer olmayan fonksiyonlar da çözülebilir. En iyi sonuç için kaç katman kullanılacağına dair kesin bir çözüm yoktur. Önceki çalışmalardan, tahminlerden ve deneme yanılma yönteminden faydalanarak seçim yapılır.

3.2.1.8.5.4. Geri beslemeli yapay sinir ağları

Yaygın olarak kullanılan bir yöntemdir. Aktivasyon fonksiyonu olarak sigmoid kullanılır. Rastgele ağırlıklar seçilerek başlanır. Seçilen başlangıç değerleri ağın öğrenme hızını etkilemektedir. Gizli katman sayısı ve her katmandaki nöron sayısının çok olması ağı ezbere yöneltir. Fakat çok sayıda nöron kullanmak gerekmektedir. Bilgi ve görüntüler arasındaki bağıntılar kullanılarak depolama yapılırsa daha az nöronla daha fazla depolama gerçekleşir. (Aydın, 2006)

Rastgele ağırlıklar seçildikten sonra çıktı değerine göre geri yayılım algoritması oluşturulur. Buna göre yeni ağırlıklar elde edilir. bu algoritma:

1) Giriş katmanı dışındaki katmanlara I girişi uygulanır.

$$net_p = \sum_{r=0}^n k_{pr} g_r \quad 3.41.$$

Denklem 3.41. deki g önceki nöron çıkışı, k ağırlıklardır.

2) Giriş katmanı dışındaki katmanlarda nöron çıkışı $y=f(t)$ dir. $f(t)$ ise:

$$f(t) = \frac{1}{1 + e^{-t}} \quad 3.42.$$

olarak kabul edilir.

3) Ağırlıkların güncellenmesi ise belli bir kurala tabidir. Delta kuralı

$$\Delta k_{pr}^{q+1} = \eta h f(t) + \alpha \Delta k_{pr}^q \quad 3.43$$

h , hata α , momentum katsayısı η öğrenme katsayısıdır.

4) İstenen çıkış değerinden ağırlık elde ettiği çıkışı çıkararak hata bulunur.

5) Elde edilen hata önce çıkış katmanına ve oradan geriye doğru girdi katmanına kadar denklem 3.43'e göre yayılım yapılarak hata değerleri ve yeni ağırlıklar hesaplanır. İkinci adıma dönülür. En iyi katsayılar bulununcaya kadar bu işlemlere devam edilir.

4.ARAŞTIRMA BULGULARI VE TARTIŞMA

Çalışmamızın bu bölümünde daha önce bildiğimiz matematiksel ifadeler ve elde ettiğimiz datalar ile veri madencilik uygulanmıştır. Bu uygulama için hem hızlı olması hem de birçok uygulamayı bünyesinde barındırması sebebiyle weka tercih edilmiştir. Üç başlık altında dört geometrik yapı analiz edilmiştir. En iyi sonucu veren algoritmalar tespit edilmiştir. Bu algoritmaların eğitim, test veri oranları değiştirilerek en iyi sonuç elde edilmiştir. Bu sonuçlara göre yorumlar yapılmıştır.

4.1 Trapez ve Dikdörtgen Şekilli Mikroekranlı Hatta Veri madenciliği uygulaması

İlk önce Trapez ve dikdörtgen şekilli mikroekranlı hatta veri madenciliği uygulaması yapılmıştır. Bu işlem iki farklı test veri oranı için tekrarlanmıştır. Birinci uygulama için %66 eğitim ve %34 test veri oranı seçilmiştir, ikinci uygulama ise %80 eğitim, %20 test veri oranı için yapılmıştır.

Çizelge 4.1. TSML de %66 eğitim, %34 test verileri ile yapılan uygulama sonucu elde edilen korelasyon ve hata

Çıkış alan adı	Zo		
	Korelasyon katsayısı	MAE	RMSE
Algoritma adı			
IBk (instance-based)	0.986	2.59	3.44
KStar	0.988	5.166	6.280
Additive Regression	0.962	4.389	5.265
Bagging	0.994	1.463	2.087
Decision Table:	0.984	2.278	3.398
M5 rules	0,993	1.634	2.235
Trees M5P	0,994	1.454	2.013
REPTree	0.999	1,936	2.680

Çizelge 4.1. ye genel hatları ile bakıldığında, korelasyon katsayı değerleri 0.99 civarında çıkmaktadır. Bu değer giriş değişkenleri ile çıkış değerleri arası ilişkinin çok iyi seviyede olduğunu göstermektedir. Bu durumda hataların da çok düşük değerlerde olduğu anlaşılmaktadır. Normalde bire yakın olması iyidir.

Hata deęerleri (MAE) 1 ve (RMSE) 2 civarındadır. Bu da 100 Ω civarı empedanslar için normal bir deęer. Daha sonra ikinci uygulama eęitim ve test veri oranları %80 - %20 için yapılmıřtır..

Çizelge 4.2. TSML de ikinci uygulama. %80 eęitim, %20 test verileri için hesaplanan korelasyon ve hata

Çıkıř alan adı	Zo		
	Korelasyon katsayısı	MAE	RMSE
Isotonic regression	0,920	6,055	7,288
Least medswq	0.950	4.531	6.672
Linear Regression	0,978	3,037	3,820
Multilayer Perceptron	0,991	1,712	2,42
Pace Regression	0,978	3,036	3,819
Simple Linear regression	0.920	6,038	7,272
SMOreg	0,978	3,261	4,093
IBk (instance-based)	0,989	2,070	2,692
KStar	0,989	4,244	5,266
LWL	0,851	8,149	9,716
Additive Regression	0972	3,668	4,322
Bagging	0,995	1,434	1,894
Ensemble selection	0,992	1,758	2,284
Conjunctive rule	0,815	8,832	10,699
Decision Table:	0,980	2,811	3,787
M5 rules	0.994	1,518	1,994
Zeros	0,994	1,518	1,994
Trees M5P	0.994	1,485	1,968

Çizelge 4.2. ışığında MAE ve RMSE değerleri hesaba katılarak en iyi sonucu veren algoritma seçilecektir. Burada MAE nin en düşük olduğu değerler en iyiden en kötüye doğru sıralarsak.

1)Trees M5P (1,454)

2)Bagging (1,463)

3)M5Rules (1,634)

Burada RMSE nin en düşük olduğu değerler en iyiden en kötüye doğru sıralarsak

1)Trees M5P (2,013)

2)Bagging (2.087)

3)REPTree (2.235)

Bu algoritmaların korelasyon değerleri ise

1)IBk (instance-based) (0,994)

2)Least medswq (0,994)

3)SMOreg (0,993)

Hata değeri hem RMSE hem de MAE olarak en düşük olan Bagging modelidir (MAE 1,434 ve RMSE 1,894) bu modelden sonra en küçük olan TressM5P (MAE 1,518 ve RMSE 1,994))dir. Daha sonra ise Zeros M5Rules (MAE 1,518 ve RMSE 1,994)) algoritmaları sonuçları alınmıştır.

Algoritmaların korelasyon katsayılarına bakalım. Bagging için(0,995) ve M5P (0,994)dir. Giriş çıkış ilişkisi için ikisi de iyi bir değer. IBk (instance-based) bizim için en iyi algoritma olmaktadır. Çünkü hata değerleri en düşük ve uygun korelasyona sahip olan algoritmadır.

Dikdörtgen ve trapez şekilli mikroekranlı hat için en iyi algoritmaM5P dir

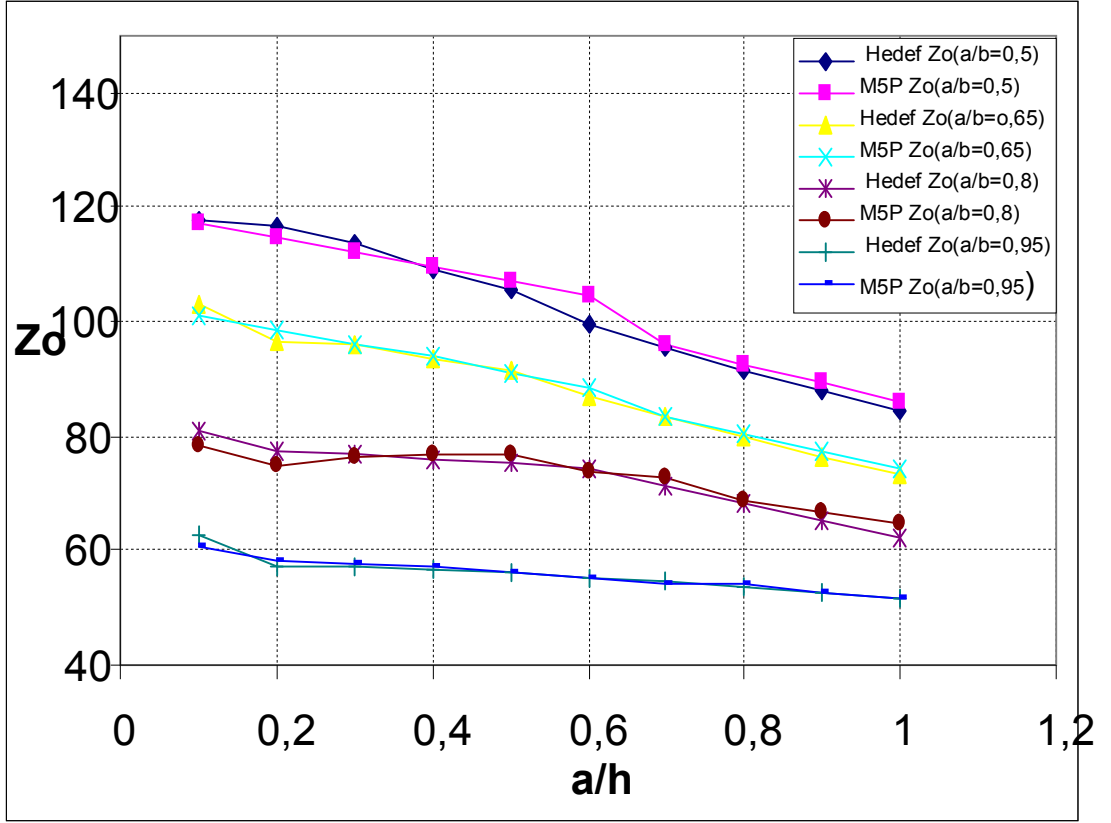
Dikdörtgen ve trapez şekilli mikroekranlı hat için a/h, a/b, D/W, β değerleri girdi olarak kabul edilirken Zo değeri çıktı olarak kabul edilmektedir. Girdiler için:

a/h=0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 birimsiz

a/b=0.5, 0.65, 0.8, 0,95 birimsiz

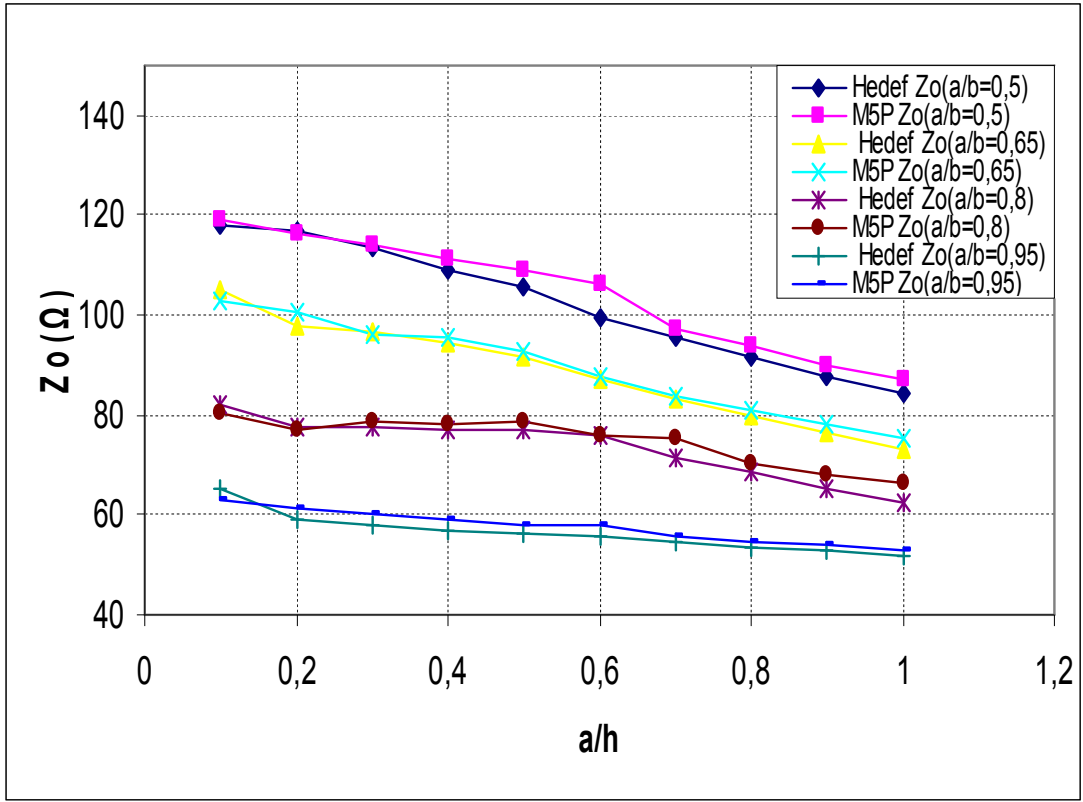
D/W=0.5,0.65, 0.8, 1 birimsiz

$\beta = -20, 0, 20$ derece Bu girdilerden β nin üç farklı değeri için üç ayrı grafik hazırlanmıştır.



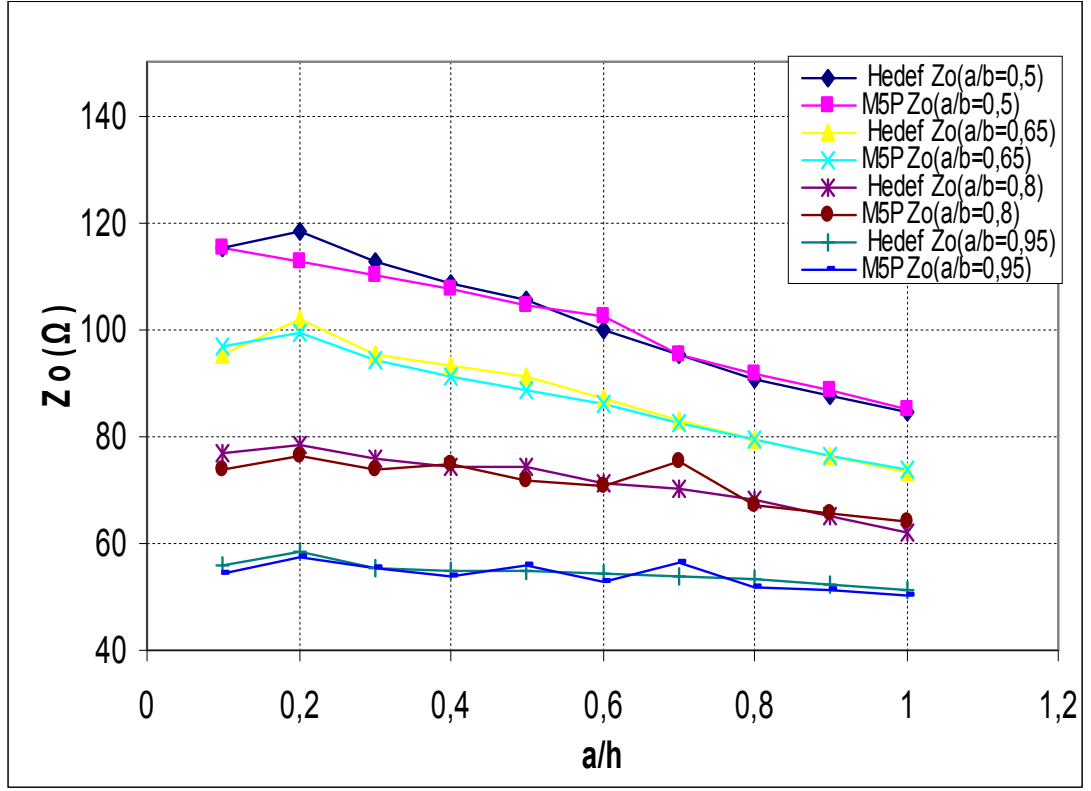
Şekil 4.1 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo. D/W=0,5, beta=0 alınmıştır. Dikdörtgen şekilli mikroekranlı hatta dönüşür.

Beta =0 alındığı için geometri dikdörtgen şekilli mikroekranlı hatta dönüşür. a/b 0,95 den 0,5 düşmesiyle çıkış empedansı da 61 Ω den 119 Ω a çıkmaktadır. Aynı şekilde a/h değerleri 0,1den başlayarak 1 değerine kadar artırılması çıkış empedansı da yine a/b=0,95 için 61 Ω en 41 Ω e ve a/b=0,5 için 119 Ω dan 84 Ω e kadar düşmesine neden olmaktadır.



Şekil 4.2 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo . D/W=0,5, beta=-20 alınmıştır. Negatif açılı trapez şekle dönüşür.

Beta =-20 alındığı için geometri negatif açılı trapez şekilli mikroekranlı hatta dönüşür. a/b 0,95 den 0,5 düşmesiyle çıkış empedansı da 62 Ω den 118 Ω a çıkmaktadır. Aynı şekilde a/h değerleri 0,1den başlayarak 1 değerine kadar artırılması çıkış empedansı da a/b=0,95 için 84 Ω en 52 Ω e ve a/b=0,5 için 119 Ω dan 62 Ω e kadar düşmesine neden olmaktadır.



Şekil 4.3 a/h ve a/b nin farklı değerlerine göre hedef Zo ve M5P algoritması çıkış Zo . D/W=0,5, $\beta=20$ alınmıştır. Pozitif açılı trapez şekle dönüşür.

Beta =20 alındığı için geometri pozitif açılı trapez şekilli mikroekranlı hatta dönüşür. a/b, 0,95 den 0,5 düşmesiyle çıkış empedansı da 54 Ω den 115 Ω a çıkmaktadır. Aynı şekilde a/h değerleri 0,1den başlayarak 1 değerine kadar arttırılması çıkış empedansı da a/b=0,95 için 84 Ω en 50 Ω e ve a/b=0,5 için 115 Ω dan 54 Ω e kadar düşmesine neden olmaktadır.

Bu geometrik şekiller için veri madenciliği performans ölçütü olarak interpolasyon ve ekstrapolasyon yapılmıştır. ekstrapolasyonda:

Çizelge 4.3. Trapez şekilli hat M5P algoritması için ekstrapolasyon test sonuçları

Test verileri	Korelasyon	MAE	RMSE
a/h=0,1	0,997	0,953	1,326
a/h=1	0,997	2,630	2,885
a/b=0,5	0,995	14,92	15,2
a/b=0,95	0,996	0,968	1,357
D/W=0,5	0,998	0,813	1,215
D/W=1	0,998	0,813	1,184
B=-20	0,993	2,129	3,046
B=20	0,995	1,390	1,925

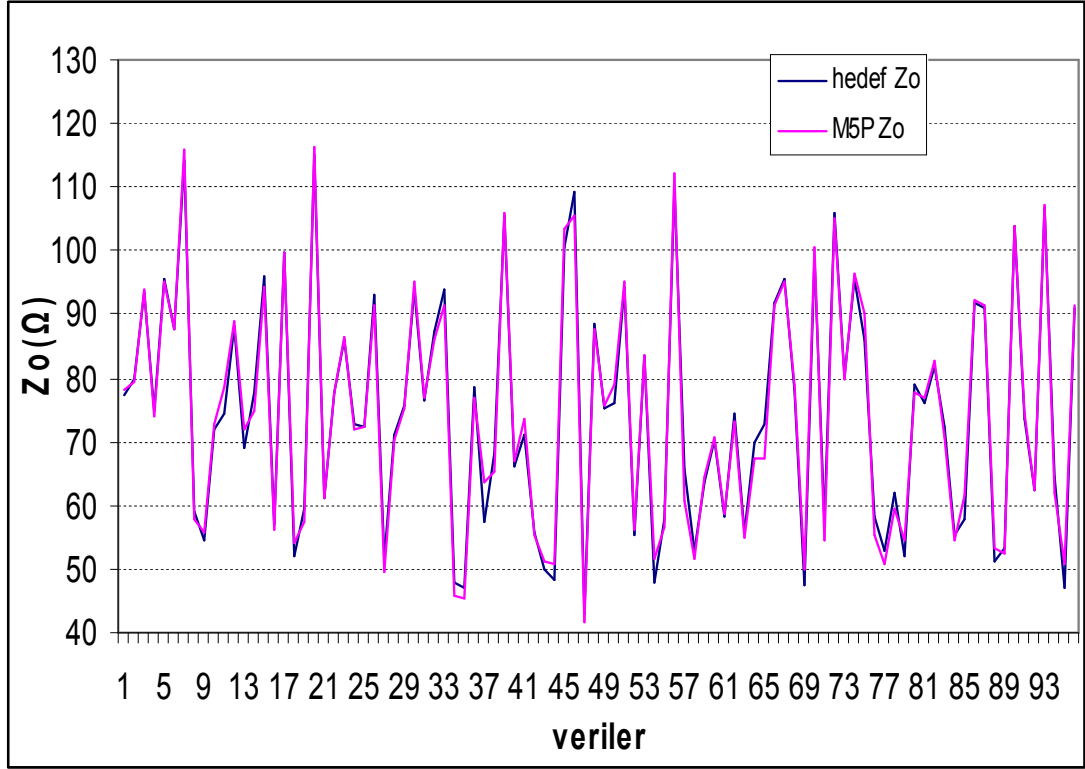
Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. Ekstrapolasyon test sonuçlarına göre trapez şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,997 mertebesindedir. Bu sonuçlar girdi çıktı arası ilişkinin mükemmel ve hatanın sıfıra çok yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. Çok iyi değerler sağlanmıştır. interpolasyon bakılırsa:

Çizelge 4.4. Trapez şekilli hat M5P algoritması için interpolasyon test sonuçları

Test verileri	Korelasyon	MAE	RMSE
a/h≠0,1 ve a/h≠1	0,991	1,690	2,571
a/b≠0,5 ve a/b≠0,95	0,998	0,968	1,357
D/W≠0,5 ve D/W≠1	0,981	3,091	4,098
b=0	0,995	1,390	1,925

Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. İnterpolasyon test sonuçlarına göre trapez şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,995 mertebesindedir. Bu sonuçlar girdi çıktı arası ilişkinin mükemmel ve

hatanın sifira çok yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. Çok iyi değerler sağlanmıştır. Son olarak hedef ve hata değerleri aynı grafikte çizilmiştir. Böylece hatanın seviyesi net olarak görülmüştür.



Şekil 4.4.hedef Zo ile M5P algoritmasının bulduğu Zo

Şekil 4.4 deki veriler M5P algoritmasının test işleminde kullandığı verilerdir. Açıkça görülmektedir ki hedef Zo ile bulunan Zo (M5P) birbirlerine çok yakındır. Hatanın seviyesi burada da açıkça görülmektedir.

4.2. V-şekilli Mikroekranlı Hatta Veri Madenciliği Uygulaması

İlk önce V-Şekilli mikroekranlı hat için weka çalıştırılmıştır. Bu işlem iki farklı test veri oranı için yapılmıştır. Biricisi %66 eğitim ve %34 test veri oranıdır. ikincisi %80 eğitim, %20 test veri oranı için test yapılmıştır.

Çizelge 4.5. V-Şekilli Mikroekranlı hat birinci uygulama. % 66 eğitim, %34 test verileri için hesaplanan korelasyon ve hata

Çıkış alan adı	Zo		
	Korelasyon katsayısı	MAE	RMSE
Isotonic regression	0.828	2.529	2.990
Least medswq	0.940	1.406	1.809
Linear Regression	0.945	1.345	1.736
Multilayer Perceptron	0.971	0.997	1.311
Pace Regression	0.945	1.345	1.736
Radial basis function network	0.300	4.189	5.096
Simple Linear regression	0.800	2.653	3.196
SMOreg	0.945	1.338	1.749
IBk (instance-based)	0.887	1.903	2.496
KStar	0.970	1.289	1.728
LWL	0.819	2.482	3.071
Additive Regression	0.963	1.074	1.431
Bagging	0.979	0.758	1.080
Ensemble selection	0.975	0.853	1.174
Single conjunctive	0.733	2.988	3.625
Decision Table:	0.9737	0.848	1.216
M5 pruned model rules	0.973	0.950	1.213
Zeros	0.115	4.339	5.340
Decision Stump	0.732	2.986	3.927
Trees MP5	0.974	0.946	1.195
REPTree	0.972	0.871	1.235

Çizelge 4.5 deki veriler %66 eğitim, %33 test verileri oranı ile elde edilmiştir. Bunun yanında bu oran değiştirilerek daha iyi sonuçlar elde etme ihtimali vardır. Deneysel olarak elde edilen %80 eğitim, %20 test veri oranı iyi sonuçlar sağlamaktadır. Bunun için en iyi sonuçları aldığımız dört algoritma %80eğitim, %20 test verileri oranı ile tekrar çalıştırılmıştır.

Çizelge 4.6. V-şekilli mikroekranlı hat için ikinci uygulama %80 eğitim, %20 test verileri için korelasyon ve hata

Çıkış alan adı	Zo		
Algoritma adı	Korelasyon katsayısı	MAE	RMSE
Bagging	0.951	0.975	1.543
Ensemble selection	0.940	1.151	1.692
M5 pruned model rules	0.954	1.114	1.546
Trees MP5	0.951	1.107	1.534

Çizelge 4.5. ile çizelge 4.6. kıyaslandığında bu dört algoritma için %66 eğitim, % 34 test veri oranı ile daha iyi sonuçlar alınmaktadır. Bu çizelge ışığında MAE ve RMSE değerleri hesaba katılarak algoritma seçilecektir.

Burada RMSE nin en düşük olduğu değerler:

- 1-Bagging (1,080)
- 2- Ensemble selection (1,174)
- 3-Trees MP5 (1,195)
- 4-M5 pruned model rules (1,213)

Bu algoritmaların korelasyon değerleri ise:

- 1-Bagging (0.979)
- 2- Ensemble selection (0.975)
- 3-Trees MP5 (0.973)
- 4-M5 pruned model rules (0.974)

Hata değeri olarak en düşük Bagging algoritmasıdır. Korelasyon katsayısı (0,979) da en iyi algoritmadır.

V- şekilli mikroekranlı hat için Bagging algoritması seçildi.

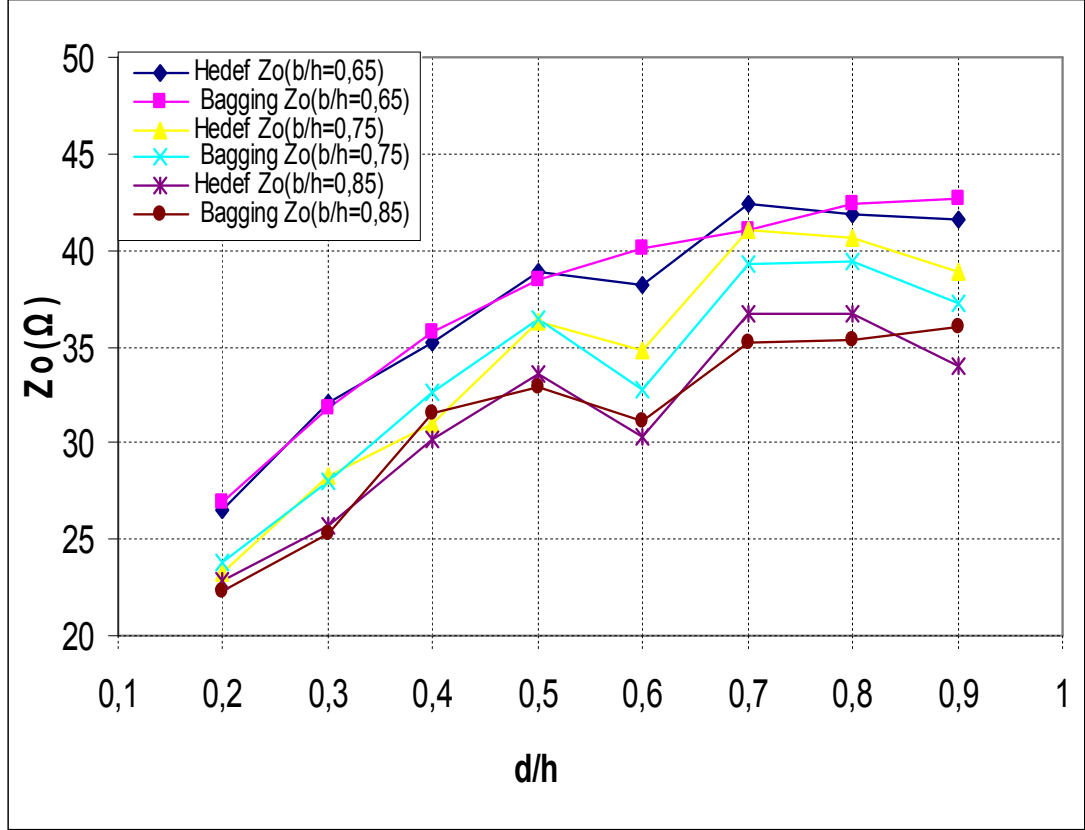
V- şekilli mikroekranlı hat için d/h, β , b/h, a/h değerleri girdi olarak kabul edilirken Zo değeri çıktı olarak kabul edilmektedir. Girdiler için:

d/h =0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9

β =30,37,45,52,60 derece

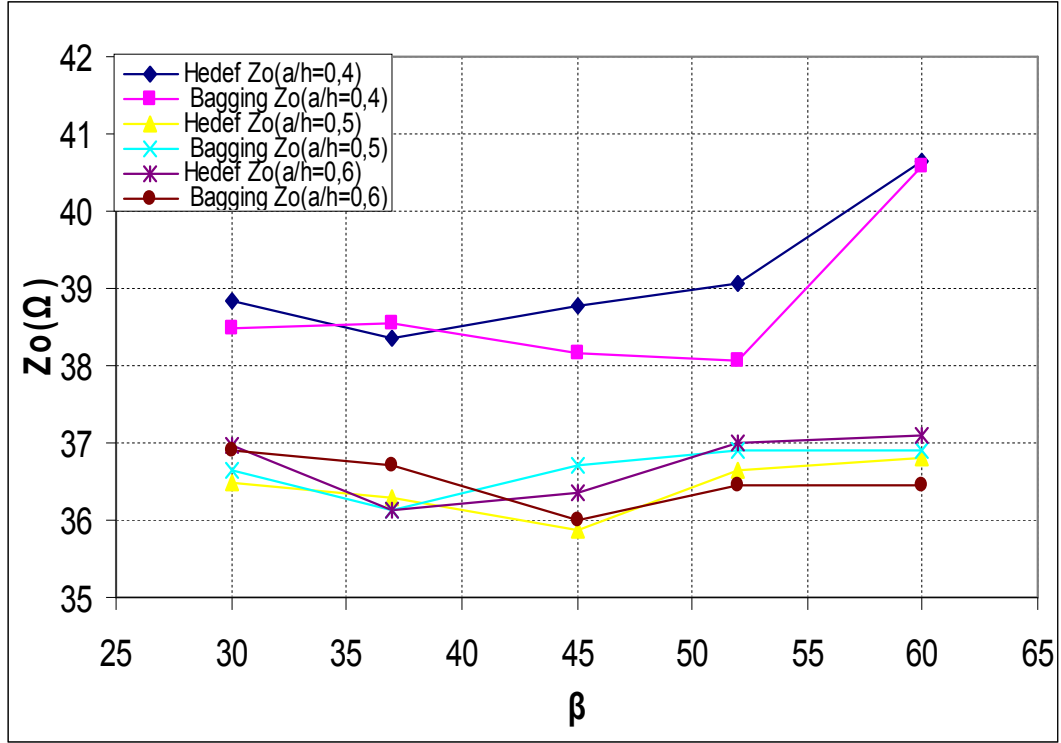
$b/h=0.65,0.75,0.85$

$a/h=0.4, 0.5, 0.6$ Bu girdilerden $d/h, b/h$ için bir grafik ve $\beta, a/h$ için bir grafik toplam iki farklı grafik hazırlanmıştır.



Şekil 4.5 d/h ve b/h in farklı değerlerine göre hedef Z_o ve Bagging algoritması çıkış Z_o dır. $\beta =30$ $a/h=0,4$ alınmıştır.

d/h değeri 0,2 den 0,9 artırıldığında $b/h= 0,65$ için çıkış empedansı da 26 Ω dan 41 Ω e kadar yükselmiştir. Aynı şekilde $d/h=0,9$ için b/h değeri 0,85 den 0,65 e düşürüldüğünde çıkış empedansı 41 Ω den 34 Ω e kadar düşmektedir.



Şekil 4.6 a/h ve β nin farklı değerlerine göre hedef Z_o ve Bagging Z_o .
 $d/h=0,5$ $b/h=0,65$ alınmıştır.

β nin artan değerleri için çıkış empedansı $a/h=0,4$ ve $\beta=60$ derece dışında genel olarak değişmemiştir. Aynı şekilde a/h ın artan değerleri için çıkış empedansı $a/h=0,4$ değeri dışında genel olarak değişmemiştir. V-şekil için hata değerleri diğer geometrilere nazaran biraz yüksek olsa da genel olarak iyi seviyededir. a/h daki değişime rağmen Z_o da fazla değişim olmadığı ve Z_o değerleri küçük olduğu için grafikteki eğriler birbirlerine karışmaktadır. Hedef ve Bagging değerleri birbirlerine yakındır bir Ω dan daha küçüktür. Bu da hatanın küçük olduğunu göstermektedir. Bu geometrik şekiller için veri madenciliği performans ölçütü olarak interpolasyon ve ekstrapolasyon da yapılmıştır. Ekstrapolasyonda:

Çizelge 4.7. V- şekilli mikroekranlı hat Bagging algoritması için ekstrapolasyon test sonuçları

	Korelasyon	MAE	RMSE
d/h=0,2	0,948	0,655	0,801
d/h=0,9	0,986	0,608	0,869
$\beta=30$	0,983	1,129	1,345
$\beta=60$	0,986	0,595	0,888
b/h=0,65	0,986	0,621	0,891
b/h=0,85	0,985	0,607	0,879
a/h=0,4	0,981	1,521	1,792
a/h=0,6	0,984	0,622	0,942

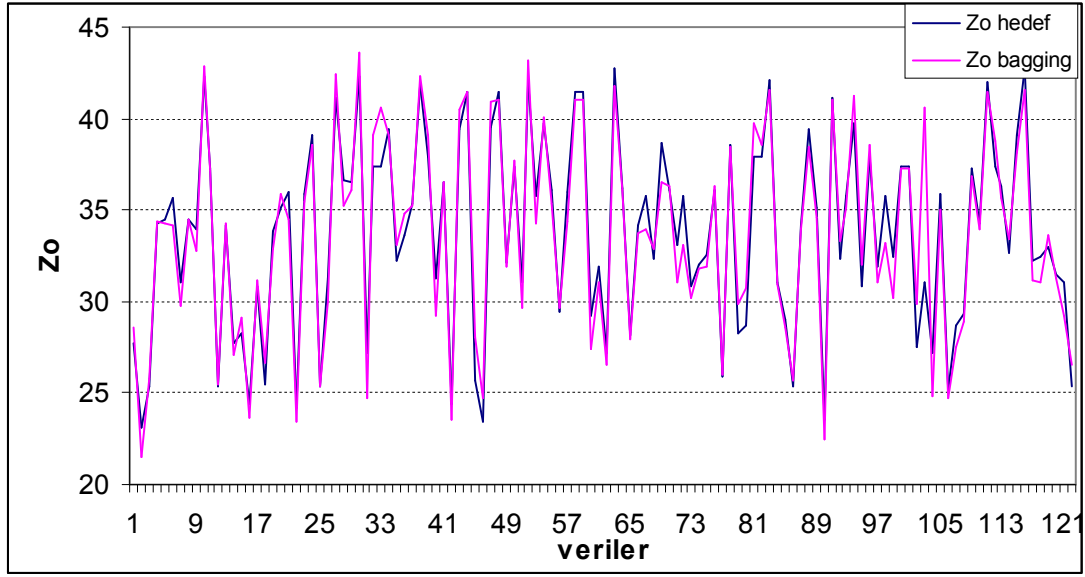
Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. Ekstrapolasyon test sonuçlarına göre V-şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,98 mertebesindedir. Bu sonuçlar girdi çıktı arası ilişkinin iyi ve hatanın sifıra yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. İyi değerler sağlanmıştır. İnterpolasyona bakılırsa:

Çizelge 4.8. Eliptik V- şekilli mikroekranlı hat Bagging algoritması için interpolasyon test sonuçları

Test verileri	Korelasyon	MAE	RMSE
d/h \neq 0,2 ve d/h \neq 0,9	0,993	0,634	0,793
$\beta\neq 30$ ve $\beta\neq 60$	0,986	0,693	0,891
b/h=75	0,988	0,596	0,869
a/h=0,5	0,979	1,172	1,459

Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. İnterpolasyon test

sonuçlarına göre V-şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,99 mertebesindedir. Bu sonuçlar girdi çıktı arası ilişkinin mükemmel ve hatanın sıfıra yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. İyi değerler sağlanmıştır. Son olarak hedef ve hata değerleri aynı grafikte çizilmiştir. Böylece hatanın seviyesi net olarak görülmüştür.



Şekil 4.7.hedef Zo ile Bagging algoritmasının bulduğu Zo

Şekil 4.7 deki veriler Bagging algoritmasının test işleminde kullandığı verilerdir. Açıkça görülmektedir ki hedef Zo ile bulunan Zo (Bagging) birbirlerine çok yakındır. Hatanın seviyesi burada da açıkça görülmektedir.

4.3. Eliptik Şekilli Mikroekranlı Hatta Veri Madenciliği Uygulaması

Çizelge 4.9. Eliptik Şekilli Mikroekranlı hat birinci uygulama. % 66 eğitim, %34 test verileri için hesaplanan korelasyon ve hata

Çıkış alan adı	Zo		
	Korelasyon katsayısı	MAE	RMSE
Isotonic regression	0.565	14,612	19.925
Least medswq	0.901	6,300	12.613
Lineer Regression	0.910	6.709	9.885
Multilayer Perceptron	0.998	1.036	1.303
Pace Regression	0.910	6.716	9.826
Simple Linear regression	0.567	14.557	19.674
SMOreg	0.907	5.836	10.814
IBk (instance-based)	0.955	5.153	7.082
KStar	0,976	4.135	7.041
LWL	0.655	13.051	18.836
Additive Regression	0.900	7.714	10.451
Meta.Bagging	0.998	1.074	1.394
Ensemble selection	0,996	1,659	2,131
Random subspace	0,994	8,194	12,047
Regresyon by discretization,	0.986	3.274	3.869
Single conjunctive	0.566	13.731	19.691
Decision Table:	0.996	0.578	1.907
M5 pruned model rules	0.998	1.046	1.505
Decision Stump	0.524	14.451	20.331
Trees M5P	0.998	0.878	1.450
REPTree	0.997	1.487	1.847

Çizelge 4.9. de Eliptik şekilli mikroekranlı hat için %66 eğitim, %33 test verileri oranı ile yapılan birinci uygulama sonucu korelasyon ve hata değerleri elde edilmiştir. Bunun yanında bu oran değiştirilerek daha iyi sonuçlar elde etme imkanı vardır. Deneysel olarak elde edilen %80 eğitim, %20 test veri oranı uygulamaları iyi sonuçlar sağlayabilmektedir. Bunun için en iyi sonuçları aldığımız dört algoritmayı %80 eğitim, %20 test verileri oranı ile tekrar çalıştırılmıştır.

Çizelge 4.10. ESML de ikinci uygulama. %80 eğitim, %20 test veri oranı için korelasyon ve hata

Multilayer perceptron	0.999	1.058	1.330
Meta bagging	0.998	1.125	1.560
Decision table	0.991	2.765	3.342
M5pruned model rules	0.997	1.229	1.746
Trees M5P	0.997	1.044	1.811

Çizelge 4.10. ile çizelge 4.6.a görüldüğü gibi birinci ve ikinci uygulamalar kıyaslandığında %66eğitim, %34 test veri oranı ile birinci uygulamada daha iyi sonuçlar alınmıştır.

Bu çizelge ışığında MAE ve RMSE değerleri hesaba katılarak algoritma seçilecektir. Burada RMSE nin en düşük olduğu değerler:

- 1) Multilayer perceptron (1.303)
- 2) Meta Bagging (1.560)
- 3) M5 pruned model rules (1.746)
- 4) Trees M5P (1.811)

Bu algoritmaların korelasyon değerleri ise:

- 1) Multilayer perceptron (0.998)
- 2) Meta Bagging (0.998)
- 3) M5pruned model rules (0.998)
- 4) Trees M5P (0.998)

Hata değeri olarak en düşük Multilayer perceptron algoritması çıktı. Ki bunun RMSE değeri 1.303 dür. Buna en yakın Meta Bagging algoritması var. Onun ise 1.560 dır. Korelasyon katsayıları aynı değerde ve iyi seviyede.

Eliptik şekilli mikroekranlı hat için Multilayer perceptron algoritması seçildi.

Eliptik şekilli mikroekranlı hat için b , w , h_1 , h_2 , ϵ_r , değerleri girdi olarak kabul edilirken Z_0 değeri çıktı olarak kabul edilmektedir. Girdiler için:

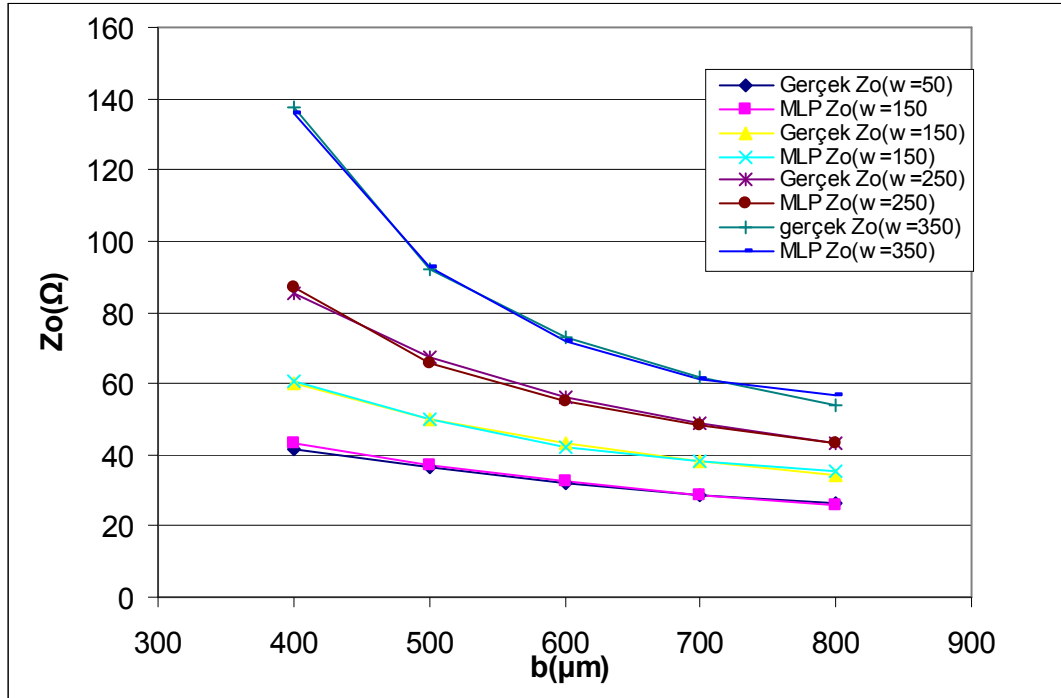
$b = 400, 500, 600, 700, 800 \mu\text{m}$

$w = 50, 150, 250, 350 \mu\text{m}$

$h_1 = 2, 3, 4 \mu\text{m}$

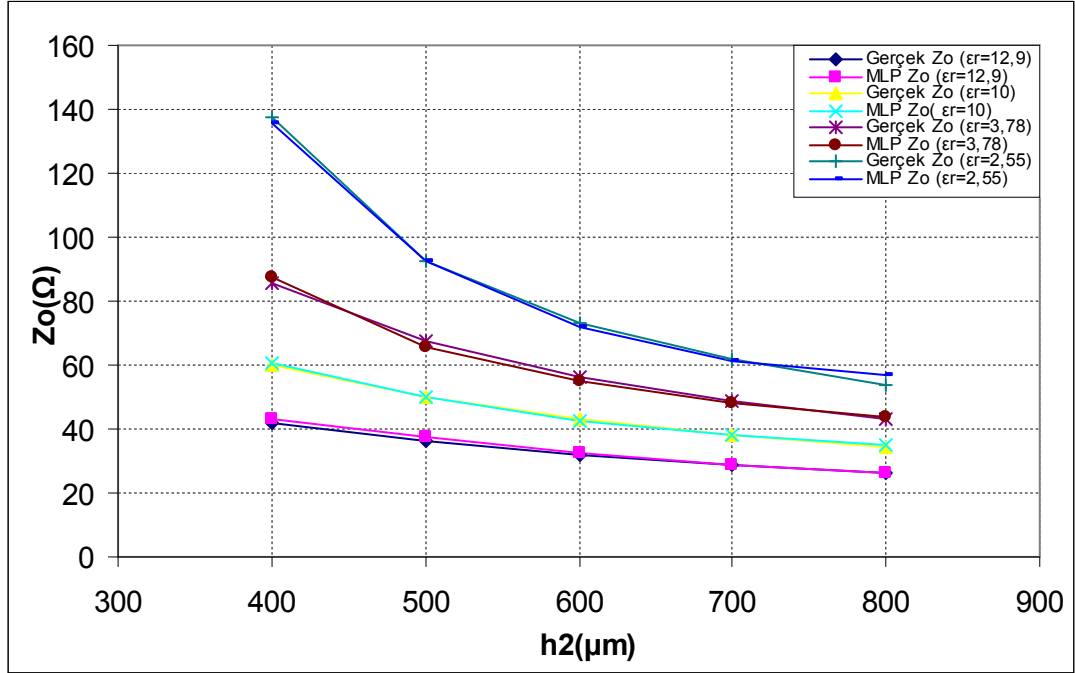
$h_2 = 400, 500, 600, 700, 800 \mu\text{m}$

$\epsilon_r = 2.55, 3.78, 10, 12.9$ Bu girdilerden b, w için bir grafik ve h, ϵ_r için bir grafik toplam iki farklı grafik hazırlanmıştır.



Şekil 4.8 b ve w nin farklı değerlerine göre hedef Z_o ve MLP algoritması çıkış Z_o dır. $h_1 = 2\text{mm}$, $b = 400\mu\text{m}$, $w = 50\mu\text{m}$ alınmıştır.

$w = 50$ için b nin değerleri 400den 800 çıkarken çıkış empedansı 43Ω dan 26Ω a kadar düşmektedir. aynı zamanda $b = 400$ için w değeri 50 den 350 arttırıldığında çıkış empedansı 43Ω den 135Ω a kadar artmaktadır.



Şekil 4.9. h_2 ve ϵ_r nin farklı değerlerine göre hedef Z_o ve MLP algoritması çıkış Z_o
 $h_1=2\text{mm}$, $b=400\mu\text{m}$, $w=50\mu\text{m}$ alınmıştır.

$\epsilon_r=12,9$ için h_2 nin değerleri 400den 800 çıkarken çıkış empedansı 43 Ω dan 26 Ω a kadar düşmektedir. aynı zamanda $h_2=400$ için ϵ_r değeri 2,55 den 12,9 arttırıldığında çıkış empedansı 43 Ω den 135 Ω a kadar artmaktadır.

Bu geometrik şekiller için veri madenciliği performans ölçütü olarak interpolasyon ve ekstrapolasyon yapılmıştır. Ekstrapolasyonda:

Çizelge 4.11 Eliptik şekilli mikroekranlı hat MLP algoritması için ekstrapolasyon test sonuçları

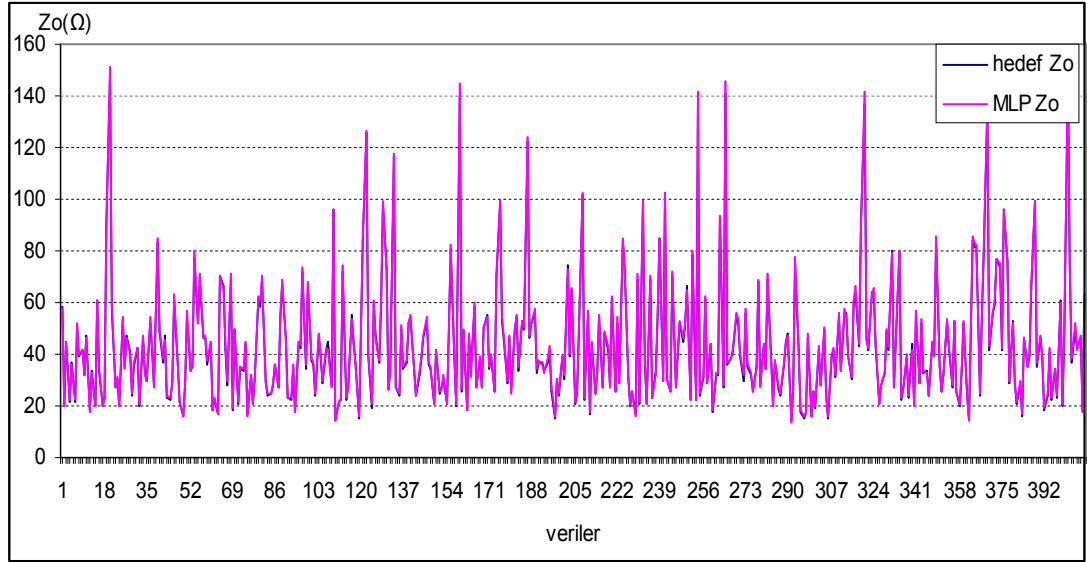
Test verileri	Korelasyon	MAE	RMSE
b=400	0,958	9,956	14,71
b=800	0,993	1,619	2,032
w=50	0,988	1,459	1,917
w=350	0,952	6,413	11,172
h ₁ =2	0,992	1,160	1,396
h ₁ =4	0,993	1,203	1,382
h ₂ =400	0,993	1,495	1,676
h ₂ =800	0,999	1,249	1,468
$\epsilon_r=2,55$	0,993	1,297	2,055
$\epsilon_r=12,9$	0,998	2,055	2,252

Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. Ekstrapolasyon test sonuçlarına göre eliptik şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,99 mertebesinde. Bu sonuçlar girdi çıktı arası ilişkinin çok iyi ve hatanın sıfıra çok yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. Çok iyi değerler sağlanmıştır. İnterpolasyona bakılırsa:

Çizelge 4.12. Eliptik şekilli mikroekranlı hat MLP algoritması için interpolasyon test sonuçları

Test verileri	Korelasyon	MAE	RMSE
b \neq 400 ve b \neq 800	0,992	1,509	2,051
w \neq 50 ve w \neq 350	0,998	1,007	1,374
h ₁ =3	0,993	0,993	1,284
h ₂ \neq 400 ve h ₂ \neq 800	0,999	0,976	1,173
$\epsilon_r \neq 2,55$ ve $\epsilon_r \neq 12,9$	0,998	2,504	3,386

Yukarıdaki çizelge hazırlanırken sol sütundaki veriler test verileri olarak ayrılmış geriye kalan veriler öğrenme verileri olarak kullanılmıştır. İnterpolasyon test sonuçlarına göre eliptik şekilli mikroekranlı hattın korelasyon katsayıları bire çok yakındır. 0,995 mertebesindedir. Bu sonuçlar girdi çıktı arası ilişkinin mükemmel ve hatanın sıfıra çok yakın olduğunu göstermektedir. MAE ve RMSE değerleri de yeterince düşük çıkmaktadır. Çok iyi değerler sağlanmıştır. Son olarak hedef ve hata değerleri aynı grafikte çizilmiştir. Böylece hatanın seviyesi net olarak görülmüştür.



Şekil 4.10.hedef Zo ile MLP algoritmasının bulduğu Zo

Şekil 4.10 deki veriler MLP algoritmasının test işleminde kullandığı verilerdir. Açıkça görülmektedir ki hedef Zo ile bulunan Zo (MLP) birbirlerine çok yakındır. Hatanın seviyesi burada da açıkça görülmektedir.

5. SONUÇ

Hedef verilerin ile veri madenciliği uygulaması yapılarak elde edilen verilerin kıyaslaması yapılarak veri madenciliği yönteminin performansını ortaya koymak amacıyla:

Trapez mikoekranlı hattı modellemede M5P algoritması en iyi korelasyon, MAE ve RMSE değerlerine sahip algoritma olduğu görülmüştür. M5P algoritmasının hesapladığı empedans ve hedef empedans bulunmuştur. $a/b, a/h$ ve β değerleri için çizilen grafiklerde M5P algoritmasının hesapladığı empedans değerleri a/h in artan değerleri için azalmıştır; β nın artan değerlerinde için değişmemiştir; a/b in artan değerlerinde azalmıştır. Bu esnada M5P algoritmasının hesapladığı empedans ile hedef empedans birbirlerine çok yakın (bir ohm dan küçük) değerler sağlamıştır. Başarı ölçütü olarak yapılan interpolasyon hesaplamasında 0,991 korelasyon değerleri sağlamıştır. Ekstrapolasyon hesaplamalarında ise 0,996 korelasyon değerleri sağlanmıştır.

V-şekilli mikoekranlı hattı modellemede Bagging algoritması en iyi korelasyon, MAE ve RMSE değerlerine sahip algoritma olduğu görülmüştür. Bagging algoritmasının hesapladığı empedans ve hedef empedans bulunmuştur. $a/h, b/h, d/h$ ve β değerleri için çizilen grafiklerde Bagging algoritmasının hesapladığı empedans değerleri d/h in artan değerleri için artmıştır; β nın artan değerlerinde için değişmemiştir; a/h ve a/b in artan değerlerinde azalmıştır. Bu esnada Bagging algoritmasının hesapladığı empedans ile hedef empedans birbirlerine çok yakın (bir ohm dan küçük) değerler sağlamıştır. Başarı ölçütü olarak yapılan interpolasyon hesaplamasında 0,985 korelasyon değerleri sağlamıştır. Ekstrapolasyon hesaplamalarında ise 0,984 korelasyon değerleri sağlanmıştır.

Eliptik şekilli mikoekranlı hattı modellemede Multilayer perceptron algoritması en iyi korelasyon, MAE ve RMSE değerlerine sahip algoritma olduğu görülmüştür. Multilayer perceptron algoritmasının hesapladığı empedans ve hedef empedans

bulunmuştur. w, b, h_2 ve ϵ_r değerleri için çizilen grafiklerde Multilayer perceptron algoritmasının hesapladığı empedans değerleri w ve h_2 nin artan değerleri için azalmıştır; b nin artan değerlerinde için artmıştır; ϵ_r in artan değerlerinde azalmıştır. Bu esnada Multilayer perceptron algoritmasının hesapladığı empedans ile hedef empedans birbirlerine çok yakın (bir ohm dan küçük) değerler sağlamıştır. Başarı ölçütü olarak yapılan interpolasyon hesaplamasında 0,993 korelasyon değerleri sağlamıştır. Ekstrapolasyon hesaplamalarında ise 0,984 korelasyon değerleri sağlanmıştır.

Bu bilgiler ışığında veri madenciliği yöntemi mikroekranlı hatların modellemesinde başarı ile kullanılabilir.

KAYNAKLAR

- Akpınar, H., Nisan 2000. Veritabanlarında Bilgi Keşfi ve Veri Madenciliği. İstanbul, İstanbul üniversitesi İşletme Fakültesi Dergisi, 22,29.
- Alpaydın, E., 2000. Zeki Veri Madenciliği. http://www.cmpe.boun.edu.tr/%7Eeethem/filespapers/veri-maden_2k-notlar.doc. Erişim Tarihi:15.02.2008.
- Alkan,A., 2004. Veri Madenini Keşfetmek. <http://groups.yahoo.com/group/verimadenciligi/files>. Erişim Tarihi: 01.08.2005.
- Asilkan, Ö., 2008.Veri Madenciliği Kullanılarak İkinci El Otomobil Pazarında Fiyat Tahmini. Antalya Akdeniz Üniversitesi doktora tezi,157 Akdeniz Üniversitesi
- Aydın, İ., 2006, Arıza Teşhisinde Veri Madenciliği Ve Yumuşak Hesaplama Tekniklerinin Kullanımı
- Berry, M., Linoff,G., 1997. Data Mining Techniques for Marketing, Sales and Customer Support. John Wiley & Sons.
- Barrett, R.M., Barnes, M.H., 1951, Microwave Circuits, National Conference on Air borne Electronics, IRE, Ohio, May
- Anonim, FGG Bilişim, 2006. Veri Madenciliği. <http://www.fggint.com/veri/verimadenciligi.jsp>. Erişim Tarihi: 09.02.2006.
- Anonim Crisp,DM., www.crisp-dm.org, 01.06.2008.
- Demir, H., Gümüşoğlu S., 2003. Üretim Yönetimi-İşlemler Yönetimi. Beta Basım
- Deniz, E., 2006. Bulanık Mantık Tabanlı Tahmin Modeli ve Uygulaması, Yüksek Lisans Tezi, Muğla Üniversitesi,
- Dilly, R., , 1999. Data Mining: An Introduction www.pcc.qub.ac.uk/tec/courses/datamining/ohp/dm-OHP-final_1.html, 01.12.2007.yayım, 6. Baskı, İstanbul,
- Fayyad, U., Piatetsky-Shapiro G., Smyth P., 1996. From Data Mining To Knowledge Discovery Ain Databases”, AI Magazine, 37-54.
- Frawley, W. J., Piatetsky-Shapiro G., Matheus C. J., 1991. Knowledge Discovery. Databases: An Overview Knowledge Discovery In Databases. Cambridge, 27
- Grieg, D.D., Engelmann, H.F., 1952, Microstrip-A New Transmission Technique for Kilomegacycle Range, Proc. IRE, 40, 1644-1650
- Gürsakal, N., Oğuzlar A., Senturk A., “Değişen Veri Kavramı ve Yeni Alanlar”, İstatistik Araştırma Sempozyumu, (Kasım 2000).
- Han, J, Kamber, M, 2001 Data Mining Concepts and Techniques, Morgan Kaufmann, USA
- Hyndman R. J., 2002. “Box-Jenkins Modelling, Informed Student Guide to Management Science, London, 1-2.
- Karakaş, M., 2002. Veri Madenciliği Üzerine. http://www.bilgiyonetimi.org/cm/pages/mkl_gos.php?nt=132. Erişim Tarihi:20.02.2008.
- Kdnuggets,l., Mart 2007, www.kdnuggets.com/polls/2007/data_mining_methods.htm, 15.06.2008
- Koyuncugil A. S., 2007. Veri Madenciliği ve Sermaye Piyasalarına Uygulanması, Sermaye Piyasası Kurulu Araştırma Raporu,
- Küçükşille, E.U., 2005. Veri Madenciliği Süreci. <http://www.sqlnedir.com/Members/ArticleDetail.aspx?Id=85>. Erişim Tarihi: 10.04.2008.

- Küçükşille, E.U., 2005. Veri Madenciliği (Data Mining) Nedir? <http://www.sqlnedir.com/Members/ArticleDetail.aspx?Id=81>. Erişim Tarihi:15.03.2008.
- Larose D.T., 2005. Discovering Knowledge in Data. An Introduction To Data John Wiley & Sons Inc., 42, 70.
- Özmen, S., 2001. Hayatı Veri Madenciliği ile _statistik Uygulamalarını Yeniden Keşfediyor. http://suleozmen.marmara.edu.tr/teblig_sunumlar/3is_hayati_veri_madencili_istatistik.pdf. Erişim Tarihi: 02.02.2006.
- Özyirmioğlu, E, 2009 İmalat Verilerinin Modellenmesi ve Analizi
- Peters, R.W, 1956, Handbook of Tri-Plate Microwave Components, Sander Associates, Nashua, New Hampshire,
- Türkmen M, 2009, Mikrodalga iletim hatlarının bulanık sinir ağları ile modellenmesi, Erciyes Üniversitesi
- Türkmen M Yıldız C., Kaya S. ve., 2006 Güney K Neural Models for the V-Shaped Conductor-Backed coplanar wagueguides
- Türkmen M., Yıldız C., Kaya S. ve. Güney, K, 2006 “Artificial neural networks for calculating the characteristic impedance of air-suspended trapezoidal and rectangular-shaped microshield lines”, Journal of Electromagnetic Waves and Applications, 20(9), s.1161-1174,
- Türkmen M., Yıldız C., Kaya S. ve Güney, K., 2008 Neural Models for the Elliptic and Circular Shaped Microshied Lines
- Türkmen M Güney, K., Yıldız C., Kaya S. ve, 2009, Adaptive-network-based fuzzy Inference system models for Computing the characteristic Impedances of air-suspended Trapezoidal and Rectangular-shaped Microshield lines
- .,Türkmen M Yıldız C., Kaya S. ve Güney K, Dairesel Mikroekranlı Hatların Quasi-Statik Analizinin Yapay Sinir Ağları ile Gerçekleştirilmesi, Erciyes Üniversitesi,
- Nihad I. Dib, and Linda P. B. Katehi,1992 mikroekranlı hat için empedans hesabı
- Naichang Yuan, Chengli Ruan and Weigan Lin, , V, Eliptik, Dairesel, Şekilli Mikroekranlı Hatların Analitik Analizi, 1994
- Kwok-Keung M. Cheng And Ian D. Robertson Asimetrik, V-Şekilli Mikroekranlı Hat Tasarımı ve Analizi İçin Basit ve Açık Formüller, 1995

ÖZGEÇMİŞ

Adı Soyad : Sedat ATEŞ

Doğum Yeri : ANKARA

Doğum Tarihi: 26.10.1976

Medeni Hali : Evli

Yabancı Dili : İngilizce



Eğitim Durumu (Kurum ve Yıl)

Lise : Veteriner Sağlık Meslek Lisesi 1995

Lisans : Süleyman Demirel Üniversitesi 2001

Yüksek Lisans : Süleyman Demirel Üniversitesi ---

Çalıştığı Kurum/Kurumlar ve Yıl

Tarım ve Köy İşleri Bakanlığı 15

Yayımları (SCI ve diğer)