



T.C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

BEYAZ EŞYA SEKTÖRÜNDE SATIŞ TAHMİNİ: BİR
VERİ MADENCİLİĞİ UYGULAMASI
YÜKSEK LİSANS TEZİ

EZGİ DEMİRER POLAT

BURSA-2022



T.C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
EKONOMETRİ ANABİLİM DALI
İSTATİSTİK BİLİM DALI

**BEYAZ EŞYA SEKTÖRÜNDE SATIŞ TAHMİNİ: BİR
VERİ MADENCİLİĞİ UYGULAMASI**
YÜKSEK LİSANS TEZİ

EZGİ DEMİRER POLAT
ORCID: 0000-0002-9557-9670

Danışman:
Prof. Dr. AYŞE OĞUZLAR

BURSA-2022

17/02/2022

Yemin Metni

Yüksek Lisans tezi olarak sunduğum "Beyaz Eşya Sektöründe Satış Tahmini: Bir Veri Madenciliği Uygulaması." başlıklı çalışmanın bilimsel araştırma, yazma ve etik kurallarına uygun olarak tarafımdan yazıldığına ve tezde yapılan bütün alıntılarının kaynaklarının usulüne uygun olarak gösterildiğine, tezimde intihal ürünü cümle veya paragraflar bulunmadığına şerefim üzerine yemin ederim.

Tarih ve İmza

Adı Soyadı: Ezgi DEMİRER POLAT

Öğrenci No: 701817002

Anabilim Dalı: Ekonometri

Programı: Tezli Yüksek Lisans Programı

Tezin Türü: Yüksek Lisans

ÖZET

Yazar Adı ve Soyadı : Ezgi DEMİRER POLAT

Üniversite :Bursa Uludağ Üniversitesi

Enstitü : Sosyal Bilimler Enstitüsü

Anabilim/Anasanat Dalı: Ekonometri

Bilim/Sanat Dalı : İstatistik

Tezin Niteliği : Yüksek Lisans Tezi

Sayfa Sayısı : 13+79

Mezuniyet Tarihi : 17/02/2022

Tez Danışmanı : Prof. Dr. Ayşe OĞUZLAR

BEYAZ EŞYA SEKTÖRÜNDE SATIŞ TAHMİNİ: BİR VERİ MADENCİLİĞİ UYGULAMASI

Büyüme hedefi olan ve sürekli gelişmeyi hedefleyen firmaların hayatta kalarak varlığını sürdürmeleri için; yeni öngörüler, stratejiler ve en önemlisi onlara yol gösterebilecek veri analizlerine ihtiyaç duymaktadırlar. Bu analizler sayesinde firmalar, yoğun rekabet ortamında hızlı ve en doğru kararı verebilir durumda olabileceklerdir. Varlıklarını devam ettirebilmek isteyen üreticiler, bir yandan artan müşteri beklentilerini karşılamak isterken bir yandan da maliyetlerini en az seviyede tutmayı amaçlamaktadırlar. Firmalar içinde buldukları mevcut yapı öncesine bir yolculuk yaptıklarında, karar mekanizmalarını daha doğru kullandıklarını düşünerek hareket ederler. Bu sebeptir ki, deneyimler bizlere her zaman doğru yolu gösterir sözü bu noktada kanıtlanabilir nitelikte olmaktadır. Bu çalışmada, firmaların geçmiş yıllara ait ciro ve karlılık durumu analiz edilerek veri madenciliği teknikleri içerisinde yer alan; yapay sinir ağı, destek vektör makineleri ve regresyon ile doğru sonuca ulaşması hedeflenmektedir. Önceki yılların verileri analiz edilerek, yeni bir yılın verileri hakkında ön görüş sahibi olmak hedeflenmektedir. Bu doğrultuda, geliştirilebilecek stratejiler sonucunda karlılık seviyesinin artacağını düşünmemek elde değildir. Bu çalışmada amaçlanan, bu doğrusal düşünceyi matematiksel olarak ispatlayarak hedefe sağlam adımlar ile ilerleyebilmektir. Durumun kanıtlanabilmesi amacıyla, beyaz eşya sektöründe faaliyet gösteren firmanın; sevk verileri bağımlı değişken, reel kesim güven endeksi, gayri safi yurtiçi hasıla, sanayi üretim endeksi, evlenme oranı, tüketici güven endeksi ve ekonomi güven endeksi ise bağımsız değişken olarak aylık bazda alınmıştır. Veriler uygulama aşamasında, Weka programında test edilmiştir. Satış tahminleri üç ayrı yöntem ile ayrı ayrı test edilmiş ve sonuçlar birbirleri ile kıyaslanmıştır.

Anahtar kelimeler: Veri madenciliği, satış tahmini, weka programı, yapay sinir ağı, destek vektör makinesi, regresyon

ABSTRACT

Name and Surname : Ezgi DEMİRER POLAT

University : Bursa Uludag University

Institution : Social Science Institution

Field : Econometrics

Branch : Statistics

Degree Awarded: Master

Page Number : 13+79

Degree Date 17/02/2022

Supervisor : Professor Dr. Ayşe OĞUZLAR

SALES FORECASTING IN THE WHITE GOODS INDUSTRY: A DATA MINING APPLICATION

For companies with growth targets and continuous development to survive and thrive; they need new insights, strategies and, most importantly, analysis of the data that can guide them. Thanks to these analyses, companies will be able to make the fastest and most accurate decision in an intensely competitive environment. In this study, the turnover and profitability status of the companies for the past years are analyzed and included in the data mining techniques; it is aimed to achieve the right result with artificial neural network, support vector machines and regression. By analyzing the data of previous years, it is aimed to have a foresight about the data of a new year. Accordingly, it is notable not to think that the level of profitability will increase as a result of the strategies that can be developed. The aim of this study is to prove this linear thinking mathematically and to move forward with solid steps towards the goal. In order to prove the situation, the company operating in the white goods sector; the referral data are taken on a monthly basis as dependent variable, real sector confidence index, gross domestic product, industrial production index, marriage rate, consumer confidence index and economic confidence index as independent variables. The data was tested in the Weka program during the implementation phase. Sales estimates were tested separately by three different methods and the results were compared with each other.

Keywords: Data mining, sales forecasting, Weka, artificial neural network, support vector machine, regression.

TEŐEKKÜR

Tez alıőmasını; planlanmasında, yřrřtřlmesinde ve oluőumunda deęerli desteklerini esirgemeyen kıymetli hocam Prof. Dr. Ayőe OęUZLAR'a sonsuz teőekkřrř bir bor bilirim.

Do. Dr. Selim TřZřNTřRK'e ve Do. Dr. Burcu řNGEN BİLİR'e de deęerli gřrřőlerinden ve desteklerinden dolayı teőekkřrř ederim.

Beni bugřnlere ulaőtıran, sonsuz sevgi ve desteklerini esirgemeyen sevgili anne ve babama, baőaracađıma olan inancıyla beni her daim destekleyen sevgili eőime, bu yola ıkmam iin cesaret veren akıl hocam Břlent řZGEN'e, her zaman yanımda olan Hanife ve Metin POLAT'a teőekkřrřlerimi sunarım.

İÇİNDEKİLER

YÜKSEK LİSANS İNTİHAL YAZILIM RAPORU Hata! Yer işareti tanımlanmamış.	
ÖZET.....	ii
ABSTRACT.....	iii
TEŞEKKÜR.....	iv
İÇİNDEKİLER.....	v
TABLolar DİZİNİ.....	viii
ŞEKİLLER DİZİNİ.....	ix
KISALTMALAR.....	x
GİRİŞ.....	1

BİRİNCİ BÖLÜM

VERİ MADENCİLİĞİ KAVRAMI

1. Veri Madenciliği.....	2
2. Veri Madenciliğinin Tarihsel Gelişimi.....	3
3. Veri Madenciliği Kullanım Alanları.....	4
4. Veri Madenciliğinin Kuvvetli ve Zayıf Yönleri.....	6
5. Veri Madenciliği Modelleri.....	6
5.1. Tanımlayıcı Modeller.....	7
5.1.1. Kümeleme Yöntemi.....	8
5.1.2. Birliktelik Kuralı.....	9

İKİNCİ BÖLÜM

TAHMİN EDİCİ MODELLER

1. Literatür Taraması.....	12
2. Sınıflama modelleri.....	19
3. Yapay Sinir Ağları.....	19
3.1. Yapay Sinir Ağlarının Öğrenme Süreci.....	22
3.2. Yapay Sinir Ağlarının Kullanım Alanları.....	23
3.3. Yapay Sinir Ağlarının Genel Özellikleri.....	24

3.4. Yapay Sinir Ağı Modelleri	24
3.4.1. Tek Katmanlı Algılayıcılar	25
3.4.2. Çok Katmanlı Algılayıcılar	26
3.4.3. İleri Beslemeli Yapay Sinir Ağları	27
3.4.4. Geri Beslemeli Yapay Sinir Ağları	28
3.5. Öğrenme Yöntemine Göre Yapay Sinir Ağları	28
3.5.1. Öğretmenli Öğrenme	29
3.5.2. Öğretmensiz Öğrenme	29
4. Regresyon Analizi	30
4.1. Basit Doğrusal Regresyon	31
4.2. Çoklu Regresyon Modeli	35
4.3. Lojistik Regresyon Modeli	37
5. Destek Vektör Makineleri	38
5.1. Destek Vektör Makinelerinin Avantajları	40
5.2. Destek Vektör Makinelerinin Dezavantajları	41
5.3. SMO Algoritması	41
6. Tahmin Performansları	42
6.1. Ortalama Mutlak Hata	42
6.2. Kök Ortalama Kare Hatası	43
6.3. Ortalama Mutlak Yüzde Hata	43

ÜÇÜNCÜ BÖLÜM

BEYAZ EŞYA SEKTÖRÜNDE SATIŞ TAHMİNİ UYGULAMASI

1. Uygulamanın Amaç ve Önemi	44
2. Weka Program Tanıtımı	44
3. Satış Tahmininde Kullanılan Değişkenler	46
4. Firma Tanıtımı	50
5. YSA Uygulaması	52
5.1. Öğrenme ve Momentum Katsayısının Belirlenmesi	53
5.2. Çevrim Katsayısının Belirlenmesi	54
5.3. Gizli Hücre ve Nöron Sayısının Belirlenmesi	55
6. Doğrusal Regresyon Uygulaması	60

7. Destek Vektörleri Makineleri Uygulaması	61
8. Uygulama Sonuçlarının Değerlendirilmesi.....	64
SONUÇ	67
EKLER.....	70
KAYNAKÇA.....	72

TABLolar DİZİNİ

Tablo 1 Veri Madenciliđi Tarihçesi	3
Tablo 2 Veri Madenciliđi Kullanım Alanları.....	5
Tablo 4 YSA Hata Oranı Sonuçları	57
Tablo 5 YSA Tahmin Sonuçları.....	58
Tablo 6 Regresyon Sonuçları	60
Tablo 7 Regresyon Tahmin Sonuçları.....	60
Tablo 8 DVM Hata Oranı Sonuçları	61
Tablo 9 DVM Tahmin Sonuçları	62
Tablo 10 2022* Tahmin Sonuçları.....	63
Tablo 11 Hata Oranı Karşılaştırması	65
Tablo 12 Tahmin Karşılaştırmaları	65

ŞEKİLLER DİZİNİ

Şekil 1 VM Uygulama Şeması.....	2
Şekil 2 Veri Madenciliği Modelleri	7
Şekil 3 Kümeleme Örneği.....	8
Şekil 4 Sinir Hücre Yapısı	20
Şekil 5 Nöron Yapısı.....	20
Şekil 6 YSA Katmanları.....	21
Şekil 7 Tek katmanlı algılayıcı modeli	26
Şekil 8 Çok katmanlı algılayıcı modeli.....	26
Şekil 9 İleri Beslemeli Yapay Sinir Ağı.....	27
Şekil 10 Geri beslemeli ağ yapısı.....	28
Şekil 11 Öğretmenli Öğrenme Modeli.....	29
Şekil 12 Basit Doğrusal Regresyon Modeli.....	31
Şekil 13 Sevk Adetleri Grafik	47
Şekil 14 RKGE Grafik	47
Şekil 15 SÜE Grafik	48
Şekil 16 GSYH Grafik	48
Şekil 17 Evlenme Oranı Grafik.....	49
Şekil 18 TGE Grafik	49
Şekil 19 EGE Grafik	50
Şekil 20 Makine Dağılımı	52
Şekil 21 Öğrenme katsayısı ve momentum katsayısı için yapılan deneme sonuçları.....	53
Şekil 22 Çevrim Katsayısının Belirlenmesi	54
Şekil 23 Gizli Hücre ve Nöron Sayısının Belirlenmesi	56
Şekil 24 YSA Gerçekleşen ve Tahmin Grafiği	58
Şekil 25 YSA Ağ Modeli	59
Şekil 26 Regresyon Gerçekleşen ve Tahmin Grafiği.....	61
Şekil 27 DVM Gerçekleşen ve Tahmin Grafiği.....	62
Şekil 28 Tahmin Sonuç Grafiği	66

KISALTMALAR

DVM: Destek Vektör Makinaları

EGE: Ekonomi Güven Endeksi

GSYH: Gayri Safi Yurtiçi Hasıla

HMDB: Hane Halkı Maddi Durum Beklentisi

KKD: Veri Tabanlarında Bilgi Keşfi (Knowledge discovery in databases)

MAE: Ortalama Mutlak Hata (Mean Absolute Error)

MAPE: Ortalama Murlak Hata (Mean Absolute Percentage Error)

MBE: Ortalama Sapma Hatası

MLP: Çok Katmanlı Algılayıcı (Multi Layer Perceptron)

MP: Yapay Sinir Ağları

RKGE: Reel Kesim Güven Endeksi

RMSE: Kök Ortalama Kare Hata (Root Mean Squared Error)

SM: Satış Miktarı

SQL: Yapılandırılmış Sorgu Dili

SRM: Yapısal Risk Minimizasyonu

SÜE: Sanayi Üretim Endeksi

TCMB: Türkiye Cumhuriyeti Merkez Bankası

TGE: Tüketici Güven Endeksi

TH: Tüketim Harcamaları

TÜİK: Türkiye İstatistik Kurumu

ÜFE: Üretici Fiyat Endeksi

VM: Veri Madenciliği

WEKA: Waikato Environment for Knowledge Analysis

YH: Yatırım Harcamaları

GİRİŞ

Günümüze kadar gelen süreçte birçok teknolojik gelişme yaşanmış ve bu teknolojik gelişmeler hayatımızın çoğunda bizlere yardımcı olacak araçları beraberinde getirmiştir. İnsanoğlunun geleceği görme isteğine de yeri geldiğinde bu teknolojik gelişmeler katkıda bulunmuştur. Ham bir veri setine bakıldığında, karmaşık sayı dizilerinden başka bir şey görmemiz pek de mümkün olmamaktadır. Gelişen teknoloji ve istatistiksel yaklaşımlar ile ham veri setleriyle önemli sonuçlar elde edecek yaklaşımlar geliştirilmiştir. Veri madenciliğinde bilgi keşfi süreci sonrasında artık ham data yığınlarından anlamlı bilgiler elde etmek mümkün olmaktadır. Çalışmamızın amacı, sahip olduğumuz verilerden anlamlı sonuçlar elde etmektir. Bunun için veri madenciliği yaklaşımları ele alınarak uygun analizlerle anlamlı sonuçlar ortaya çıkarılmaya çalışılacaktır.

Tez çalışmasının birinci bölümünde, Veri Madenciliği kavramı açıklanmıştır. Tarihçe ve tanımlayıcı modeller hakkında bilgi verilmiştir.

Çalışmanın ikinci bölümünde, literatür taraması yapılmıştır. Satış veya talep tahminini içeren benzer çalışmalar taranmıştır. Toplam yirmi üç adet çalışma incelenmiştir. Çalışmalarda kullanılan yöntem ve değişkenler detaylı olarak açıklanmıştır. Veri Madenciliği içerisinde yer alan tahmin edici modeller açıklanmıştır. Tez çalışmasında kullanılan Yapay Sinir Ağları, Destek Vektör Makinaları ve Çoklu Doğrusal Regresyon yöntemlerine ait detaylı bilgi verilmiştir.

Çalışmanın üçüncü bölümü; Weka program tanıtımı, veri seti alınan firmanın tanıtımı ve uygulama aşamalarını içermektedir.

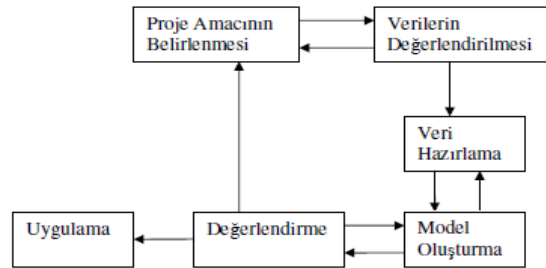
BİRİNCİ BÖLÜM

VERİ MADENCİLİĞİ KAVRAMI

1. Veri Madenciliği

Veri; henüz anlamlı bir bilgi biçimine dönüşmemiş veri yığınlarını ifade eder. Araştırmacıyı sonuca götürmek için işlenecek olan kayıtlar topluluğudur. Veri madenciliği, (data mining) ham bir datadan kullanışlı bilgileri çıkarmak için birçok yöntem kullanan bir disiplindir. Çok büyük veri setlerindeki yapının açıklanabilmesi için çok sayıda veri analiz aracını kullanır (Oğuzlar, 2004: 4). Verilerdeki işlenebilir, yararlı ve anlaşılabilir potansiyel örüntüleri tanımlamak için kullanılan önemli bir yöntem olarak ifade edilmektedir. Veri madenciliği (VM) kavramı verinin anlamını yeniden şekillendirmiş bir bilgi keşfi sürecidir. VM veriden bilgiye ulaşma sürecinde araştırmacıya yol gösteren bir araçtır. Sürecin çeşitli aşamalarında kullanılmak üzere çok sayıda araç ve teknik geliştirilmiş ve geliştirilmeye devam etmektedir. VM bu noktada birçok disipline dayanmaktadır.

Larose (2005:3) çalışmasında, veri madenciliği sürecini altı adımda açıklamıştır. Şekil 1.1.'de belirtilmiş olan aşamalar sırasıyla; uygulaması yapılacak çalışmanın amacının belirlenmesi, veri değerlendirilmesinin yapılması, veri setinin hazırlanması, veri setine uygun modelin oluşumu, modelin değerlendirilmesi ve uygulamanın yapılmasıdır.



Şekil 1 VM Uygulama Şeması

Kaynak: Larose (2005:3)

Geleneksel yaklaşımın temellerini ortaya çıkaran analizlerin amacı istatistiksel veri özelliklerini ortaya çıkarmak ve sayısal çıkarımlar yapmaktır. Bu amaçla makine öğrenmesi, istatistik ve örüntü tanıma gibi yöntemlerin karışımıyla oluşturulmuş bir disiplini ifade eder.

Veri kümesinin büyüklüğü, standart istatistiksel uygulamalar ve veri madenciliği arasındaki en temel farklılığı oluşturmaktadır. İstatistikle uğraşan biri için birkaç yüz veya bin ile ifade edilen veriler büyük bir veri setini oluşturmaktadır. Ancak veri madenciliğinde, büyük veri kavramı milyon veya milyarlar ile ifade edilmektedir. (Oğuzlar, 2003:69).

2. Veri Madenciliğinin Tarihsel Gelişimi

Veri madenciliğinin (VM) süreci 1950’li yıllardan başlayarak günümüze kadar gelişerek devam etmiştir. 1960 ve öncesinde temel düzeyde dosya işleme yapılmaktadır. Veri madenciliği birden fazla disiplinin birleşerek bir blok oluşturması ile oluşmuştur. Bu bloğun içerisinde ilk zamanlarda veri tabanı yönetim sistemleri, yapay zekâ sistemleri ve makine öğrenmesi yer almaktaydı. (Venkatadri & Reddy, 2011: 19 – 22).

Tablo 1 Veri Madenciliği Tarihçesi

1950-1960	İlk bilgisayarlar (sayım için)
1960-1970	Veri tabanı ve verilerin depolanması
	Perseptronlar (Algılayıcı, Fark edici)
1970-1980	İlişkisel veri tabanı yönetim sistemleri
	Basit kurallara dayanan uzman sistemler ve makine öğrenimi
1980-1990	Büyük miktarda veri içeren veri tabanları SQL sorgu dili
1990-2000	Veri tabanlarında bilgi keşfi çalışma grubu ve Sonuç bildirgesi
	Veri madenciliği için ilk yazılım
2000-...	Tüm alanlar için veri madenciliği uygulamaları

1970 ve 1980’li yıllarda, hiyerarşik ve ağ veri tabanı sistemleri, ilişkisel veri tabanı sistemleri, veri modelleme araçları, sorgu dilleri: SQL, vb., online işlemsel veri işleme ve kullanıcı ara yüzleri gibi yenilikler dahil edilmiştir.

Veri madenciliği terimi ise bilgisayar bilimi alanı içerisinde ilk olarak 1980’lerde bir disiplin olarak adlandırılmıştır. 1990’ların başlarında veri madenciliği süreç olarak veri tabanlarından bilgi keşfi (knowledge discovery in databases – KDD) olarak adlandırılan sürecin bir alt dalı olarak ortaya çıkmıştır (Coenen, 2004: 1 – 24).

1990’lı yıllardan günümüze kadar olan kısımda veri madenciliği yazılımları, bilgi alışverişi, veri ve bilgi entegrasyonu gibi konular önem kazanmıştır.

3. Veri Madenciliği Kullanım Alanları

Teknolojik ve bilimsel yenilikleri akademik tezlerde görmemiz mümkün olmaktadır. Son yıllarda veri madenciliği teknik ve açıdan çözülmeye çalışan problemlere yanıt bulduğu için akademik çalışmalarda aktif olarak kullanılmaya başlanmıştır. Lisansüstü tezlerin, çeşitli alanlarda güncel konular ve gelişmeler üzerine yapıldığı varsayımı belirtilebilir. Dolayısıyla, seçilmiş bir alanda çalışma gerçekleştirilen lisansüstü tezlerinin detaylı bir biçimde incelenmesi, söz konusu alandaki seçimlerin değerlendirilmesi, değişen tekniklerinin incelenmesi ve uygulama alanlarının belirlenmesi açısından önem oluşturmaktadır.

Veri madenciliği oldukça yeni bir disiplin olmasına karşın geniş bir kullanım alanına sahiptir. Bunun en büyük nedeni de günümüzde özel sektör ve bilimsel araştırmalarda çalışılması gereken veri miktarının çok fazla olmasıdır. (Durmaz & Kocamış, 2008:4)

Çeşitli amaçlarla VM uygulamaları görülebilmektedir. Bunlardan bazıları şu şekildedir:

- a) Karar verme konularında yardımcı olmak ve veri tabanı için destek sağlamak,
- b) Pazar araştırması: uygun pazarın belirlenmesi, müşteriler arası ortak noktaların tespit edilmesi, sepet analizi ve çapraz pazar incelemesi,
- c) Risk analizi: öngörü geliştirmek, kalite kontroller sonucu sahtekarlıkları saptamak ve önlemek, rekabet analizi geliştirmek,
- d) Kredi risk arařtırmaları,
- e) Kaynakların olabilecek en iyi düzeyde kullanımını sağlamak,
- f) Geçmiş veriler ile mevcut yapının analizi sonuçlarıyla gelecek hakkında çıkarımlarda bulunmak amacıyla kullanılmaktadır.

Gorunescu (2011) yaptığı çalışmada, VM çalışmalarını değerlendirilmiş ve kullanım alanlarını Tablo 2'deki şekilde özetlemiştir.

Tablo 2 Veri Madenciliği Kullanım Alanları

Kullanım Alanları	Kullanım Oranları (%)
CRM/Müşteri Analitiği	32.8
Bankacılık	24.4
Direk Pazarlama	16.1
Kredi Puanlama	15.6
Telekomünikasyon	14.4
Dolandırıcılık Tespiti	13.9
Satış	11.7
Sağlık	11.7
Finans	11.1
Bilim	10.6
Reklamcılık	10.6
E-Ticaret	10.0
Sigortacılık	10.0
Web Madenciliği	8.3
Sosyal Ağlar	7.8
İlaç	7.8
Biyoteknoloji	7.8

Kaynak: Gorunescu, 2011:41

4. Veri Madenciliğinin Kuvvetli ve Zayıf Yönleri

Teknolojik ve bilimsel yenilikleri akademik tezlerde görmemiz mümkün olmaktadır. Son yıllarda veri madenciliği teknik ve açıdan çözülmeye çalışan problemlere yanıt bulduğu için akademik çalışmalarda VM aktif olarak kullanılmaya başlanmıştır. VM'nin kuvvetli ve zayıf yönlerini değerlendirdiğimizde;

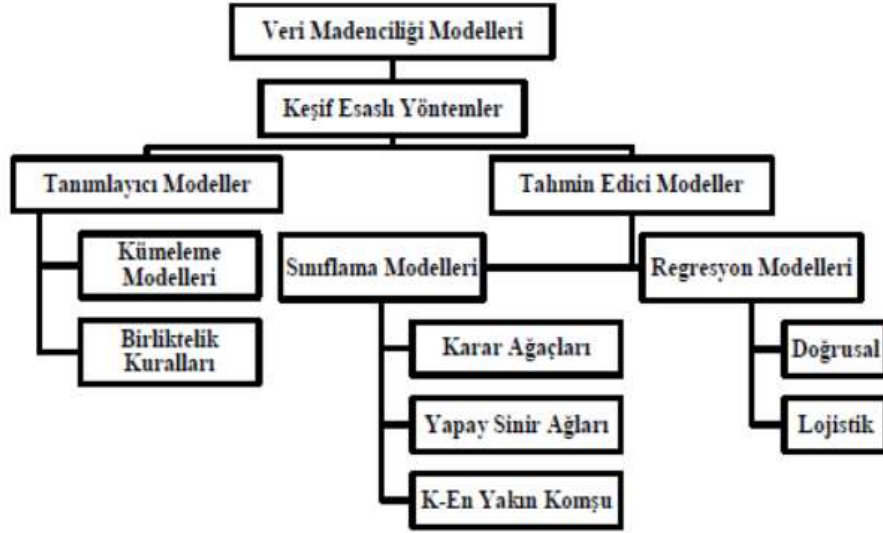
- i.** VM teknikleri geniş açılı sorunların çözümü için uygun bir araç olması,
- ii.** VM teknikleri karmaşık durumların çözümlenmesinde kullanılmak için uygun olması,
- iii.** VM araçlarının hem sayısal hem de kategorik değişkenlerin çözümünde kullanılması kuvvetli yönlerini oluşturmaktadır.

Veri madenciliği zayıf yönleri ise şunlardır:

- i.** VM tekniklerinde 0-1 arasında giriş verisi olması gereklidir.
- ii.** VM teknikleri ürettikleri sonuçların açıklamasını sunamazlar.
- iii.** VM tekniklerinde elde edilen sonucun en iyi sonuç olma garantisi yoktur. Sonuçlar farklı teknikler ile tekrarlanarak en iyi sonuç seçilebilir.

5. Veri Madenciliği Modelleri

VM, büyük verileri setlerini işleyerek bu verilerden anlamlı bir bilgiyi ortaya çıkarma ve keşif sürecidir. Bu süreçte kullanılması gereken çeşitli modeller bulunmaktadır. Kullanım amacına göre en uygun modeli belirleyebilmek için çeşitli algoritmalar kullanılır. Bu modeller; “Tahmin Edici (Predictive)” ve “Tanımlayıcı (Descriptive)” olmak üzere iki başlık altında incelenebilir.



Şekil 2 Veri Madenciliği Modelleri

5.1. Tanımlayıcı Modeller

Tanımlayıcı modellerin temel amacı bir sonuç tahmini yapmaktan ziyade, veri grubu içerisinde bulunan veriler arasındaki bağlantıları, ilişkileri ve davranışları belirlemektir. Tanımlayıcı modeller analiste daha önceden bir hipoteze sahip olmaksızın, veri kümesinin içinde ne tür ilişkiler olduğunu anlama imkânı sunar. Analizcinin çok geniş veri tabanlarındaki bilgileri incelemek, örüntüleri keşfetmek için doğru soruları sorup hipotezler geliştirmesi pratikte zor olduğundan, ilginç örüntüleri keşfetme inisiyatifi veri madenciliği programına bırakılır. Keşfedilen bilginin kalitesi ve zenginliği, uygulamanın kullanılabilirliğini ve gücünü oluşturur (Güvenç 2001:120).

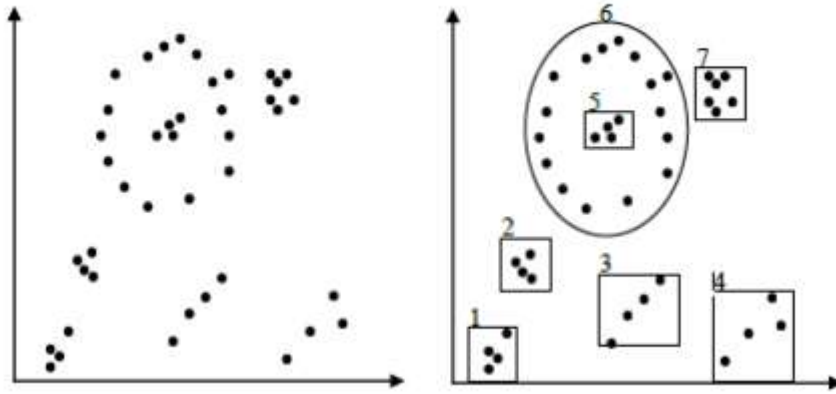
Tanımlayıcı modellerde genellikle kullanılan yöntemler aşağıdaki gibidir:

- Kümeleme yöntemi
- Birliktelik Kuralı

5.1.1. Kümeleme Yöntemi

Kümeleme yöntemi, veriyi gruplara veya sınıflara ayırma işlemidir. Kümeleme yönteminde objenin sınıf etkisi bilinmemektedir.

Kümeleme analizinde, her veri var olan kümeler ile kıyaslanır ve kendi özelliklerini taşıdığı düşünülen kümeye taşınır. Taşınan bu veri, kümesinin tamamlayıcı değerini alır. Var olan tüm kayıtlar optimum biçimde atanana kadar yapılan bu işlem tekrarlanır. Karmaşık ve büyük veri setleri gruplanırken, bu veri setleri ayrılarak benzer gruplar ile bir arada ele alınmaktadır. Bu şekilde yapılan çalışmada, heterojen yapıya sahip veri setlerinin homojen yapıya dönüştürülmesi hedeflenmektedir. Uzman kişiler tarafından, oluşan kümelerin hangi anlamları taşıdığı belirlenmektedir (Çelik, 2009: 37-38).



Şekil 3 Kümeleme Örneği

Şekil 3’de heterojen yapıya sahip verilerin homojen bir yapı altında gruplandıkları görülmektedir. Benzer özelliklere sahip verilerin gruplandırılmaları sonucunda kümeleme işlemi gerçekleşmektedir.

Kümeleme yöntemine ilişkin genel özellikler şu şekildedir;

- Küme sayısı bilinmemektedir.
- Kümelere ilişkin net bir bilgi olmayabilir.
- Kümeleme sonuçları dinamiktir.

Önceden belirlenmiş sınıfların bulunmaması kümelemedeki öğrenmenin denetimsiz öğrenme olmasının nedenidir. Bu durum sınıflama yönteminin kümeleme yönteminden ayrılmasının sebebidir. sınıflar belli iken yapılan öğrenmeye denetimli öğrenme, herhangi bir sınıflama yapılmadan yapılan öğrenmeye ise denetimsiz öğrenme adı verilir. Örneğin veri tabanlarındaki kayırların her biri gruplara ayrılmış olduğunda (kadın veya erkek) bu kural çıkarma işlemi denetimli öğrenme olarak adlandırılmaktadır. Ancak herhangi bir kaydın yanında kadın veya erkek olduğu bilgisi yok ise bu durum denetimsiz öğrenme olarak ifade edilmektedir. Bu işlem aynı zamanda kümeleme işlemidir. Bu işlemde kadın/erkek gibi bir etiket ya da sınıf olmayacağı için, kümeleme kayıtlar arasındaki benzerlik ölçütüne göre yapılmaktadır (Silahtaroglu, 2013:60).

5.1.2. Birliktelik Kuralı

Veri içerisindeki birliktelik özelliklerini açıklayan örüntüleri ortaya çıkarmak için kullanılan bir analiz yöntemidir. Ortaya çıkarılan bu örüntüler alt grupları veya özel çıkarılmış kuralları ifade edebilir. Birliktelik analizinin amacı önemli örüntülerin ortaya çıkarılmasını sağlamaktır (Dunham, 2003:8).

Birliktelik kuralları, bir arada olan olayların ya da o olaya ait özelliklerin belirlenme işlemidir. İlişki veya pazar sepet analizi olarak isimlendirilmektedir. Birliktelik kuralları genellikle birbirini etkileyen olayları içermektedir. Bir olayın olabilmesi başka bir olayın oluşuna bağlanmaktadır. Birliktelik kuralları içerisinde Apriori ve GRI en çok tercih edilen algoritmalarıdır.

Ortaya konan bir stratejiye göre; birbiriyle bağlantılı ürünler satışlarının artması amacıyla yan yana dizilmektedir. Örneğin bilgisayar almak için gelen bir müşteri aynı zamanda bir yazılım satın alma eğilimi gösterebilir. Bu sebeple bilgisayar ve yazılımlar yan yana sergilenmektedir. Alternatif bir strateji olarak ise; birlikte alma eğilimi gösteren ürünleri farklı iki uca koyarak müşterinin birini aldıktan sonra diğerine gitmek için yol boyunca başka ürünleri satın alarak satışlarının artmasına sebep olabilir. Örneğin; bilgisayar satın almak isteyen bir müşteri satın alma kararı verdikten sonra yazılım ürünlerinin olduğu tarafa giderken yol üzerinde güvenlik sistemi için satılan ürünleri görerek almak isteyebilmektedir. Birliktelik kuralları içerisinde yer alan analizi hangi ürünler için iskonto yapılacağını belirlemek konusunda yardımcı olmaktadır. Eğer müşterilerin yazıcı ve bilgisayarı birlikte satın alma gibi bir eğilimi varsa yazıcı da yapılan bir indirim bilgisayar satışlarını da etkileyerek satın almaya teşvik edecektir (Han ve Kamber, 2001:227).

Bu kısma kadar olan alanda veri madenciliği hakkında bilgi verilmiş ve tanımlayıcı modellerden bahsedilmiştir. Tez çalışmasında tahmin edici modeller kullanılmıştır. Bu sebeple tahmin edici modeller bir sonraki bölüm altında detaylı olarak açıklanmıştır.

İKİNCİ BÖLÜM

TAHMİN EDİCİ MODELLER

Tahmin, geçmiş tecrübe veya verilerden elde edilen bilgilerin gelecek ön görüşü için kullanılmasına denir. Karar alma süreçlerinde tahmin edici modeller önemli bir rol oynamaktadır. Tahmin edici modellerde var olan verilerden hareket edilerek bir model geliştirilmesi hedeflenir. Oluşturulan bu model kullanılarak sonuçları bilinmeyen veri yığınları için geçerli bir sonuç ortaya çıkarması amaçlanmaktadır.

Uygulamada kullanılan başlıca tahmin edici teknikler şunlardır:

1. Yapay Sinir Ağları
2. Destek vektör makineleri
3. Regresyon Analizi

Uygulama içinde kullanılan üç model haricinde, tahmin edici modeller içerisinde yer alan ve uygulamaya dahil edilmeyen ancak en çok kullanımı sağlanan diğer modeller hakkında da kısaca bilgiler verilmiştir.

Tahmin edici modeller sınıflama ve regresyon modellerinden oluşmaktadır. Sınıflama içerisinde karar ağaçları, Yapay Sinir Ağları ve k-en yakın komşu algoritmaları yer almaktadır. Regresyon ise doğrusal ve lojistik regresyon olmak üzere ikiye ayrılmaktadır. Çalışmada tahmin edici modellerden Yapay Sinir Ağları, doğrusal regresyon ve destek vektör makineleri uygulamaya dahil edilmiştir. Sınıflandırma ve regresyon olarak ayrılan tahmin edici modellerden burada kısaca bahsedilmiştir. Uygulamada yer alan diğer modeller detaylı bir biçimde bir sonraki bölümde sunulmuştur. Tahmin edici modeller ile ilgili bilgi verilmeden önce, bu modeller ile ilgili yapılan literatür taramaları paylaşılacaktır.

1. Literatür Taraması

Bir çalışmaya başlamadan önce, yapılan diğer çalışmaların taramasının yapılması önem arz etmektedir. Yapılacak çalışmanın literatüre uygun olması ve daha önce yapılan çalışmalarla desteklenmesi gerekmektedir. Satış tahminleri ile ilgili literatür taramaları yapıldığında, veri madenciliği yöntemlerinin sıklıkla kullanıldığı görülmüştür. Talep tahmininin aynı zamanda satış tahminini ifade etmesi sebebi ile, araştırılan çalışmalarda “talep tahmini” ve “satış tahmini” olarak değerlendirilmiştir. Ürün satışını veya talebini belirleyen faktörler temelinde yapılan çalışmalar incelenmiştir. Değişkenlerin belirlenebilmesi amacıyla literatür taraması gerçekleştirilmiştir.

Yılmaz, Hilal (2018), aylık konut satışının tahminini gerçekleştirmiştir. Çalışmada Yapay Sinir Ağları ve regresyon yöntemi kullanılmıştır. Veri seti 2013-2017 yılları arasında aylık olarak alınmıştır. Değişken olarak; konut fiyatı, kredi durumları, nüfus sayısı, GSYİH, işsizlik, ailedeki kişi sayısı baz alınmıştır. Yapılan analizin sonucunda tahmin edilen değerler 2017 gerçek veri seti ile karşılaştırılmıştır. YSA modelinin gerçeğe daha yakın tahminler ürettiği görülmüştür.

Pavlyshenko (2019), çalışmasında Rossmann mağazalarının satış talebini incelemiştir. Çalışma R programı ve Jupyter Notebook kullanılarak gerçekleştirilmiştir. Çalışmada kullanılan modeller; Yapay Sinir Ağları, Rastgele Orman Algoritması ve Ağaç modelleridir. Analizin sonucunda satış tahmini yapılırken regresyon modellerinin zaman serisi modellerinden daha doğru sonuç verdiği ortaya konmuştur.

Efendigil vd. (2009), beyaz eşya tedarik zinciri talep tahmini gerçekleştirmiştir. 24 aylık veri her perakendeciden anket yöntemiyle toplanmıştır. 79 eğitim ve 17 data test verisi olarak bölünmüştür. Değişken olarak; ürün fiyatı, ürün kalitesi ve promosyonları kullanmıştır. Yöntem olarak Bulanık Mantık ve Yapay Sinir Ağları kullanılmıştır.

Çalışmanın sonucunda Bulanık Mantık uygulaması içinde test edilen Gauss üyelik fonksiyonunun daha etkin bir çözüm sunduğu ortaya konulmuştur.

Rincon-Patino & Lasso & Corrales (2018), avokado satışlarını incelemişlerdir. Hava koşullarının etkileri dikkate alınarak gerçekleştirilen bu çalışmada, tarım sektörüne ilişkin veriler makine öğrenimi teknikleriyle analiz edilmiştir. Çalışmada; Doğrusal Regresyon, Yapay Sinir Ağları içerisinde yer alan çok katmanlı algılayıcı (MLP), Destek Vektör Makineleri ve Çok Değişkenli Regresyon modeli kullanılmıştır. Sonuçlar değerlendirildiğinde, destek vektör makineleri ve çok değişkenli regresyon modellerinin daha iyi sonuç verdiği görülmüştür.

Koochakpour & Tarokh (2016), bir işletmenin satış tahminini gerçekleştirmiştir. Değişken olarak; milli gelir, döviz kuru, enflasyon ve geçmiş satış verilerini kullanmıştır. Tahmin yöntemi olarak Bulanık Mantık ve Yapay Sinir Ağları tercih edilmiştir. Çalışmanın sonucunda Bulanık Mantık modellerinin daha iyi sonuç verdiği ortaya konmaktadır.

Emre TÜRK & Farzad KIANI (2019), beyaz eşya sektöründe satış tahmini gerçekleştirmiştir. Beyaz eşya satış miktarı (SM) bağımlı değişken olarak kullanılırken, tüketici güven endeksi (TGE), reel kesim güven endeksi (RKGE), gayri safi yurt içi hasıla (GSYH), sanayi üretim endeksi (SÜE), tüketim harcamaları (TH), hane halkı maddi durum beklentisi (HMDB), ekonomi güven endeksi (EGE) ve evlenme istatistikleri de bağımsız değişken olarak kullanılmıştır. Tahmin yöntemi olarak YSA ve regresyon kullanılmıştır. Yapılan çalışmanın sonucunda YSA modellerinin, regresyon analizlerine göre daha başarılı olduğu ortaya konmuştur. Tahminler ile gerçek değerler karşılaştırıldığında, bulunan sonuçların gerçek değerlere yakın olduğu gözlemlenmiştir.

Hicham & Bouhorma (2012), ambalaj satış tahmini gerçekleştirmiştir. Değişken olarak; imalat tüketici endeksi, rekabet endeksi ve geçmiş satış verileri alınmıştır. Delphi ve ysa ile tahmin gerçekleştirilmiştir. Sonuçlar MAPE ve RMSE oranlarına göre değerlendirilmiştir. Delphi yönteminin YSA yöntemine göre daha başarılı sonuçlar koyduğu savunulmaktadır.

Azadeh vd. (2013), demir çelik sektöründe talep tahmini gerçekleştirmiştir. Değişken olarak; milli gelir, büyüme oranı, enflasyon ve çelik üretim verileri kullanılmıştır. Bulanık Mantık ve YSA ile tahmin yapılmıştır. Çalışmada ortaya konan mükemmellik modelinin esnek olduğunu belirterek hem küçük hem de büyük şirketlerde uygulanabileceğini öne sürmüştür.

Ecemiş (2018), paslanmaz çelik satış tahmini gerçekleştirmiştir. Değişken olarak; hammadde fiyatları, USD/TRY, ÜFE ve sanayi üretim endeksi alınmıştır. Model Yapay Sinir Ağları ve Destek Vektör Makinaları ile test edilmiştir. Sonuçlar MAPE, MAE ve RMSE oranlarına göre kıyaslanmıştır. Yapay Sinir Ağları yönteminin, makine sektörü haricinde kalan tüm sektörlerde destek vektör makineleri yöntemine göre daha başarılı sonuçlar elde ettiği ortaya konmuştur.

Karaatlı Meltem & Helcacioğlu Ceyda & Ömürbek Nuri & Toköz Gönül (2012), Yapay Sinir Ağları kullanarak otomobil satış tahmini gerçekleştirmiştir. Veri seti olarak Ocak 2007, Haziran 2011 yılları arasındaki veriler aylık olarak alınmıştır. Değişken olarak; gayri safi yurt içi hasıla (GSYH), reel kesim güven endeksi (RKGE), yatırım harcamaları (YH), tüketim harcamaları (TH), tüketici güven endeksi (TGE), döviz kuru ve bağımsız değişkenler ve satılan toplam otomobil sayısı bağımlı değişken olarak alınmıştır. Modelin Yapay Sinir Ağları ile tahmini gerçekleştirilmiştir. MAPE değeri baz alınarak çalışmanın doğruluğu analiz edilmiştir. Gerçek değerler ile tahmin değerleri kıyaslandığında, yapılan tahminin başarılı olduğu ortaya konulmuştur.

Aksel (2000), İstanbul ve Türkiye geneli için uzun dönem elektrik talep tahmini gerçekleştirmiştir. 2001, 2003 ve 2005 yıllarına ait puan yük enerji tüketim değerleri tahmin edilmiştir. Tahmin yöntemi olarak; çok değişkenli regresyon analizi ve geriye yayılım algoritmaları, iki gizli katmanlı yapay sinir ağı yöntemini kullanmıştır. Değişken olarak; nüfus, kişi başına düşen gayri safi milli hâsıla, gelişim hızı, sanayi üretim endeksi, petrol varil fiyatı ve elektrik enerjisi birim satış fiyatı kullanılmıştır. Kullanılan çok değişkenli regresyon analizi ve geriye doğru yayılım algoritması sonucu ağırlıklı ortalama yöntemiyle tek sonuca dönüştürülmüştür.

Demirel (2009), yaptığı tez çalışmasıyla elektrik enerjisi talep tahmini gerçekleştirmiştir. Değişken olarak; Gayri Safi Milli Hâsıla, üretilen enerji, tüketilen enerji, nüfus ve kurulu güç verilerini kullanılmıştır. Tahmin edilen sonuçları ilgili kurumun ön görüşleriyle karşılaştırmışlardır.

Catal, Ece & Arslan & Akbulut (2019), satış tahmini gerçekleştirmiştir. Araştırmada, klasik zaman serisi yöntemleri makine öğrenmesi algoritmalarıyla karşılaştırılmıştır. Çalışmada; Bayes, Yapay Sinir Ağları, Karar Ağacı Yöntemleri kullanılmıştır. Ortalama kare hata ve kök ortalama kare hata sonuçları karşılaştırılarak tahmin güvenilirlikleri incelenmiştir. Tahmin başlangıcında karar ağacı regresyon modeli en düşük hatayı verirken tahmin sonlarında arttırılmış karar ağacının en düşük hatayı verdiği tespit edilmiştir. Bu durumda çalışmada, regresyon modellerinin klasik zaman serisi modellerinden daha iyi tahminler ürettiği varsayımından hareket edilmiştir.

Tozak, E. Yedigâr (2021), veri madenciliği programlarını kullanarak bir tekstil firmasının satış verilerini tahmin etmiştir. Veriler yaklaşık dört yıllık olarak aylık bazda alınmıştır. Veriler karar ağacı kullanılarak tahmin edilmiştir. Analize dâhil edilen değişkenler; müşteri sayısı, üretim yeri, fabrika, yıl, ay, sevk adeti, birim fiyattır. Elde edilen tahmin sonuçlarına göre, firmanın yıllık üretim kapasitesinin ne kadar olması

gerektiđi saptanmıřtır. Bu durumda üretici için dođru bir planlama yapılması hedeflenmiřtir.

Gavcar vd. (1999), alıřmalarında Türkiye'deki kâđıt ve karton talebi tahmin edilmiřtir. Analizde deđiřken olarak; kâđıt ürünleri, ithalat ve ihracat miktarları, eřya fiyat endeksi, gayrisafi milli hasıla ve nüfus verileri kullanılmıřtır. alıřmada oklu regresyon analizi ile tahmin gerekleřtirilmiřtir. Modelde yer alan deđiřkenlerin iliřkisi korelasyon analiziyle incelenmiřtir. alıřma sonucunda 21. yüzyılda talep edilecek kâđıt miktarına ait analiz gerekleřtirilmiřtir.

Carlson & Umble, (1980), Amerika'da, beř farklı segmentte bulunan otomobillerin (standart ve lüks otomobil segmentinde) beř senelik talep tahminini regresyon analiziyle gerekleřtirmiřlerdir. Deđiřken olarak; benzin fiyatı, benzin kıtlıđının piyasaya etkisi, tařıt fiyatları, tüketici gelirleri ve sektördeki iřçi grev sayıları alınmıřtır.

Kılı, Fatih vd. (2018), yemekhane talebini yapay zekâ tekniklerini kullanarak tahmin etmiřtir. alıřmanın amacı, yemek miktarının tahmini karřılayabilecek řekilde üretilmesini sađlamaktır. Üniversitelerdeki talep miktarı deđiřkenlik göstermektedir. Bu sebeple yapılacak talep tahmini ile sorunun özümü hedeflenmektedir. Deđiřken olarak; tüketici tipleri, yemekhaneden faydalanan öđrenci sayıları, yemekhaneden faydalanan personel sayıları alınmıřtır. Teknik olarak Yapay Sinir Ađı, Destek Vektör Makineleri ve Lineer Regresyon kullanılmıřtır. Tahmin sonuçları deđerlendirildiđinde Destek Vektör Makinelerinin bu analiz için daha iyi sonuç verdiđi ortaya konmaktadır.

Demirbař, F. Pınar (2011), kombi satıř adetlerini tahmin etmiřtir. alıřmanın amacı talep tahminin belirlenerek müřteri kaybı riskini azaltmaktır. alıřmada tahmin zaman serisi analizi yöntemiyle gerekleřtirilmiřtir. Deđiřken olarak 2008-2010 yılları

arası satış verileri kullanılmıştır. Araştırma sonuçlarına göre en başarılı tahmini en küçük kareler tahmincilerinin gerçekleştirdiği ortaya konmaktadır.

Yücesoy, Mihriban (2011), kâğıt talebi tahmini gerçekleştirmiştir. Yapay Sinir Ağları ve Çoklu Doğrusal Regresyon modelleri uygulamada kullanılmıştır. Değişken olarak; üretim miktarı, ithalat miktarı, kâğıt basım sanayi toptan fiyat endeksi, GSYH, nüfus bilgileri, okuryazar oranı kullanılmıştır. Yapılan tahmin sonucunda YSA modeliyle %98,89 başarı oranı elde edilirken, regresyon modelindeki başarı oranı %61,89 olarak bulunmuştur.

Yeğen, Neslihan (2020), perakende alanında faaliyet gösteren bir firmanın satış verilerinin tahminini gerçekleştirmiştir. Modelde zaman serisi analizi yöntemi kullanılmıştır. Sonuçlar MAPE oranına göre kıyaslanarak en az hata oranına sahip bulgu değerlendirilmiştir. Modelde eksik değişkenin olduğu tespit edilmiş, ürün çeşidi değişkeninin modele eklenmesi gerektiği sonucuna ulaşılmıştır. Çalışma sonucunda, zaman serisi modelinin daha anlamlı sonuçlar elde ettiği görülmüştür.

Poyraz, İlker (2020), eczaneler için ilaç talep tahmini gerçekleştirmiştir. Yöntem olarak makine öğrenmesi kullanılmıştır. Çalışmanın amacı sonraki yılların satış miktarının tahmin edilmesiyle stok yönetiminin sağlanabilmesidir. Makine öğrenme algoritmalarından; Linear Regresyon, Gaussian Process, M5Rules, Yapay Sinir Ağları içerisinde yer alan çok katmanlı algılayıcı, SMOreg, M5P, Random FOREST kullanılmıştır. Veri seti 2015-2019 yılları arasında haftalık olarak alınmıştır. Sonuçlar MAPE değerine göre karşılaştırılmıştır. M5Rules en düşük hata oranına sahip algoritma olarak bulunmuştur. Gaussian Processes algoritması ise üç hafta boyunca en iyi sonucu veren algoritma olarak bulunmuştur.

Kargı A., V. Sinem (2013), üretim sayılarındaki hata oranlarını düşürmek için tahmin gerçekleştirmiştir. Tahmin modeli olarak çoklu doğrusal regresyon ve Yapay Sinir Ağlarını kullanmıştır. Yapılan çalışma sonucunda Yapay Sinir Ağlarının Multi Layer Perceptron (MLP) ile yapmış olduğu tahminin %93 oranında başarılı olduğu görülmüştür. MLP yani Yapay Sinir Ağları tahmincisi çoklu doğrusal regresyona göre tüm hata oranlarında daha başarılı sonuç sergilemiştir. Firma üretimine başlayacağı ürünün hata oranını öncesinde tespit edebilecek ve gerekli önlemleri alabilecektir.

Karahan Mehmet (2011), Yapay Sinir Ağları kullanarak talep tahmini gerçekleştirmiştir. Modelde kullanılan ağ; bir gizli katmana sahip, ileri beslemeli, geri yayılım gösteren bir sinir ağıdır. Çalışmada 2004-2010 yılları arasındaki veriler kullanılmıştır. Oluşturulan modelde beş girdi değişkeni ve bir çıktı değişkeni yer almaktadır. Çıktı değişkeni tahmini gerçekleştirilmiş olan kuru kayısı talepleridir. Bağımsız değişkenler ise; geçmiş talep miktarları, tarih, kur bilgisi, aylık ortalama fiyat, Pazar sayısı, mevsim etkileri modele dahil edilmiştir. Ağın eğitimi için momentum ve öğrenme katsayısı değeri 0,1 olarak alınmıştır. Yapay Sinir Ağlarının doğruluk sonucu %86 olarak bulunmuştur. Bu durumda gerçekleşen veriler ile tahmin edilen adetler arasındaki sapma çok düşüktür ve tahmin edilen değerlerin gerçek değerlere yakın olduğu görülmektedir. Çalışmanın amacı, elde edilen tahmin bilgisine göre firmanın üretim planını bu doğrultuda önceden planlamasını sağlayabilmektir.

Ürün satışını veya talebini belirleyen faktörler temelinde yapılan çalışmalar incelenmiştir. Başta beyaz eşya sektörü olmak üzere farklı sektörlerde yapılan çalışmalara da yer verilmiştir. Yapılan çalışmalarda kullanılan, satışlara etki eden ekonomik faktörler olarak genellikle, ülke milli geliri, sanayi üretim endeksi, ekonomi güven endeksi, enflasyon, döviz kuru, sektörün üretimleri, ürün fiyatları ve geçmiş satış verileri olduğu görülebilir.

2. Sınıflama modelleri

Sınıflandırma, veriyi daha öncesinde belirlenen gruplarla eşleştirme sürecidir. Sınıflandırma işleminde önceden belirlenmiş kategorik bir değişken bulunur. VM modeli geniş veri gruplarını sorgular. Veri seti içerisinde yer alan her veri hedef değişken üzerinde bilgi içerir.

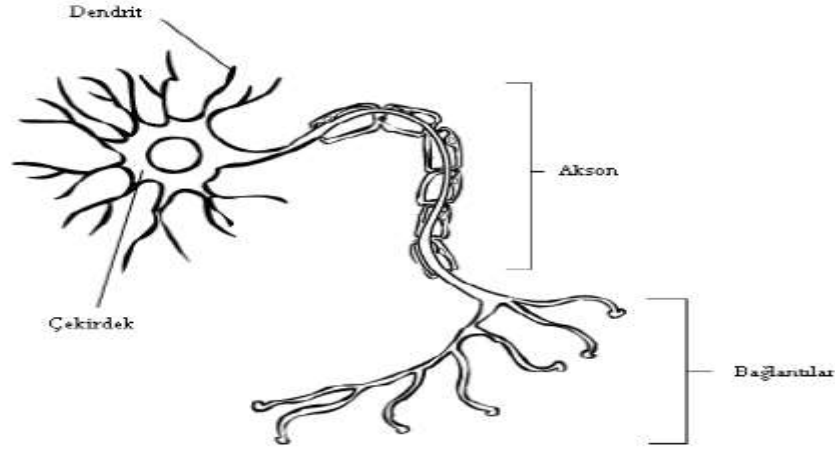
Sınıflandırma teknikleri veri madenciliğinde çok kullanılan yöntemlerdendir. Kullanıldığı yerlere örnek olarak aşağıdaki durumlar verilebilir.

- a. Bankacılık: Kredi kullanım durumlarında müşterinin risk düzeyinin tespitinde,
- b. Eğitim: Yeni öğrencileri özelliklerine göre sınıflara yerleştirmede ve buna göre eğitim programları hazırlamada,
- c. Sağlık: Hastalık teşhisinde,
- d. Hukuk: Karar verme konularında

3. Yapay Sinir Ağları

YSA, insan beyninin çalışma prensibinden etkilenilerek çözülmesi zor problemlerin çözümü oluşturulmuş yapılardır. Yapay Sinir Ağları (Artificial Neural Networks), sistemi öğrenerek hücreler arasında bilgi akışı gerçekleştirerek bilgisayar ortamında modellenmektedir. Yaklaşık 1011 sayıda nöron (sinir hücresi) ve birleşme noktalarından oluşan yapıya sahip olan insan beynidir. Yapay Sinir Ağları (YSA), insan beyni örnek alınarak kurgulanmış ve beynin işlemlerini yerine getirebilecek bir sistem olarak tasarlanmıştır. İnsani öğretilerin düşünme, öğrenme, hatırlama ve yorumlama gibi davranışlar ele alınarak hataları en aza indirme konusunda tasarlanmış bir yöntem olarak karşımıza çıkmaktadır. YSA modelini daha iyi anlayabilmek için biyolojik sinir sistemi hakkında az çok bilgi sahibi olmak gerekmektedir. Çünkü YSA modelleri oluşturulurken

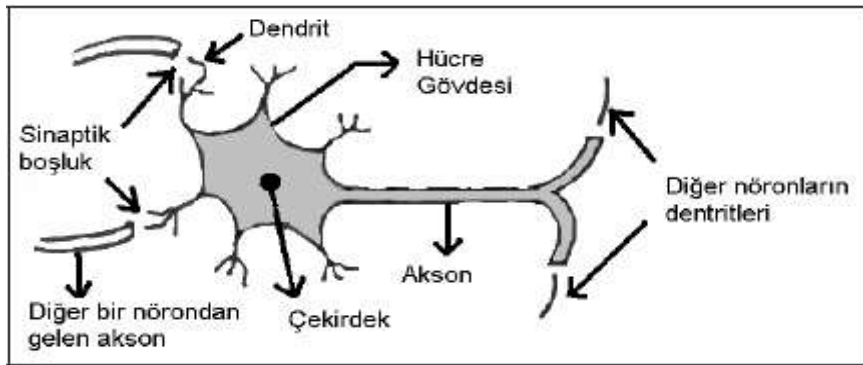
insan beyni yapısı düşünülerek oluşturulmuştur. Amaç ağ öğrenimini insan beyin yapısı gibi şekillendirmek ve ağın kendi kendine öğrenebilmesini sağlamaktır.



Şekil 4 Sinir Hücre Yapısı

Kaynak: Çayiroğlu 2015:4

Şekil 4'te sinir hücre yapısı görülmektedir. YSA prensibi beyin hücresinde bulunan sinir yapısından uyarlanmıştır. Sinir hücre yapılarının bağlantılar yoluyla aktarımı gerçekleşmektedir. YSA modellemesinde ise aktarım katmanlar arasında gerçekleşmektedir. YSA yapısında basit sinirler olarak işlem elemanları (düğüm) örnek gösterilebilir. YSA modeli birbiriyle ilişkili birden fazla bağlantılı düğümlerin bir araya gelmesiyle oluşmaktadır.

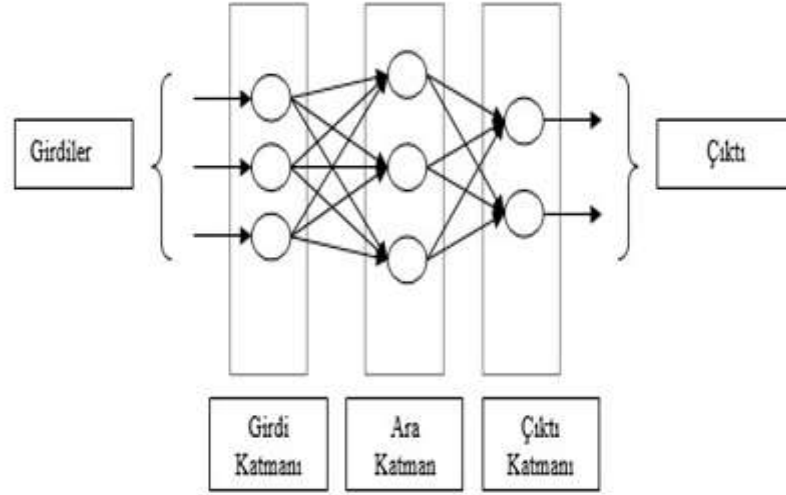


Şekil 5 Nöron Yapısı

Kaynak: Fausett, 1994:6

Nöronlar sinir sisteminin temelini oluşturmaktadır. Şekilde görülen dendrit giriş bilgilerini diğer nöronlardan almaktadır. Akson ise çıkış sinyallerini değerlendirmektedir.

Yapay Sinir Ağları iki alan üzerinde ayrılmaktadır. İleri beslemeli çok katmanlı ağlar ve geri beslemeli çok katmanlı ağlar olarak ayırım yapılabilir. YSA modelleri sınıflama ve tahmin modellerini içeren birçok problemin çözümünde kullanılabilir. Biyolojik bir nöronun yapısı aşağıda gösterilmiştir.



Şekil 6 YSA Katmanları

Şekil 5'te yapay sinir ağı yapısı görselleştirilmiştir. Belirtilen bu yapıda YSA yapısının üç ana katmandan oluştuğu görülmektedir. Bu ana katmanlar sırasıyla; girdi katmanı, ara katman ve çıktı katmanı olarak ifade edilmektedir. Modelde yer alan her girdi değişkeni bağıntılı ağırlık katsayısı ile çarpılarak ara katmana gönderilmektedir. Gizli katman içerisinde bir veya daha fazla katman yer alabilir. Gizli katmanda gelen değişkenlerin bilgilerine göre doğrusal veya doğrusal olmayan fonksiyon olarak ayrıştırılarak çıktı katmanına dönüştürülmektedir.

Yapıyı tasarlamamanın en basit yolu, elemanları uygun gruplara yerleştirmektir. Katmanlar üç bölümden oluşmaktadır. Sinir hücrelerini katmanlara bölüştürmek, aralarındaki

bağlantıları ayırmak ve transfer ve kümeleme fonksiyonlarını gruplamaktır. Rastgele sinir hücre kombinasyonlarının yapılması mümkün değildir. Paralel bağlanmaları sonucunda ağ giriş, gizli ve çıkış katmanı olarak oluşmaktadır (Öztemel, 2012:58).

Öztemel, (2012:58) Yapay Sinir Ağları katmanları aşağıdaki şekilde açıklamıştır;

Giriş Katmanı: Giriş katmanı içinde yer alan nöronlar, bilgiyi alarak gizli katmanlara aktarımı sağlar. Giriş katmanı içindeki hiçbir bilgiyi işlemeyen ağlar mevcut olup, bu ağlar giriş değerlerini sonraki katmana gönderirler. Bazı analizciler bu sebeple giriş katmanını ağ katman sayısı içine dahil etmemektedir.

Gizli (Orta) Katman: Orta katmana giriş katmanından gelen bilgiler işlenerek çıkış katmanına iletilir. Ağ modelinde giriş ve çıkış bir katmandan oluşmaktadır. Ancak gizli katmanlar birden fazla katman içerebilir. Gizli katmanda yer alan nöron sayısının belirlenmesi, ağ performansını belirlemek için önemli bir kriterdir. Nöron sayıları ve katman sayılarında yapılan artış ya da azalışlar, ağın basit ya da karmaşık olup olmasını sağlar.

Çıktı Katmanı: Çıkış katmanı içerisinde bulunan nöronlar, orta katmandan gelen bilgileri işleyerek, giriş katmanında üretilen örnek için uygun çıkışı oluşturmaktadır. YSA sınıflandırma, tahmin, öğrenme işlemlerini yapabilir. Bu sebeple öğrenme sürecinde verileri düzenleyerek girdiler için daha fazla doğru karar alma performansı göstermektedir.

3.1. Yapay Sinir Ağlarının Öğrenme Süreci

Süreç elemanlarının birbiriyle olan ilişkisinin, ağırlık değerlerinin belirlenmesi işlemine ağın eğitimi adı verilir. Ağ eğitiminde, bir modelin en optimum düzeyde kullanılabilmesi için çeşitli denemeler yapılmaktadır. Denemeler sonucunda bulunan değer ağırlıkları, ağa birçok kez tanıtılarak ağ eğitimi gerçekleştirilmektedir. Ağ öğrenimi iki aşamada

yapılmaktadır. Örnek olarak alınan çıktılarının sonuçları kıyaslanır ve doğruluk derecesi en yüksek olan çıktıya göre ağ bağlantılarının ağırlıkları hesaplanır. Bu sayede en fazla doğruluğa sahip değer ağı öğretilmiş olmaktadır.

Ağın öğrenimi sonrasında öğrenme durumunun ölçülmesi ağın test edilmesi anlamına gelmektedir. Ağ, içerisine tanıtılmayan farklı örneklerle test edilmektedir. Yapılan bu testte, daha önce deneme yoluyla elde edilmiş olan ağ ağırlığı değiştirilmemektedir. Yapılan testin örnekleri ağı gösterilir. Önceki çıktılar sonucunda, yeni verilen örnekler için sonuçlar üretmesi beklenir. Alınan sonuçlar değerlendirilerek yani tahmin derecesine bakılarak, ağ eğitimi ile ilgili bilgi edinilmiş olunur. Tahmin edilen değerler ile gerçekleşen değerler arasındaki sonuç ne kadar yakın olursa, ağın öğrenme gücü de bu denli iyi olmaktadır. YSA modellerinde ağ eğitimi için yararlanılan veriler ağ eğitim verisi, test aşamasında kullanılan veriler ise test verisi olarak adlandırılmaktadır. Bilinmeyen verileri, ağı sunulan veriler ile yorumlama durumuna YSA modellerinde adaptif öğrenme adı verilmektedir (Öztemel, 2012: 55).

3.2. Yapay Sinir Ağlarının Kullanım Alanları

Günümüzde YSA modellerinin kullanım alanları gittikçe yaygınlaşmaktadır. YSA hem doğrusal hem de doğrusal olmayan problemlerin çözümünde kullanılması sebebiyle geniş bir çalışma alanına sahiptir. Yapay Sinir Ağlarının kullanıldığı alanlar şu şekilde sıralanabilir;

- Sınıflandırma problemlerinin çözümünde,
- Veri filtreleme ve sıkıştırma konularında,
- Veri yorumlama konularında,
- Ağ analizleri sonrası veri gruplamada konularında,
- Zaman serileri analizlerinde,
- Üretim uygulamalarında,
- Konu ilişkilendirmelerinin ardından tespit edilen eksik bilgilerin veri ilişkilendirmesi ile giderilmesi konularında,
- Girdi değerlerinden çıktı değerlerinin tahminleri konularında kullanılır.

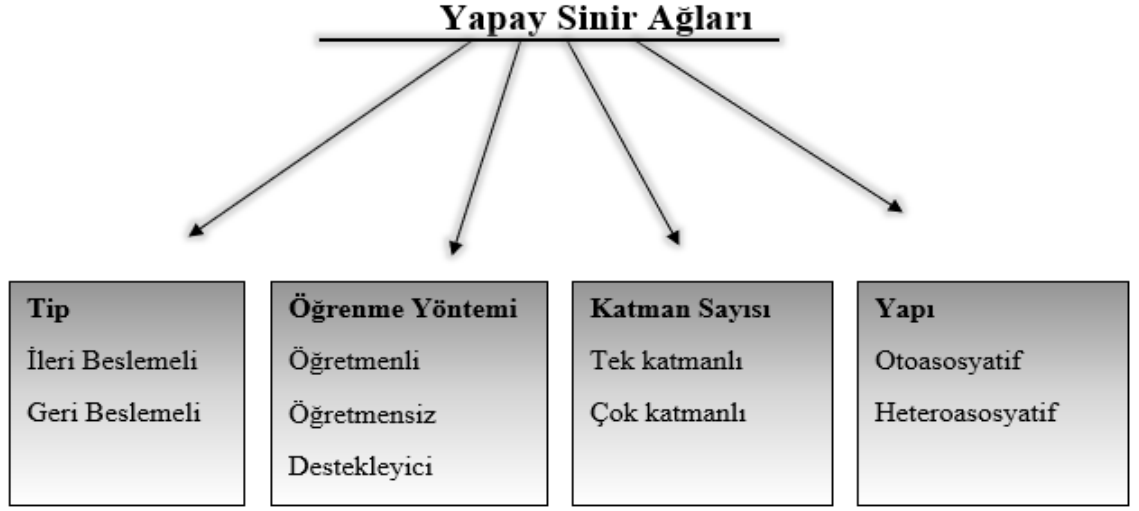
3.3. Yapay Sinir Ağlarının Genel Özellikleri

Günümüzde birçok alanda kullanılan Yapay Sinir Ağları genel özellikleri sebebiyle birçok alanda tercih edilmektedir. Birçok kaynaktan toplanan bilgilerin değerlendirilmesi sonucunda; YSA modellerinin genel özelliklerini özetlemek gerekirse, aşağıdaki sonuçlara ulaşılması mümkün olmaktadır.

- YSA makine öğrenmesini bilgisayar yoluyla gerçekleştirir.
- YSA klasik yöntemlere göre farklılık gösterir, doğruluk oranı deneme sayılarına bağlı olarak artırılabilir.
- YSA'da ağ yapısı bilgiyi saklar.
- YSA, örnekler yoluyla öğrenmeyi gerçekleştirir ve gelecek hakkında tahminde bulunur.
- YSA, veri setinde zaman faktörünün yer aldığı uygulamalarda kullanılır (Tüzüntürk, Selim 2010:79).
- YSA yapısı genelleme yeteneğine sahiptir.
- YSA, ağa tanıtılmasının ardından kendi kendine öğrenme yapabilir.
- YSA, şekil ilişkilendirebilme, sınıflandırma ve tahmin yapabilir.
- YSA, eksik bilgi ile de çalışabilmektedir.
- YSA, hatalara karşı tolerans payı vardır.
- YSA, hem doğrusal hem de doğrusal olmayan her türlü problemlerde kullanılabilir.

3.4. Yapay Sinir Ağı Modelleri

Yapay Sinir Ağları, tahmin ve sınıflama olarak doğrusal ve doğrusal olmayan problemlerin çözümünde kullanılmaktadır. Tahmin veya sınıflama işlemi yapılırken, veri setine uygun olarak YSA modelleri seçilmektedir.

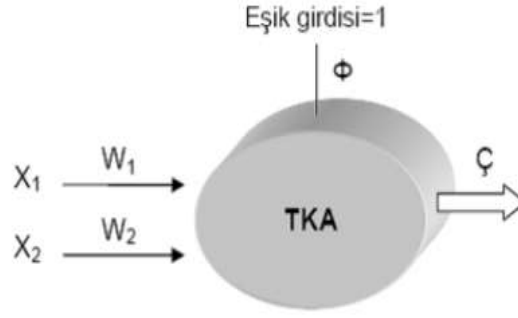


Kaynak: Hamzaçelebi, 2011:19.

Yapay Sinir Ağları katman yapılarına göre; tek katmanlı algılayıcılar ve çok katmanlı algılayıcılar olmak üzere iki grup altında incelenebilir.

3.4.1. Tek Katmanlı Algılayıcılar

Girdi ve çıktı katmanından oluşan modellere tek katmanlı algılayıcılar adı verilmektedir. Modelde yer alan ağlar, giriş ve çıkış olarak birden fazla ağ yapısı ile ilişki kurabilir. Kurulan her ilişkinin ağ katsayısı bulunmaktadır. Eşik değeri, ağın çıktı değerinin sıfır olmasını engellemektedir. Eşik değeri, +1 olarak giriş değerine sahiptir. (Öztemel, 2012:69). Tek katmanlı algılayıcılar, genellikle lineer problemler için tercih edilirler. Tek katmanlı algılayıcılar, doğrusal olmayan olayları öğrenememektedir. Bu durum modelin en önemli problemi olarak gösterilmektedir.



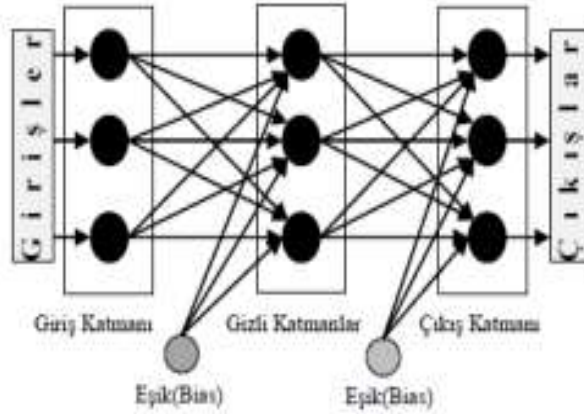
Şekil 7 Tek katmanlı algılayıcı modeli

Kaynak: Arı, Ayşe & Berberler, Murat (2017:57)

Şekil 6’da bir girdi ve bir çıktısı olan tek katmanlı yapay sinir ağı modeli gösterilmiştir.

3.4.2. Çok Katmanlı Algılayıcılar

Giriş ve çıkış katmanı arasında en az bir gizli katmanın bulunduğu yapılara çok katmanlı algılayıcılar adı verilmektedir. (Richard & P.,Lippmann,1987:15). Tek katmanlı algılayıcıların, doğrusal olmayan problemlerde yeterli olmaması sonucunda geliştirilmiş bir modeldir.



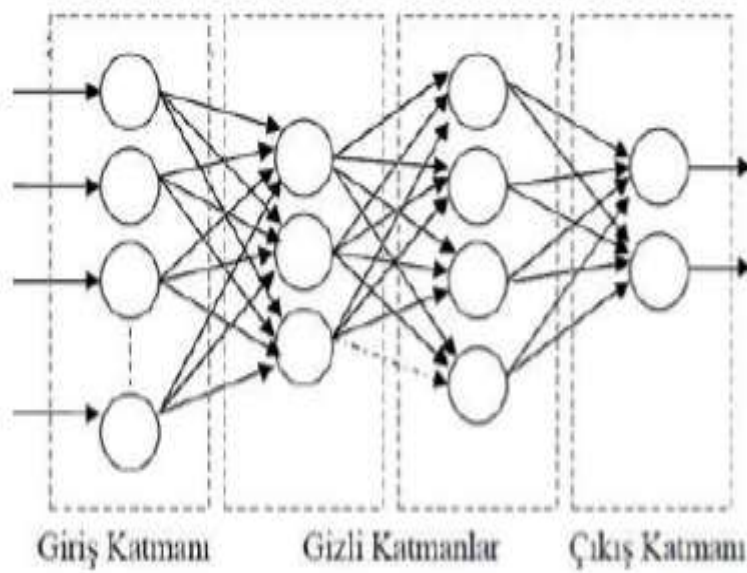
Şekil 8 Çok katmanlı algılayıcı modeli

Kaynak: Arı, Ayşe & Berberler, Murat (2017:57)

Şekil 7’de çok katmanlı bir yapay sinir ağı yapısı görülmektedir. Giriş ve çıkış katmanı arasında gizli katman yer almaktadır. Şekil 7’de akışın ileri doğru olduğu görülmektedir. YSA modelleri, sinirler arasındaki ilişkilerin yönlerine göre ileri beslemeli (feed forward) ve geri beslemeli (feedback recurrent) olarak ikiye ayrılmaktadır.

3.4.3. İleri Beslemeli Yapay Sinir Ağları

İleri beslemeli ağlarda akış ileri yönlü olarak hareket eder ve işlemciler katmanlar halinde bulunmaktadır. Çok katmanlı algılayıcılar (MLP) ve sayısallaştırılmış öğrenme vektörü (LVQ) ağları ileri beslemeli ağlara örnek olarak verilebilir (Saygılı, 2008).



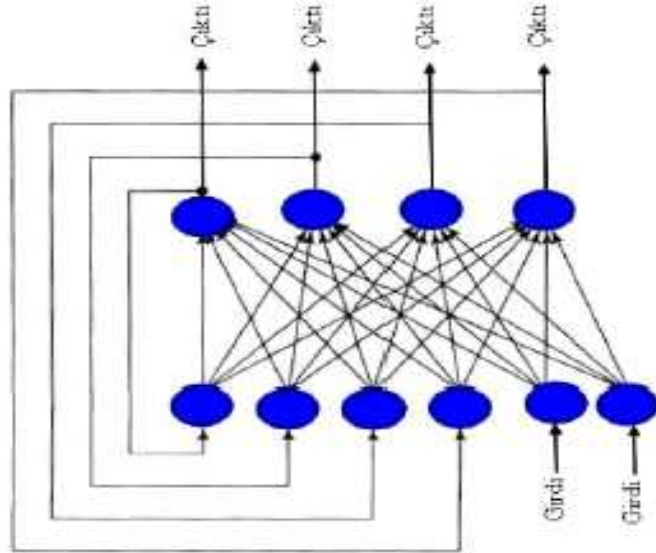
Şekil 9 İleri Beslemeli Yapay Sinir Ağı

Kaynak: Hamzaçelebi, 2011: 44.

Bu tip ağlarda döngüler bulunmamaktadır ve katmanlar arasında tek yönlü bağlantılar bulunmaktadır. Girdilerden çıktılara doğru akış tek yönde gerçekleştirilmektedir (Tüzüntürk, Selim & Eteman S. Fatma & Sezen, Kemal 2016:133). Giriş katmanındaki veriler herhangi bir işleme sokulmadan direkt gizli katmana aktarılmaktadır. Gizli katmandaki eleman sayısı deneme sonuçlarıyla belirlenmekte ve en yüksek doğruluğa sahip katman sayısı tercih edilmektedir.

3.4.4. Geri Beslemeli Yapay Sinir Ağları

İleri beslemeli ağlarda nöronun çıktısı sıradaki nöron katmanına girdi olarak aktarılmakta iken, geri beslemeli ağlarda kendinden önceki veya aynı katmandaki bir başka nörona aktarımı yapılabilmektedir. Bu durum modelin dinamik bir yapıda olmasını sağlamaktadır. Önceki giriş değerleri de nöronların çıkışını etkilemektedir. Bu sebeple geri beslemeli Yapay Sinir Ağları tahmin problemleri için daha uygun bir yapı sergilemektedir.



Şekil 10 Geri beslemeli ağ yapısı

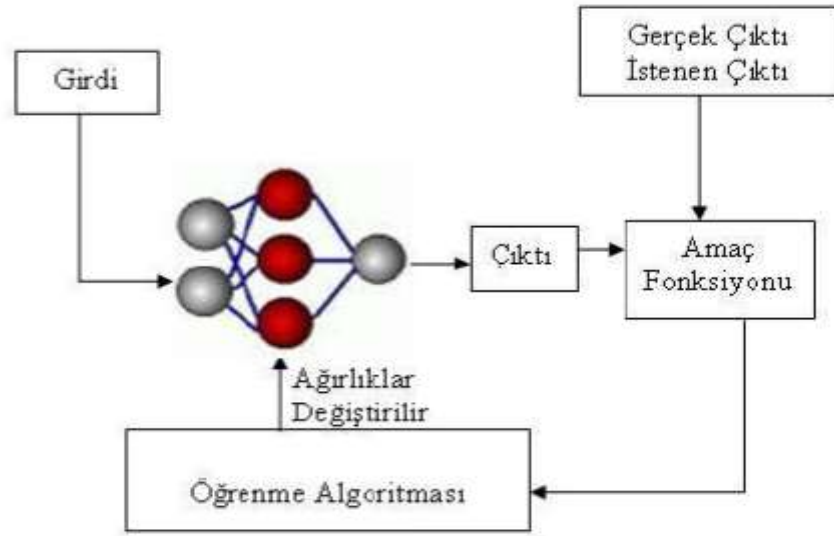
Kaynak: Hamzaçelebi, 2011: 20.

3.5. Öğrenme Yöntemine Göre Yapay Sinir Ağları

Yapay Sinir Ağları modellerinde amaç ağın öğrenmesini sağlayarak, ağa tanıtılmayan verilerin tahminini de sağlayabilmektir. YSA modellerinde ağın öğrenme yöntemleri öğretmenli öğrenme ve öğretmensiz öğrenme olarak ikiye ayrılmaktadır.

3.5.1. Öğretmenli Öğrenme

Bu öğrenme yönteminde ağın analizi için girdi ve talep edilen çıktı değerleri ağa tanıtılır. Çıktı değerlerinin tanıtılmasıyla birlikte ağdan aynı değerleri vermesi beklenir. Bu yöntemdeki en önemli amaç, girdi değişkeninden üretilecek çıktı değişkeni ile gerçek çıktı değişkeni arasındaki hatanın da ağa tanıtılmasıdır. Böylelikle, en minimum seviyede ağın öğrenmesi hedeflenmektedir. (Hamzaçelebi, 2011: 19)



Şekil 11 Öğretmenli Öğrenme Modeli

3.5.2. Öğretmensiz Öğrenme

Öğretmensiz öğrenme yönteminde ağa herhangi bir çıktı değeri tanıtılmamaktadır. Ağa sadece girdi değeri tanıtılmaktadır. Tanıtılan girdi değeriyle ağın belirli ağırlıklar oluşturarak çıktı değerini kendisinin bulması beklenmektedir.

4. Regresyon Analizi

Regresyon, süreklilik gösteren nicel değişkenlerin aralarında bir ilişki (doğrusal veya doğrusal olmayan) olduğunu kabul eden ve diğer değişkenler yardımıyla tahmin yapan bir yöntem olarak açıklanabilir (Bidgoli 2004:208). İki veya daha fazla sayıda nicel değişken arasındaki ilişkiyi ölçmek için kullanılır. Eğer analiz tek bir değişken baz alıyorsa tek değişkenli regresyon, birden çok değişken baz alınarak yapılıyorsa çok değişkenli regresyon analizi olarak adlandırılır.

Model çıktısı, tahmin değeri Y_x ile tahmin girdileri X_1, X_2, \dots, X_n arasındaki ilişkiyi inceler. Burada Y_x bağımlı değişken X_1, X_2, \dots, X_n ise bağımsız değişkenlerdir. Regresyonun amacı tahmin etmektir. Bağımlı değişkenin bağımsız değişkenler yoluyla tahmin edilmesi amaçlanmaktadır. Gözlem değerleri “gerçek değerler”, tahmin edilen değerler ise “teorik değerler” dir. Gerçek değerler ve tahmin edilen değerlerin eşit olması durumunda değişkenlerin arasındaki ilişki tam ve kesin olmaktadır. Ancak değişkenlerin arasındaki ilişkinin eşit olmaması durumu, hata payını (ϵ) değerini göstermektedir.

Bağımlı değişkenin tipine göre regresyon analizleri doğrusal, doğrusal olmayan ve lojistik regresyon olarak adlandırılmaktadır. Regresyon analizinin başlıca amaçlarını özetlemek gerekirse aşağıdaki sonuçlara ulaşılabilir:

1. Bağımsız değişkenlerden alınan bilgiler ile bağımlı değişkenin tahmin edilmesi asıl amaçtır.
2. Bağımlı değişken üzerindeki etkilerin tespit edilmesi. İlişkinin yönü ve büyüklüğünün araştırılmasının sağlanması.
3. Bağımlı değişkenin tahmin edilmesi amacıyla bağımsız değişkenler üzerine değerler verilmesidir.

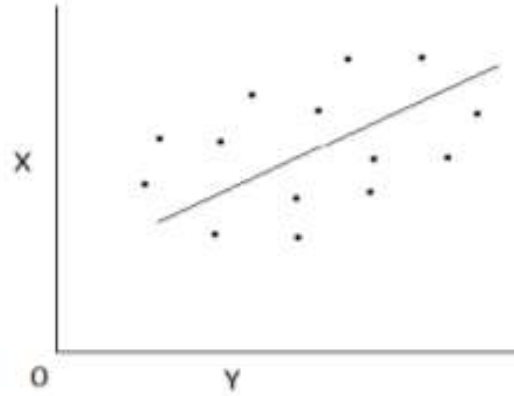
Regresyon yöntemi basit doğrusal regresyon, çoklu regresyon ve lojistik regresyon olmak üzere üç grup altında incelenmektedir. Tez çalışmasında çoklu doğrusal regresyon modeli kullanılmıştır.

4.1. Basit Doğrusal Regresyon

Bağımlı ve bir bağımsız değişkeni bulunan model iki değişkenli regresyon veya basit regresyon modeli olarak isimlendirilmektedir. Basit regresyonun istatistiksel modeli aşağıdaki gibidir.

$$Y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

ε değeri gerçek değerler ile tahmini yapılan değerlerin arasındaki farkı ifade eder. Ayrıca bu değer hata terimi olarak bilinmektedir. Söz konusu model, anakütle regresyon modeli olarak adlandırılır ve bu modelde yer alan değişkenlerin dağılımı dikkate alındığında aralarındaki ilişkinin doğrusal olduğu kabul edilmektedir.



Şekil 12 Basit Doğrusal Regresyon Modeli

Kaynak: Güriş & Çağlayan, 2010: 80.

Şekil 7'de belirtilen grafik aynı zamanda serpilme diyagramı olarak adlandırılır. Değişkenlerin arasındaki ilişkinin tam doğrusal olmaması sebebiyle, arada oluşan

sapmalarını açıklamak için kullanılır. β_0 sabit değer olarak adlandırılır. Modelde yer alan bağımsız değişkenin sıfır olması durumunda bağımlı değişkenin alacağı değeri ifade eder. Aynı zamanda bağımsız değişkenin y eksenini kestiği noktayı da ifade etmektedir.

β değeri bağımsız değişkenin parametresini ifade eder. Bağımsız değişkende ortaya çıkan değişimlerin bağımlı değişken üzerindeki etkisini inceler. Bu etkinin hangi yönde ve hangi düzeyde etkili olduğunun araştırılmasını sağlar. İlişkinin yönünün tespiti için değişken katsayısının işaretine bakılmaktadır. Eğer işaret pozitif yönde ise aralarındaki ilişkinin doğru yönlü, negatif ise ilişkinin ters yönlü olduğu ifade edilir. β_0 ve β_1 hesaplanan ana kütle parametrelerini ifade etmektedir. Eğer, β_0 ve β_1 bilinmiyorsa, anakütleden örneklem alınarak parametreler için istenen bilgiler üretilebilir. Bu durumda tahmini değerler olarak b_0+b_1 kullanılmaktadır.

Ana kütle doğrusal regresyon modelinin örneklem tahmini aşağıdaki gibi ifade edilir. Burada \hat{y} y'nin tahmini değerini ifade etmektedir.

$$\hat{y}=b_0+b_1x \quad (2)$$

Modele hata terimi ilave edildiğinde örneklem tahmin fonksiyonu aşağıdaki gibi olmaktadır.

$$\hat{Y}=b_0 + b_1x + e \quad (3)$$

e değeri modelde artık veya hata terimi olarak adlandırılır. \hat{Y} tahmin edilen değerleri ifade etmektedir. Artıklar ise; $e= Y - \hat{Y}$ olarak ifade edilir.

Sevüktekin (2013:244-246) eserinde, hata teriminin sebeplerini aşağıdaki şekilde açıklamıştır.

- i. Hata terimi, modele katılmamış ya da ihmal edilmiş değişkenlerin etkisini yansıtmaktadır. Bir ekonomik modelde ilgili bütün değişkenlerin yer alması istenmektedir. Ancak model kurulurken hangi değişkenin bağımlı değişkeni ne kadar etkileyeceği ölçülmediğinden, bir değişken önemsiz gibi görülerek modele dahil edilmeyebilir. Bunun sonucunda modelde, değişken eksikliğinden dolayı hata etkisi gözlenebilmektedir.
- ii. Matematiksel modelin yapısı hata terimini etkilemektedir. Modelde doğrusal olmayan bir ilişki mevcut iken doğrusal bir model seçilmesi hataya neden olabilmektedir.
- iii. Değişkenlerin ölçme hatalarındaki etkiler hata terimi üzerinde toplanmaktadır. Bağımsız değişkenin bağımlı değişkeni tahmininde, rassal hatalar e ile absorbe edilmektedir.
- iv. Hata terimini etkileyen bir diğer husus insanın öngörülemeyen davranışları sebebiyledir. Gerçekte var olmayan veya olması tahmin edilemeyen olayların etkileri de hata terimi üzerinde görülmektedir.

Basit doğrusal regresyon analizinde parametre tahminlerinin gerçekleştirilmesinden önce, çeşitli varsayımların kabul edilmesi gerekmektedir. Tahmincilerin her biri bir örnekleme dağılımına sahiptir. Bu örnekleme dağılımları bazı varsayımlar ile geçerlilik kazanmaktadır. (Sevüktekin,2013:247-264)

1. Doğrusallık: x'in her değeri için y'nin ortalama değeri parametrelerine göre doğrusal bir regresyon ile tanımlanmaktadır. Bu varsayımın ihlali sonucu modelde tanımlama (spesifikasyon) hatası ortaya çıkmaktadır.
2. Rassallık: Hata teriminin rassal bir değişken olduğu kabul edilmektedir. Aynı zamanda Y hata terimini içerdiğinden, Y değerinin de rassal olduğu kabul edilmektedir.
3. Hata teriminin beklenen değeri sıfırdır. Bu durum rassal hataların anakütle ortalaması civarında hareket ettiği anlamına gelmektedir. Bu durumun ihlal edildiği durumlarda β_0 sapmalı olmaktadır.

4. x 'lerin gözlem değerlerinin en az bir tanesinin diğerlerinden farklı olduğu kabul edilmektedir. x 'lerin gözlemlerinde bir değişme olmaz ise bağımlı değişkeni açıklamak mümkün olmamaktadır.
5. Rassal hata teriminin sabit bir varyansla dağıldığı kabul edilmektedir. Yani saçılımı ve dağılımının eşit olması durumunu (homoskedastisite) ifade etmektedir. Bu durumun ihlali sonucunda değişken varyans (heteroskedastisite) ortaya çıkmaktadır.
6. Rassal hata teriminin korelasyonsuz yani birbirinden bağımsız olduğu kabul edilmektedir. Eğer bir hata terimi ortalama değerinin üstündeysen bir başka hata teriminin de ortalamanın üstünde olacağı anlamına gelmez. Bu durumda hata teriminin rassal olması durumunu açıklamaktadır.
7. Hata teriminin ortalamasının sıfır ve sabit bir varyansla dağıldığı kabul edilmektedir.
8. Tekrarlanan bir örneklem sürecinde x 'lerin sabit kaldığı, y ve hata teriminin ise değiştiği kabul edilmektedir. Bu varsayımın ihlali sonucunda; bağımsız değişkenlerin ölçümünde hatalar, bağımlı değişkenlerin gecikmeli değerlerinin bağımsız değişken olarak kullanılması ve bağımlı değişkenin birkaç ilişkiyi aynı anda tanımlaması hataları ortaya çıkabilmektedir.
9. Açıklayıcı değişkenlerin hatasız olduğu varsayılmaktadır.
10. İki veya daha fazla değişken arasında tam doğrusal bir ilişkinin olmadığı varsayılmaktadır.
11. Ele alınan değişkenler toplan değerlerinden oluşuyorsa, bu değerlerin doğru olarak alındığı varsayılmaktadır.

12. Bağıntıların katsayılarının tahmin edilmesi için oluşturulan matematiksel modelin tek olduğu varsayılmaktadır.
13. Modelin matematiksel biçiminin doğru olduğu ve tüm açıklayıcı değişkenleri modelin içerdiği varsayılmaktadır.
14. Gözlem sayısının tahmin edilecek değişken katsayısından fazla olmalıdır. Modelin bu şekilde doğruluğu kabul edilmektedir.
15. Tahminin beklenen değeri, açıklayıcı değişkenlerin parametrelerine bağlıdır.
16. Bağımlı değişkene ait gözlemlerin korelasyonsuz olduğu varsayılmaktadır.
17. y değerinin kendi ortalamasında ve normal dağıldığı varsayılmaktadır.

4.2. Çoklu Regresyon Modeli

Ekonometrik olarak, bir analiz yapılmak istendiğinde iki değişkenli modelin (basit regresyon modeli) yetersiz kaldığı görülmektedir. Çünkü gerçek hayatta bir analiz yapıldığında, tek bir değişken kullanarak bir konuyu tahmin etmemiz mümkün değildir. Bir bağımlı değişken ve birden fazla bağımsız değişkenin olduğu modeli ifade eder. Çoklu regresyon modelinde amaç bağımsız değişkenlerden en çok hangisinin bağımlı değişkeni etkilediğinin tespit edilmesidir. Diğer bir amaç ise, bağımsız değişkenler kullanılarak bağımlı değişkenin tahminidir.

X_i 'ler bağımsız değişkenleri ve Y bağımlı değişken olmak üzere, çoklu regresyon denklemi aşağıdaki şekilde gösterilmektedir.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon \quad (4)$$

Dördüncü modelde β değerleri parametreleri ifade eder. Modelde ifade edildiği gibi; k sayıda değişken ve k sayıda parametre mevcuttur. β_0 değişkeni modelde yer alan diğer

tüm deęişkenlerin sıfır olması durumunda, baęımlı deęişkenin alacağı deęeri yani sabit deęeri ifade eder. Ayrıca anakütle regresyon doğrusunun y eksenine ne kadar uzakta olduğunu belirtmektedir. Marjinal etkileri yansıtan kısmi regresyon katsayılarına ise β_1, \dots, β_k parametreleri denmektedir. Hata terimi y ile x'ler arasındaki ilişkiler tanımlanırken y'deki ölçme hatalarını ifade etmektedir.

Basit doğrusal regresyon modelinde regresyon analizinin varsayımlarından bahsedilmiştir. Bu varsayımların çoęu çoklu doğrusal regresyon modeli için de geçerlidir. Çoklu regresyon modellerinde baz alınan varsayımlar şu şekilde özetlenebilir (Sevüktekin 2013:396-404):

1. Doğrusallık: Çoklu gerçek ilişki regresyon denkleminde X_1, X_2, \dots, X_k 'anın her deęeri için y'nin ortalama deęeri $\beta_0, \beta_1, \dots, \beta_k$ 'a göre doğrusaldır.
2. Rassal deęişkenler: hata terimi bir rassal deęişkeni ifade eder. Yani hata teriminin alacağı deęerler olasılıklı ve şansa baęlıdır. Y de hata terimini içermesi sebebi ile rassal yani deęişken olarak ifade edilmektedir.
3. Hata teriminin anakütle ortalaması ve beklenen deęeri sıfırdır. Bu varsayımla, ihmal edilen ve modelin belirlenmesi sürecindeki hataların etkisinin sıfır olduęu varsayılmaktadır.
4. Açıklayıcı deęişkenlerin gözlem deęerlerinden en az bir tanesi dięerlerinden farklıdır.
5. Açıklayıcı deęişkenler baęımsız olduęu ve hata terimi ile arasında bir korelasyon bulunmadıęı varsayılmaktadır.
6. Hata teriminin varyansının sıfır olduęu kabul edilmektedir. Benzer şekilde y'nin olasılık dağılımının varyansının da sabit olduęu kabul edilmektedir.
7. İki farklı gözleme ait rassal hata arasındaki kovaryansın sıfır olduęu kabul edilmektedir.

8. Sabit bir varyansla ve sıfır ortalamayla dağılan hata terimi, normal dağılıma sahiptir.

9. x'ler tekrarlanan örneklem kümesinde sabit değerler aldığı varsayılmaktadır.

10. Açıklayıcı değişkenlerin ölçümlerinin hatasız olduğu varsayılmaktadır. Modele değişkenlerin doğru bir şekilde yerleştirildiği kabul edilmektedir.

11. Çoklu doğrusal regresyon modelinde açıklayıcı değişkenlerin arasında herhangi bir doğrusal bağlantının olmadığı varsayılmaktadır.

4.3. Lojistik Regresyon Modeli

Çok değişkenli bir modelde bağımlı ve bağımsız değişken ayrımının yapılması durumunda ve bağımlı değişken nominal ölçekli olduğunda en küçük kareler (EKK) tekniği yetersiz kalmaktadır. Bu durumda varyanslar, minimum olma özelliğini yitirmiş olmaktadır. EKK tekniklerine alternatif olarak lojistik modelleri yer almaktadır. Lojistik regresyon modellerinin parametreleri analitik olarak elde edilememektedir. Bu nedenle maksimum olabilirlik (maximum likelihood- ML) yöntemi kullanılmaktadır (Kalaycı 2010:273).

İki değişkenli lojistik regresyon modeli aşağıdaki gibi tanımlanmaktadır.

$$P(Y) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \Rightarrow P(Y) = \frac{e^{\beta_0 + \beta_1 X}}{1+e^{\beta_0 + \beta_1 X}} = \frac{1}{1+e^{-(\beta_0 + \beta_1 X)}} \quad (5)$$

Çok değişkenli modelde lojistik regresyon modeli ise aşağıdaki gibi açıklanmaktadır.

$$\frac{P(Y)}{Q(Y)} = \frac{P(Y)}{1-P(Y)} = \frac{e^z / (1+e^z)}{1 - e^z / (1+e^z)} = e^z = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X} \quad (6)$$

Her iki tarafın logaritması alındığında en sade şekliyle denklem aşağıdaki formuna ulaşmaktadır.

$$\ln\left(\frac{P(Y)}{Q(Y)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (7)$$

$$OR = \frac{P(Y)}{Q(Y)} = e^z = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p} = \text{Exp}(\beta) \quad (8)$$

Üstünlük oranı $\text{Exp}(\beta)$ olarak ifade edilmektedir. 1'den büyük olduğu durumlarda y değeri üzerindeki etkisinin önemli olduğunu ifade etmektedir.

5. Destek Vektör Makineleri

Destek Vektör Makineleri (DVM), makine öğrenmesinin alt dalı olarak karşımıza çıkmaktadır. Makine öğrenmesi; olasılık, istatistik ve optimizasyon gibi matematiğin alt dalları üzerine kurulmuştur. Model kurulur ve makineye geçmiş veriler verilerek geleceği tahmin etmesi beklenir. Yapay zekâ, bilgisayar bilimlerinin bir alt dalıdır. Bilgisayarların akıllı bir biçimde davranmasını sağlayarak, insanlara benzer şekilde öğrenmesi amaçlanarak endüstride kullanılmaktadır.

N boyutlu hiper düzlem oluşturmak için iki kategoriye ayıran DVM modelleri YSA yapılarıyla yakından ilişkilidir. DVM ileri beslemeli, iki katmanlı bir ağ yapısına sahiptir (Haykin, 1999: 318). DVM, çok büyük veri setleri ya da ultra büyük veri setleri içinde yüksek performans göstermektedir. Kolay uygulanabilir olmasıyla pazarlama, görüntü ve sınıflama gibi birçok alanda kendine yer edinmektedir. Ayrıca aykırı değerlere karşı dayanıklı olması sebebi ile de tercih edilmektedir. Sınıflandırma ve regresyon problemlerinin çözümlenmesinde de kullanılmaktadır. Bu problemlerin çözümünde de iyi performans göstermektedir.

İstatistiksel öğrenme teorisindeki gibi doğrusal ayırıcı bir fonksiyonun bulunması, eldeki eğitim verilerinin mümkün olduğunca gerçeğe yakın bir şekilde yansıtılmasını sağlamaktadır. Bu durum, DVM regresyon metodundaki temel fikirdir. DVM, doğrusal olmayan yapılar için çekirdek fonksiyonları kullanmaktadır (Çomak, 2008: 37). Ele alınan eğitim kümesinde her bir değişkenin iki kategoriden birine ait olduğu varsayıldığında, DVM eğitim algoritması yeni kategorilere yeni örnekler atayan bir model oluşturmaktadır. Bu durum DVM modelinin ikili doğrusal sınıflandırıcı olmasını sağlar. Haritalama yaklaşımını benimsemesi sebebi ile, verileri geniş aralıklarla ayrı kategorilere bölmektedir. Bu sayede, yeni örneklem verilerinin de hangi kategoriye ait olduğunun tahmin edilmesini gerçekleştirir (Jordan vd., 2008:30-3). DVM teorisi yapısal risk minimizasyonu (SRM) fikrine dayanmaktadır. Amaç, hiper bir düzlem meydana getirmektir. (Suykens & Vandewalle, 1999:293-300). İki sınıfın örneklerini ayıran birçok hiper düzlem olabilir. Amaç, bu hiper düzlemin optimal ve marjı maksimize etmesidir (Lin & Wang, 2002: 465).

DVM için model kurulduğunda;

$x_i \in R^n \rightarrow$ gelecek vektörü

$y_i \in R \rightarrow$ hedef çıktı $\{(x_1, y_1), \dots, (x_l, y_l)\}$

Girdi ve çıktı değişkenleri arasındaki doğrusal olmayan ilişki doğrusal bir fonksiyonla formüle edilmelidir.

$$f(x) = w^t \Phi(x) + b \quad (5)$$

Denklemden; $f(x)$: öngörülen tahmin değerleri, Φ : doğrusal olmayan haritalama fonksiyonunu ifade eder. DVM standart formu $C > 0$ ve $\epsilon > 0$ olmak üzere aşağıdaki gibi tanımlanır:

$$\text{Min}_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + c \sum_{i=1}^{\ell} \xi_i + \xi_i^* \quad (6)$$

ξ_i^* değeri, ε üzerindeki eğitim hatalarını, ξ_i değeri ise ε altındaki eğitim hatalarını ifade eder.

Lagrange çarpanları ile çözümden sonra elde edilen denklem aşağıdaki gibi olacaktır:

$$w = \sum_{i=1}^{\ell} (\lambda_i^* - \lambda_i) \phi(x_i) \quad (7)$$

Böylelikle DVM formülü elde edilmiş olur.

$$f(x) = \sum_{i=1}^{\ell} (\lambda_i^* - \lambda_i) K(x_i, x_j) + b \quad (8)$$

5.1. Destek Vektör Makinelerinin Avantajları

- Tahmin yöntemi olarak basittir ve tahmin doğruluk seviyesi diğer yöntemlere göre daha üstündür.
- Karmaşık karar modellemelerini gerçekleştirebilme yeteneğine sahiptir.
- Bağımsız değişken sayısı fazla iken de çalışmasını sürdürebilir.
- Doğrusal ve doğrusal olmayan veriler ile çalışma imkânı sunar.
- Genelleme yeteneği fazladır.

5.2. Destek Vektör Makinelerinin Dezavantajları

- Çekirdek seçimi en büyük sınırlamasıdır.
- Test ve eğitim verisi arttıkça, algoritmanın hızı düşmektedir.

5.3. SMO Algoritması

Ardışık minimal optimizasyon algoritması (SMO) algoritması Destek Vektör Makineleri yöntemini kullanarak regresyonu gerçekleştirmektedir. Söz konusu algoritma, “Platt“ tarafından geliştirilmiştir. Alt kümeler ile çalışır. İki boyutlu çalışma kümelerini tekrarlı olarak seçer. Bu seçime uyan amaç fonksiyonları uygunlaştırarak parçalamayı en uç noktaya kadar gerçekleştirir. İki boyutlu veri setleri ile çalıştığı için uygunluk alt problemi analitik olarak çözülebilmektedir. Algoritma, tüm eğitim örnekleri uygun olma koşulunu sağlayıncaya kadar yinelenir.

SMO algoritması, her aşamada lagrange çarpanı için en optimal değeri hedefler ve en küçük optimizasyon sonucunu amaçlar. İki lagrange çarpanını çözebilmesi, SMO algoritmasının avantajıdır. Matris algoritması kullanmaması sebebiyle hassasiyete daha tepki gösterir. (Elmas,2003:19).

Tahmin, elde var olan veri setiyle geleceği ön görme işlemidir. Veri setinin uygunluğuna göre birçok algoritma ve model yardımıyla tahmin işlemi gerçekleştirilebilmektedir. Gerçekleştirilen tahminlerin hangisinin daha başarılı olduğunun tespit edilmesi için hata oranlarının karşılaştırılması yöntemi kullanılabilir. Tez çalışmasının uygulama kısmında ele alınan algoritmaların hangisinin daha başarılı tahmin gerçekleştirdiğini bulmak için de hata oranları; ortalama mutlak hata, kök ortalama kare hatası ve ortalama mutlak yüzde

hata yöntemleri kullanılacaktır. Bunun gerçekleştirilebilmesi için, tahmin performanslarının açıklanması gerekmektedir.

6. Tahmin Performansları

Tahmin edilen değerler ile gerçekleşen değerlerin arasındaki farkın kıyaslanmasıyla tahmin performanslarının etkinliğinin ölçümü gerçekleştirilebilir. Tahmin modellerinin doğruluklarının denetimi yapılan tahminin güvenilirliği açısından önemlidir. Tahmin öngörü tekniklerinin güvenilirliğini sınamak için birçok hata ölçüm teknikleri mevcuttur. Tez çalışmasında yer verilecek hata ölçüleri sırasıyla açıklanmıştır.

6.1. Ortalama Mutlak Hata

Ortalama Mutlak Hata, (Mean Absolute Error - MAE), öngörüdeki hataların ortalama büyüklüklerini ölçmektedir. Tüm bireysel farklılıkların eşit ağırlığa sahip olduğu tahmin ile, gerçek gözlem arasındaki mutlak farklılıkların test örneği üzerindeki ortalamasıdır. MAE değeri 0'dan ∞ 'a kadar değişebilir. Negatif yönelimli puanlar yani daha düşük değerlere sahip tahminleyiciler daha iyi performans gösterir. Ortama mutlak hata aşağıdaki formül ile gösterilmektedir.

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad (9)$$

Formülde mutlak değer alınmasının nedeni; ortalama sapma hatası (MBE) dönüşümünü engellemek içindir. MBE model sapmasını ortalama olarak vermektedir. Ancak burada mutlak değer alınmadığından olumlu ve olumsuz hataların birbirini götürme riski bulunmaktadır. MBE formülü aşağıdaki gibi açıklanabilir.

$$MBE = \frac{1}{n} \sum_{i=1}^n (m_i - c_i) \quad (10)$$

6.2. Kök Ortalama Kare Hatası

Kök Ortalama Kare Hata (RMSE), tüm hatanın karesinin ortalamasının kare köküdür. RMSE kullanımı çok yaygındır ve sayısal tahminler için mükemmel bir genel amaçlı hata metriği olarak kabul edilir. RMSE değeri 0'dan ∞ 'a kadar değişebilir. Negatif yönelimli puanlar yani daha düşük değerlere sahip tahminleyiciler daha iyi performans gösterir. RMSE değerinin sıfır olması modelin hiç hata yapmadığı anlamına gelir. Gerçek değer ve tahmin değeri arasındaki karesel farkın karekökü olarak ifade edilmektedir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (m_i - c_i)^2} \quad (11)$$

6.3. Ortalama Mutlak Yüzde Hata

Ortalama Mutlak Yüzde Hatası, (Mean Absolute Percentage Error- MAPE) tahmin doğruluğunu kontrol etmek için en yaygın kullanılan ölçüdür. Ölçekten bağımsız olan ve farklı ölçeklerdeki serileri karşılaştırmak için kullanılabilen yüzde hatalarının altında gelir. Ölçülen ve tahmini değer arasındaki yüzdesel sapmanın mutlak değeri ortalaması şeklinde açıklanabilir.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{m_i - c_i}{m_i} \right| \times 100}{n} \quad (12)$$

MAPE değeri %10'un altında ise model yüksek doğruluk seviyesine, %20-%50 arası doğruluk seviyesi var ise kabul edilebilir seviyede olduğu ve %50 üzerinde bir değere sahip ise yanlış ve hatalı model kullanıldığı ifade edilir (Lewis, 1982:40).

ÜÇÜNCÜ BÖLÜM

BEYAZ EŞYA SEKTÖRÜNDE SATIŞ TAHMİNİ UYGULAMASI

1. Uygulamanın Amaç ve Önemi

Bu çalışma ile 2021 yılına ait firmanın satış adetleri tahmin edilmeye çalışılmış ve bu tahminler ışığında üretim için gerekli hammadde temini, nakliye hizmeti ve stok maliyetlerinin planlanması hedeflenmiştir. Firmaların kar etme amacı ile maliyetlerini minimum da tutmaları gerekmektedir. Bunun gerçekleşebilmesi için ise iyi bir planlama ve bir sonraki yılı yüksek oranda tahmin ederek yol izlemek önemlidir.

Bu çalışmanın amacı, veri madenciliği teknikleri içerisinde yer alan yapay sinir ağı, destek vektör makineleri ve çoklu doğrusal regresyon yöntemlerini kullanarak satış tahmini yapmaktır. Yapılan bu tahminleri inceleyerek birbiri ile kıyaslayarak arasındaki farklılıkları incelemek ve satıcı açısından maksimum kâr sağlayarak, zaman kaybını minimuma indirmeye çalışmaktır.

2. Weka Program Tanıtımı

Weka (“The Waikito Environment for Knowledge Analysis”), programı Yeni Zelanda’da bulunan “Waikito Üniversitesi” tarafından geliştirilmiştir. Weka programının logosunda ve işlemleri gerçekleştirirken pencerelerin çeşitli yerlerinde gördüğümüz kuş simgesi, Yeni Zelanda’ya özgü bir kuştur. Bu kuş türü aynı zamanda programın kısa adı olan Weka ile aynı isme sahiptir.

Weka programı farklı amaçlar için ve farklı dillere sahip birçok insan için uygun bir programdır. Bu sebeple son yıllarda popülerliği gittikçe artan bir program olmaktadır.

Programın kullanımı ücretsizdir. Açık kaynak kodlu bir programdır. Hazır algoritmalar ile çalışılması mümkün olduğu gibi, Java kodu ile algoritma oluşturulması da mümkündür.

Weka programında veri madenciliği uygulamalarının birçok işlemi yapılabilmektedir.

Bu işlemlerden en çok kullanılanlarını özetlemek gerekirse;

1. Weka ile veri kümesindeki ön çalışma gerçekleştirilebilir.
2. Model oluşturulabilir.
3. Veri setine göre özel sınıflandırma ve algoritma yöntemi seçilebilir.
4. Tahmin günlük, haftalık, aylık ve yıllık olarak yapılabilir.
5. Model çıktı sonucu analiz edilebilir ve grafikleri oluşturulabilir.

Bu sebepler sonucunda Weka programı kullanımı gittikçe yaygınlaşmaktadır. Özellikle tahmin çalışmalarında ve ağaç diyagramlarının oluşturulmasında sıklıkla tercih edilen bir programdır. Weka'da ara yüz kullanımının basitliği de kullanıcılar için tercih sebeplerinden birini oluşturmaktadır. Weka programında kullanılan sekmeler hakkında aşağıda kısaca bilgi verilmiştir.

Explorer: Bu buton Veri Madenciliği araştırması için kullanılır.

Experimenter: Farklı makine öğrenimi metotları ve veri türleri üzerinde performans karşılaştırmalarının yapılabilmesi amacıyla bu buton kullanılır.

Knowledge Flow: Program araçlarının grafiksel ara yüzlerine bu buton ile ulaşılır.

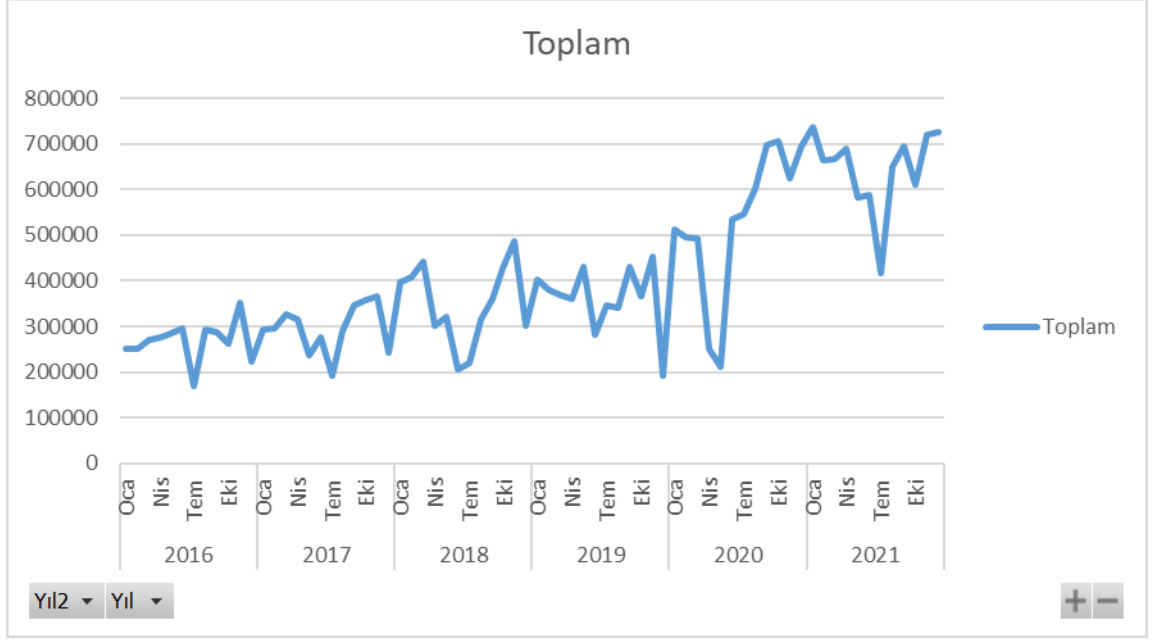
Simple CLI: Komut satırı ara yüzüne bu buton vasıtasıyla ulaşılır.

Zaman serilerinde öngörü “Time Series Forecasting” modülü, zaman serisi veri setlerinde, periyodik bazda sıralı gelecek değerlerin tahmin edilmesi amacıyla kullanılmaktadır. Weka programı yüklendikten sonra bu modülün ayrıca kurulması gerekmektedir. WEKA uygulaması indirildiğinde, uygulamanın içinde “Time Series Forecasting” eklentisi gelmemektedir. Eklentinin, “ Weka GUI Chooser – Tools – Package Manager” dan yüklenmesi gerekmektedir. Package manager sekmesine giriş yapıldıktan sonra açılan ekranda uygulanmak istenilen yöntem seçilerek yüklenir. Uygulamanın sağ tarafında bulunan menüde yer alan “Explorer” butonu tıklanarak açılan ekrana dosya açma ya da doya aktarım işlemi uygulanabilmektedir. Çalışmanın amacına göre veri işleme, sınıflandırma, veri ilişkilendirme ve veri kümeleme algoritmaları kullanılarak analizler yapılmaktadır. Weka programında veri girişi yapılabilmesi için arff dosyasının hazırlanması ve verilerin bu dosya da düzenlenmesi gerekmektedir. Ancak Weka'nın güncel sürümünde CVS dosya uzantısı ile veri girişi yapılması mümkündür. Excel üzerinde veriler hazırlanarak CVS uzantısı ile kaydedilerek veri girişi yapılabilir.

3. Satış Tahmininde Kullanılan Değişkenler

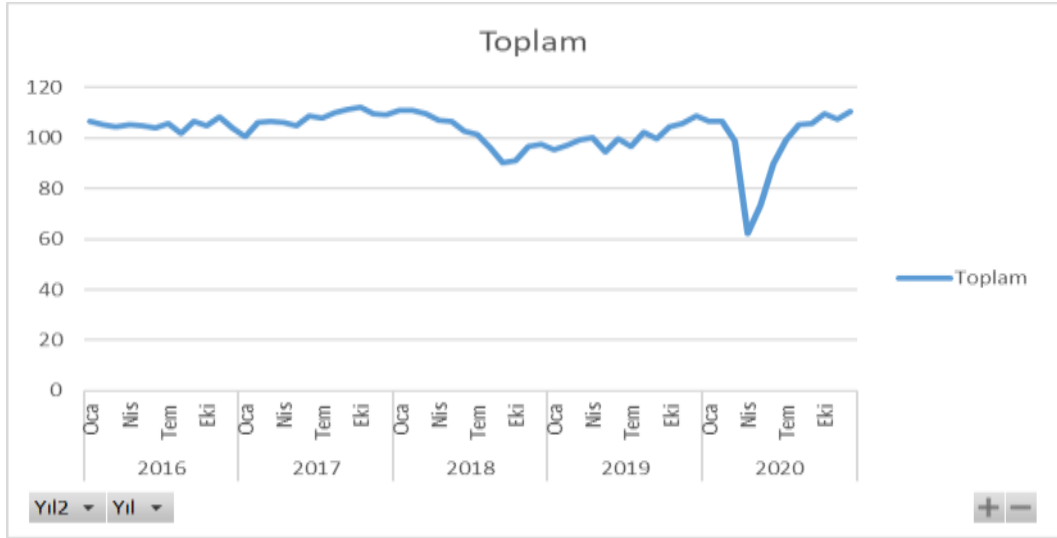
Satış tahmininin gerçekleştirilebilmesi için hangi değişkenlerin modele etki ettiği belirlenmeye çalışılmıştır. Uygulamada bağımlı değişken olarak kullanılacak veri satış adetleridir. Verilerin alındığı firma ihracat ve yurtiçi piyasasında önemli tedarikçi konumundadır. Satış değişkenine etki eden değişkenlerin tespiti için literatür taramasındaki çalışmalardan yararlanılmıştır. Bu sebeple firma satışlarında sadece yurtiçi piyasası baz alınmış ve iç piyasa satışlarını etkileyen değişkenler modele dahil edilmiştir. Modelde kullanılan değişkenler sırasıyla açıklanmıştır.

Aylık Ürün Satışları: 01.2016-12.2021 yılları arasındaki ürün satışları bağımlı değişken olarak alınmıştır. Uygulamada firmanın yurtiçi piyasasına yapmış olduğu satış rakamları baz alınmıştır. Bağımlı ve bağımsız değişkenler sırasıyla açıklanmıştır.



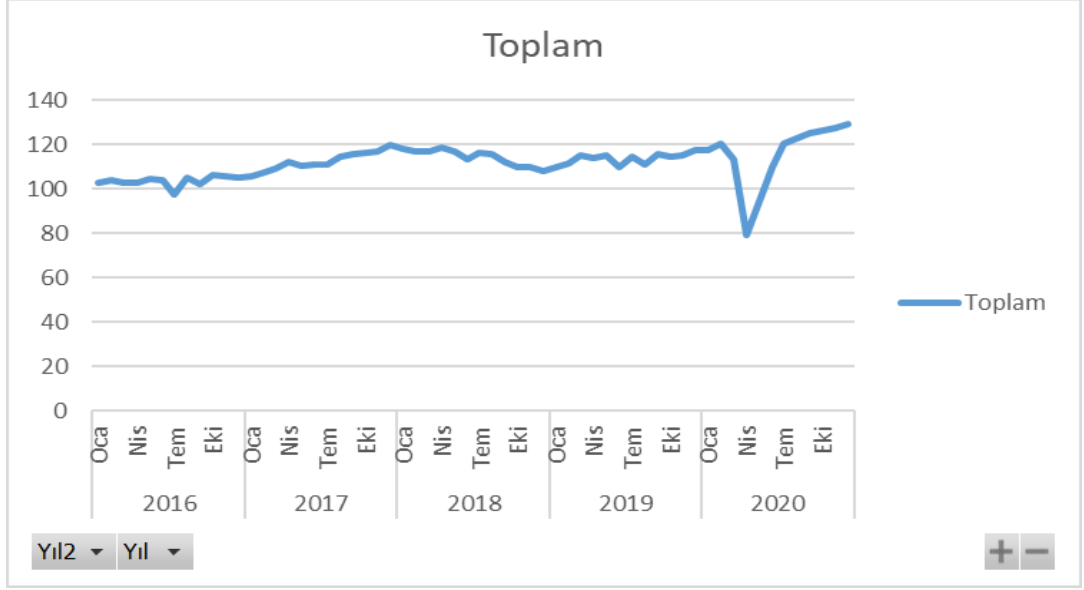
Şekil 13 Sevk Adetleri Grafik

Reel Kesim Güven Endeksi (RKGE): Veriler aylık olarak Türkiye Cumhuriyeti Merkez Bankası (TCMB) tarafından yayınlanmaktadır.



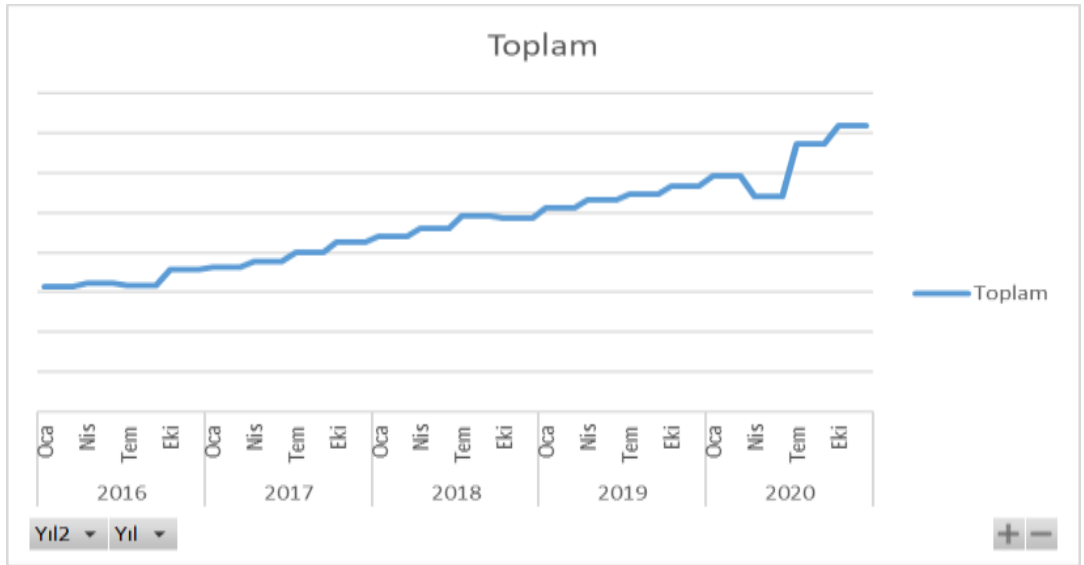
Şekil 14 RKGE Grafik

Sanayi Üretim Endeksi (SÜE): Veriler aylık olarak Türkiye İstatistik Kurumu (TÜİK) tarafından yayınlanmaktadır.



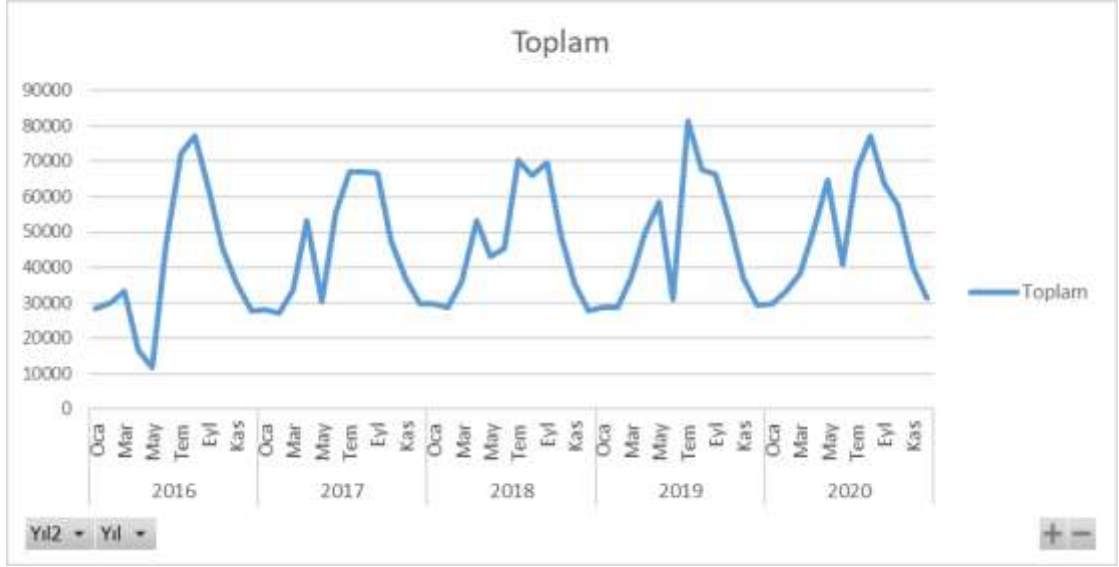
Şekil 15 SÜE Grafik

Gayri Safi Yurtiçi Hasıla (GSYH): Veriler aylık olarak Türkiye İstatistik Kurumu (TÜİK) tarafından yayınlanmaktadır.



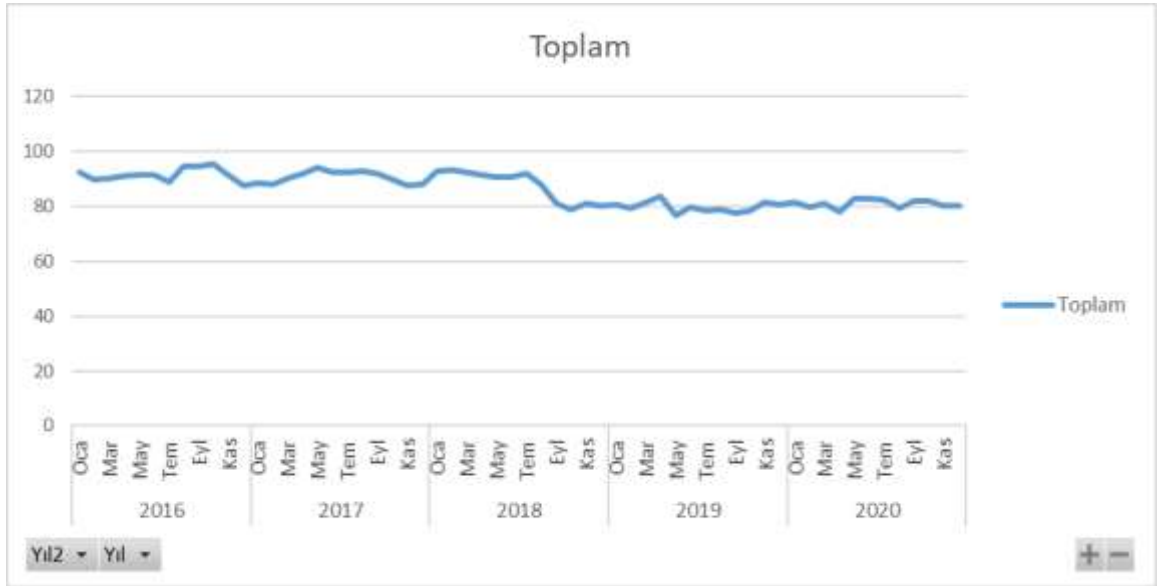
Şekil 16 GSYH Grafik

Evlenme Oranı: Veriler aylık olarak Türkiye İstatistik Kurumu (TÜİK) tarafından yayınlanmaktadır.



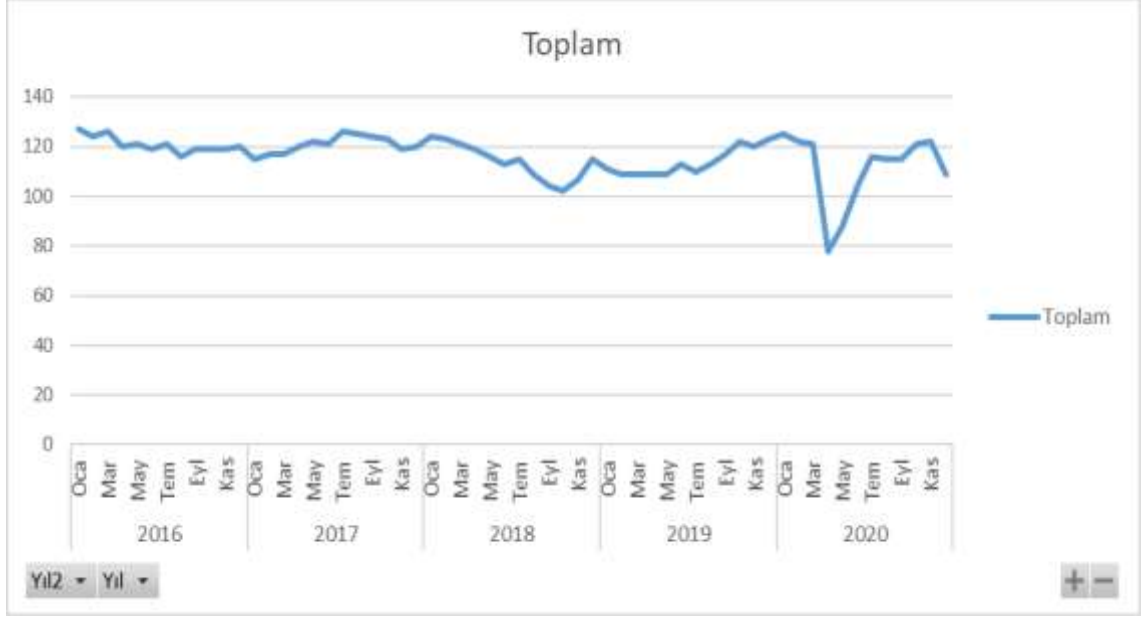
Şekil 17 Evlenme Oranı Grafik

Tüketici Güven Endeksi: Verileri aylık bazda TÜİK yayınlamaktadır. TGE, tüketicilerin ekonominin seyrini görmelerinin ardından harcama veya tasarruf hakkında verdikleri kararı gösteren bir orandır.



Şekil 18 TGE Grafik

Ekonomik Güven Endeksi (EGE): Verileri aylık bazda TÜİK yayınlamaktadır. Tüketici ve üreticilerin ekonomik duruma karşı beklentilerini yansıtan bir oranı ifade etmektedir.



Şekil 19 EGE Grafik

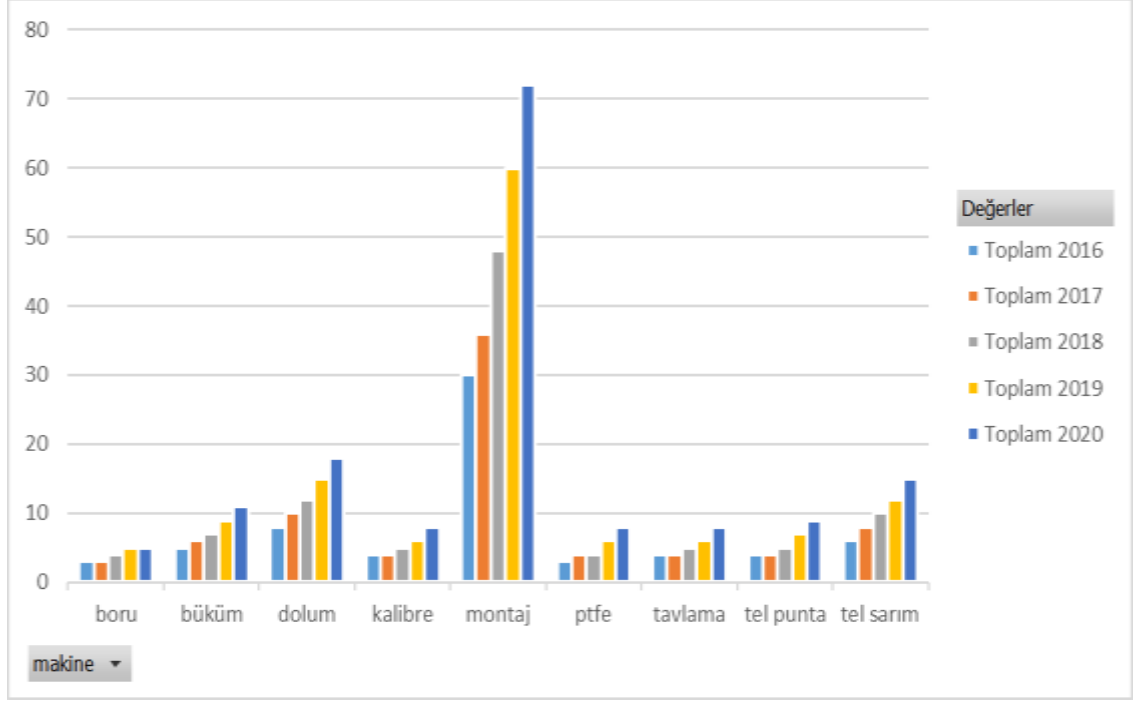
4. Firma Tanıtımı

Etrafımızda var olan maddelerin bazen hazır bazen de dönüştürülmüş hali ile hizmetinin gerçekleştirilmesi üretim olarak adlandırılabilir. Büyüyen ve sürekli zorlaşan piyasa şartlarında, firmaların varlığını sürdürebilmeleri için stratejik kararlar almaları gerekmektedir. Üretim yönetimi en az maliyetle, standart bir niteliğe sahip ve en fazla kazanç ile ürünleri var etmek olarak ifade edilebilir. Planlama yapılırken müşterilerinin isteklerinin de göz önüne alınması gerekmektedir. Müşterinin o üründen beklentisi, sağladığı fayda düzeyinin artırılması konusunda yardımcı olacaktır. Geliştirilen bir ürün müşteri geri dönüşleriyle daha iyi seviyeye getirilebilir. İyi bir üretim yönetimi anlayışıyla, firma artan sermayesini doğru bir yatırım aracına dönüştürebilir veya eksik yanlarını telafi fırsatını yakalayabilir. Bu doğrultuda zaman, para ve doğru kararlar varlık gösterebilmek için önemli unsurlardandır. Hepimiz geleceğimiz ile ilgili hayaller kurmuş ve kendimizi birkaç yıl sonra nelerin beklediğini düşünmüştür. Firmalar da varlıklarının devamı için bu düşüncüyü gütmektedirler. Önümüzdeki yılın satışları ne kadar olacak? Elde edilecek ciro nasıl değerlendirilmeli? Elde ne kadar stok tutulmalı? gibi kritik sorulara cevap aranmaktadır. Bu sebeple uygulamada ürünlerin satış tahminleri gerçekleştirilerek belirtilen sorunların çözümü için cevap aranacaktır.

Tüketicilerin belirli bir fiyattan piyasadan almak istedikleri ürünlere talep denmektedir. Talep tahmini işletmelerin gelecek dönemlere ait durumunun önceden belirlenmesi işlemidir. Talebi etkileyen değişkenlerin çoğalmas ve talep için ön görüsel değerlendirmelerin yetersizliğı istatistiki yöntemlere başvurulmasını zorunlu kılan nedenlerdendir. Bu çalışmada belirli bir sektörde faaliyet gösteren firmanın satış adetlerinin tahmini farklı metotlarla test edilecek ve tahmin edilen değerler ile gerçek değerler kıyaslanarak tahmin performansları incelenecektir. Bu durumda tahmin yöntemleri arasında da bir kıstas ortaya konmuş olacaktır.

Araştırılacak olan firma, metal sanayi sektöründe faaliyet göstermektedir. Rezistans üretimini gerçekleştiren firma Türkiye ve Dünyada beyaz eşya sektöründe isim yapmış ana sanayilere parça üretimi yapmaktadır. Bu sebeple, 37 yıldır önemli bir tedarikçi konumundadır. Firmanın çalışma prensibi sıcak satış yerine yıllık anlaşma bazlı satıştır. Yıl sonunda yapılan kapasite planı doğrultusunda siparişler kabul edilir ve üretim bir sonraki yıla bu şekilde başlar. Dolayısıyla firmanın bir sonraki yıl ne kadar üretmesi gerektiğı içinde bulunduğu yılın başarısına ve gelecekteki yılın ön görüleri ile olabilecektir. Üretimde kullanılan ham maddeler yurt dışından ithal edilmekte ve transit süreleri ortalama 4-6 ay arasındadır. Üretimin devamlılığı için hammadde siparişleri ve takibi önemli bir rol üstlenmektedir. Gelecek planı doğru yapılarak verilecek hammadde siparişleri yönetilebilir ve sifıra yakın risk ile üretimin devamlılığı sağlanabilir.

Rezistans üretimi üretim tesisi içerisinde 9 ayrı bölümden geçerek tamamlanmaktadır. Bu bölümler ve makine dağılımları aşağıda yer alan grafikte gösterilmiştir.



Şekil 20 Makine Dağılımı

Uygulama, veri madenciliğinde kabul edilmiş bir standart olan CRISP-DM basamakları takip edilerek sürdürülmüştür. Bu basamaklar;

- Yapılacak işi veya araştırmayı anlama,
- Kullanılacak veriyi anlama,
- Veriyi hazırlama,
- Modelleme,
- Değerlendirme,
- Sonuçları yayma şeklindedir.

5. YSA Uygulaması

Öğrenme katsayısını bulabilmek için 0.1-0.8 değerleri arasında deneme yapılmıştır. Bulunan öğrenme katsayısı da dikkate alınarak momentum katsayısı 0.1-0.8 değerleri arasında test edilmiştir. Değerlerin arasındaki optimizasyonun değerlendirilmesi

amacıyla MAPE, MAE ve RMSE deęerleri ölçüt alınmıştır. Uygulamada kullanılan gerçek veri setleri, sinir aęlarına girmeden önce normalizasyon işlemi yapılmıştır. Bu işlem sonucunda giriş verilerinin tümü $[0,1]$ aralığına indirgenmiştir.

5.1. Öğrenme ve Momentum Katsayısının Belirlenmesi

Deneme No	Öğrenme Katsayısı	Momentum Katsayısı	MAPE	MAE	RMSE
1	0,1	0,1	0,2090	0,0738	0,0878
2	0,1	0,2	0,2148	0,0757	0,0897
3	0,1	0,3	0,2196	0,0769	0,0899
4	0,1	0,4	0,2130	0,0746	0,0862
5	0,1	0,5	0,2045	0,0722	0,0852
6	0,1	0,6	0,2480	0,0844	0,1012
7	0,1	0,7	0,3240	0,1094	0,1320
8	0,1	0,8	0,6731	0,2237	0,2530
9	0,2	0,1	0,2107	0,0722	0,0875
10	0,2	0,2	0,2170	0,0733	0,0910
11	0,2	0,3	0,2168	0,0723	0,0916
12	0,2	0,4	0,2499	0,0852	0,1000
13	0,2	0,5	0,2253	0,0757	0,0935
14	0,2	0,6	0,2457	0,0773	0,0949
15	0,2	0,7	0,3275	0,1122	0,1218
16	0,2	0,8	0,3382	0,1207	0,1458
17	0,3	0,1	0,1819	0,0639	0,0810
18	0,3	0,2	0,1913	0,0669	0,0823
19	0,3	0,3	0,2273	0,0776	0,0926
20	0,3	0,4	0,2611	0,0845	0,0992
21	0,3	0,5	0,3153	0,0968	0,1243
22	0,3	0,6	0,4103	0,1262	0,1871
23	0,3	0,7	0,3599	0,1268	0,1476
24	0,3	0,8	0,4637	0,1672	0,1806
25	0,4	0,1	0,1726	0,0648	0,0904
26	0,4	0,2	0,1994	0,0774	0,1069
27	0,4	0,3	0,2726	0,1008	0,1166
28	0,4	0,4	0,3089	0,114	0,1326
29	0,4	0,5	0,3640	0,1294	0,1571
30	0,4	0,6	0,5735	0,1868	0,21
31	0,4	0,7	0,8037	0,2745	0,2948
32	0,4	0,8	0,9616	0,3294	0,3386
33	0,5	0,1	0,1580	0,0542	0,0717
34	0,5	0,2	0,2151	0,0759	0,0984
35	0,5	0,3	0,3100	0,1056	0,1344
36	0,5	0,4	0,5504	0,1898	0,2223
37	0,5	0,5	0,4258	0,1274	0,1615
38	0,5	0,6	0,3372	0,124	0,1666
39	0,5	0,7	0,2369	0,0758	0,095
40	0,5	0,8	0,5238	0,1905	0,2115
41	0,6	0,1	0,2158	0,0748	0,0946
42	0,6	0,2	0,2243	0,0797	0,1002
43	0,6	0,3	0,2634	0,0986	0,1278
44	0,6	0,4	0,4005	0,1437	0,1723
45	0,6	0,5	0,5747	0,1859	0,2022
46	0,6	0,6	0,7796	0,2405	0,2639
47	0,6	0,7	0,2687	0,0769	0,0961
48	0,6	0,8	0,7399	0,2186	0,2393
49	0,7	0,1	0,2461	0,0838	0,1121
50	0,7	0,2	0,2360	0,0827	0,1023
51	0,7	0,3	0,3460	0,0983	0,1184
52	0,7	0,4	0,3718	0,1272	0,1534
53	0,7	0,5	0,8095	0,2549	0,2756
54	0,7	0,6	1,3269	0,3914	0,4234
55	0,7	0,7	0,5842	0,1656	0,1873
56	0,7	0,8	2,8689	0,7406	2,5981
57	0,8	0,1	0,1873	0,0598	0,0783
58	0,8	0,2	0,1942	0,0666	0,0871
59	0,8	0,3	0,2290	0,0752	0,1068
60	0,8	0,4	0,8412	0,255	0,2731
61	0,8	0,5	1,5099	0,4574	0,4932
62	0,8	0,6	0,3099	0,0893	0,1054
63	0,8	0,7	0,8279	0,2478	0,267
64	0,8	0,8	0,9616	0,3294	0,3386

Şekil 21 Öğrenme katsayısı ve momentum katsayısı için yapılan deneme sonuçları

YSA modelleri verinin tanıtılmasının ardından ağırlıklarının belirlenmesine ağ etiğini denilmektedir. Öğrenme katsayısının doğru seçimi ağı öğrenme performansında oldukça etkilidir. Öğrenme katsayısının belirlenebilmesi için, optimal sonuca ulaşılan kadar deneme gerçekleştirilmiştir. Denemeler 0.1-0.8 değerleri arasında gerçekleştirilmiştir. Tahmincilerin kıyaslamasında MAPE, MAE, RMSE ölçü birimi kullanılmıştır. MAPE, MAE ve RMSE değeri küçük olan tahminci için daha iyi performans ortaya koyduğu söylenebilir. Yapılan denemeler sonucunda; MAPE, MAE ve RMSE değerleri kıyaslandığında tüm hata düzeylerinde en optimal sonucu veren öğrenme katsayısı 0,5 ve momentum katsayısı 0,1 olduğu görülmektedir.

5.2. Çevrim Katsayısının Belirlenmesi

Belirli bir deneme sonrasında ağı öğrenmesinin tamamlandığı düşünülmektedir. Çevrim katsayısı 500-2000 arasında alınmıştır. Baz alınan öğrenme ve momentum katsayıları sırasıyla belirlenen çevrim katsayılarıyla test edilmiştir.

Deneme No	Öğrenme Katsayısı	Momentum Katsayısı	Çevrim Katsayısı	MAPE	MAE	RMSE
65	0,5	0,1	500	0,1580	0,0542	0,0717
66	0,5	0,1	750	0,2358	0,0848	0,116
67	0,5	0,1	1000	0,261	0,0922	0,1286
68	0,5	0,1	1250	0,2846	0,0961	0,1239
69	0,5	0,1	1500	0,3514	0,1183	0,1498
70	0,5	0,1	1750	0,3489	0,1172	0,1493
71	0,5	0,1	2000	0,3368	0,1139	0,1439

Şekil 22 Çevrim Katsayısının Belirlenmesi

Deneme yapılırken daha önce belirlenen öğrenme katsayısı ve momentum katsayısı kullanılmıştır. Öğrenme katsayısı 0,5 ve momentum katsayısı 0,1 olarak alınmıştır. Çevrim katsayısı ise 500-2000 arasında denenmiştir. Yapılan denemeler hata katsayıları ile kıyaslanmıştır. Sonuçlar incelendiğinde MAPE, MAE ve RMSE değeri en düşük 500

çevrim katsayısını veren değerdir. Bu sebeple çevrim katsayısı 500 olarak belirlenmiştir. Bu durumda modelin;

Öğrenme katsayısı: 0,5

Momentum katsayısı: 0,1

Çevrim katsayısı:500 olarak belirlenmiştir.

Bu aşamadan sonra ağız gizli hücre sayısı ve nöron sayısının bulunması gerekmektedir.

5.3. Gizli Hücre ve Nöron Sayısının Belirlenmesi

Momentum katsayısı 0.5, öğrenme katsayısı 0.1, çevrim sayısı 500 alınarak devamında nöron sayısı ve gizli hücre sayısının hesaplanmasına geçilmiştir. Ağız nöron sayısı yapısal özelliklerinden birini oluşturmaktadır. Bir katmanda bulunan nöron sayısının minimum düzeyde olması gerekmektedir. Gereğinden çok olması ağız ezberlenmesine neden olabilir. Gizli hücre sayısı 1' den giriş hücre sayısının iki katı olan 2'ye kadar denemeler yapılmıştır. En düşük MAPE, MAE ve RMSE değerine sahip olan hücre sayısının 2, bu katmanlardaki nöron sayılarının ise sırasıyla 3 ve 3 olduğu belirlenmiştir. Bu aşamaya kadar bulunan değerler aşağıda özetlenmiştir.

Öğrenme katsayısı: 0.5

Momentum katsayısı: 0.1

Çevrim katsayısı: 500

Gizli hücre sayısı: 2

1. Nöron sayısı: 3

2. Nöron sayısı: 3

Deneme No	Öğrenme Katsayısı	Momentum Katsayısı	Çevrim Katsayısı	Gizli Katman	Nöron sayısı1	Nöron sayısı2	MAPE	MAE	RMSE
71	0,5	0,1	500	1	1	-	0,2793	0,0993	0,1149
72	0,5	0,1	500	1	2	-	0,2045	0,0629	0,0858
73	0,5	0,1	500	1	3	-	0,1933	0,0715	0,0924
74	0,5	0,1	500	1	4	-	0,158	0,0542	0,0717
75	0,5	0,1	500	1	5	-	0,1941	0,065	0,0863
76	0,5	0,1	500	1	6	-	0,2289	0,0787	0,1025
77	0,5	0,1	500	1	7	-	0,193	0,068	0,1047
78	0,5	0,1	500	1	8	-	0,183	0,0609	0,0772
79	0,5	0,1	500	1	9	-	0,2309	0,0698	0,0883
83	0,5	0,1	500	2	1	1	0,4741	0,1339	0,1518
84	0,5	0,1	500	2	2	1	0,1669	0,0558	0,0732
85	0,5	0,1	500	2	3	1	0,2507	0,0907	0,1064
86	0,5	0,1	500	2	4	1	0,474	0,1338	0,1518
87	0,5	0,1	500	2	5	1	0,4741	0,1338	0,1518
88	0,5	0,1	500	2	1	2	0,165	0,0578	0,0765
89	0,5	0,1	500	2	2	2	0,2755	0,0994	0,1131
90	0,5	0,1	500	2	3	2	0,1753	0,0617	0,0783
91	0,5	0,1	500	2	4	2	0,1588	0,0544	0,0714
92	0,5	0,1	500	2	5	2	0,2642	0,0944	0,1088
93	0,5	0,1	500	2	1	3	0,1797	0,0637	0,0807
94	0,5	0,1	500	2	2	3	0,2812	0,7398	0,0998
95	0,5	0,1	500	2	3	3	0,1285	0,0446	0,061
96	0,5	0,1	500	2	4	3	0,259	0,0928	0,1068
97	0,5	0,1	500	2	5	3	0,1501	0,0509	0,0654
98	0,5	0,1	500	2	1	4	0,2684	0,0935	0,1046
99	0,5	0,1	500	2	2	4	0,1756	0,0614	0,0784
100	0,5	0,1	500	2	3	4	0,1317	0,0455	0,0617
101	0,5	0,1	500	2	4	4	0,2781	0,0991	0,1103
102	0,5	0,1	500	2	5	4	0,1384	0,048	0,0639
103	0,5	0,1	500	2	1	5	0,2545	0,0887	0,1005
104	0,5	0,1	500	2	2	5	0,2722	0,0961	0,1062
105	0,5	0,1	500	2	3	5	0,268	0,0942	0,1044
106	0,5	0,1	500	2	4	5	0,2615	0,0922	0,1027
107	0,5	0,1	500	2	5	5	0,1321	0,0457	0,0609

Şekil 23 Gizli Hücre ve Nöron Sayısının Belirlenmesi

Gizli hücre ve nöron sayısının da belirlenmesiyle YSA uygulaması için gereken tüm girdi değerleri bulunmuştur. Bu aşamadan sonra WEKA ile uygulama yapılarak sonuçlar paylaşılacaktır.

Uygulamada bağımlı değişken olarak beyaz eşya sektöründe faaliyet gösteren firmanın yurtiçi piyasasındaki 2016-2021 arasındaki 6 yıllık satış verileri aylık olarak alınmıştır. Verilerin 2016-2020 yılları arasındaki verileri eğitim verisi olarak ağa tanıtılmıştır. 2021 yılına ait veriler ağa tanıtılmamış öğretilimsiz öğrenme yaklaşımı uygulanmıştır. Literatür taramasından elde edilen sonuçlara göre veri setine uygun olan değişkenler bağımsız değişken olarak seçilmiştir. Sırasıyla bu değişkenler; reel kesim güven endeksi, gayri safi yurtiçi hasıla, sanayi üretim endeksi, evlenme oranı, tüketici güven endeksi ve ekonomi güven endeksidir. Verilerin %75'i eğitim verisi %15'i ise test verisi olarak kullanılmıştır. Öğrenme katsayısı 0.5, momentum katsayısı 0.1, çevrim katsayısı 500, gizli hücre sayısı 2,1. nöron sayısı 3, 2. nöron sayısı ise 3 olarak belirlenmiştir. Bu sonuçlara göre 2021 yılı tahmin edilerek gerçekleşen değerler kıyaslanacaktır.

YSA			
MAPE	MAE	RMSE	Korelasyon
0,1285	0,0446	0,061	0,808

Tablo 3 YSA Hata Oranı Sonuçları

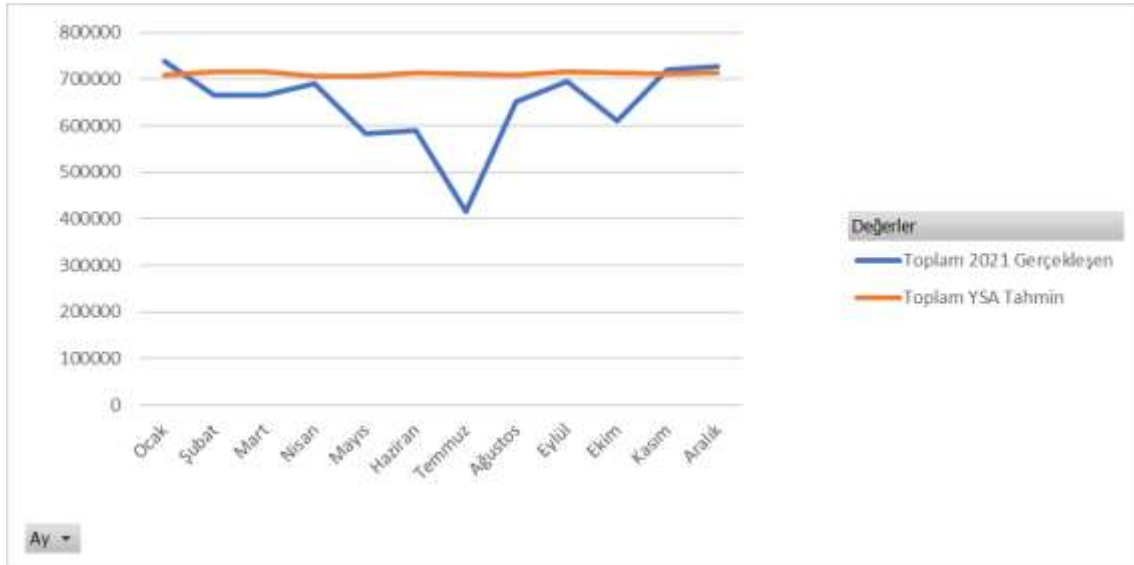
YSA hata oranı sonuçları incelendiğinde MAPE değerinin 0.1285, MAPE değerinin 0.0446 ve RMSE değerinin ise 0.061 olduğu görülmektedir. Hata oranlarının sıfıra yakın olması sebebi ile, tahminin düşük bir hata oranına sahip olduğu söylenebilir.

Korelasyon katsayısının ise 0,808 olduğu görülmektedir. Bağımsız değişkenlerin arasındaki ilişkinin pozitif yönlü ve aralarındaki ilişkinin güçlü olduğu söylenebilir. İstatistiki olarak anlamlıdır. Değişkenler birbiri ile uyumludur ve aynı doğrultuda hareket etmektedir.

Ay	2021 Gerçekleşen	YSA Tahmin	Sapma
Ocak	737.985	708.200	29.785
Şubat	665.124	714.400	49.276
Mart	665.435	715.900	50.465
Nisan	690.074	706.900	16.826
Mayıs	583.149	706.800	123.651
Haziran	588.805	714.100	125.295
Temmuz	416.467	710.500	294.033
Ağustos	650.684	707.900	57.216
Eylül	695.787	714.800	19.013
Ekim	609.367	713.700	104.333
Kasım	719.240	710.500	8.740
Aralık	725.846	712.700	13.146

Tablo 4 YSA Tahmin Sonuçları

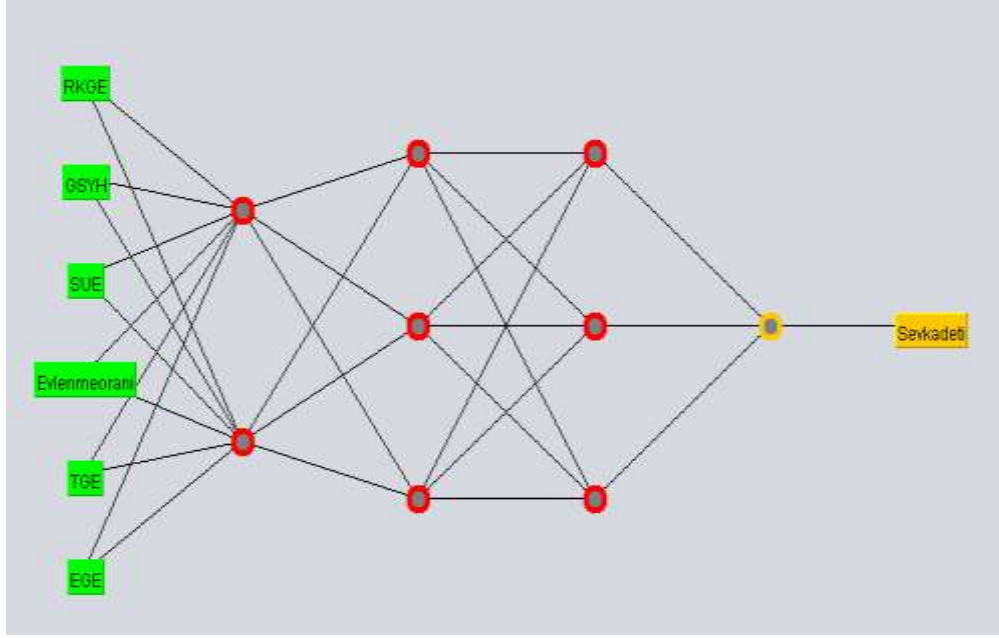
YSA tahmin sonuçları incelendiğinde; gerçekleşen değerler ile tahmin edilen değerlerin çok yakın olduğu gözlemlenmektedir. Bu durum YSA modellerinin iyi bir tahminci olduğunun kanıtıdır. Tablo 5'te elde edilen bilgilere göre YSA tahmin başarısının ortalamasının %89 olduğu görülmektedir.



Şekil 24 YSA Gerçekleşen ve Tahmin Grafiği

Şekil 23'de yer alan grafik incelendiğinde, YSA tahminin belirli bir ortalama etrafında dağıldığı gözlemlenmiştir. Ocak, Nisan, Kasım ve Aralık aylarında gerçekleşen değerler ile

tahmin deęerlerinin oldukça yakın olduęu gözlemlenmektedir. Veri seti arttırıldığında tahmin doęruluęunun giderek artacaęı varsayılabilir.



Şekil 25 YSA Ağ Modeli

Elde edilen ağ modelinde giriş katmanında reel kesim güven endeksi, gayri safi yurtiçi hasıla, sanayi üretim endeksi, evlenme oranı, tüketici güven endeksi ve ekonomi güven endeksi yer almaktadır. Gizli katman sayısı 2 nöron katmanları ise 3 ve 3 şeklindedir. En optimal sonucu veren gizli katman ve nöron sayıları ağa öğretilerek model oluşturulmuştur.

YSA modeli ağ öğreniminin yapılarak sonuçlara ulaşılması sayesinde iyi bir tahminci özellięi göstermektedir. Ayrıca YSA modeldeki hatalara karşı duyarlı deęildir ve hatalar ile de yüksek başarıda tahminleme yeteneęine sahiptir. Bu sebeple son yıllarda birçok araştırmacı tarafından öncelikle tercih edilen bir yöntem olmaktadır.

6. Doğrusal Regresyon Uygulaması

2016-2021 arasındaki 6 yıllık satış verileri alınmıştır. Literatür taramasından elde edilen sonuçlara göre ise veri setine uygun olan değişkenler bağımsız değişken olarak seçilmiştir. Sırasıyla bu değişkenler; reel kesim güven endeksi, gayri safi yurtiçi hasıla, sanayi üretim endeksi, evlenme oranı, tüketici güven endeksi ve ekonomi güven endeksidir.

Regresyon			
MAPE	MAE	RMSE	Korelasyon
0,1576	0,0509	0,0614	0,7892

Tablo 5 Regresyon Sonuçları

Doğrusal regresyon hata oranı sonuçları incelendiğinde MAPE değerinin 0.1576, MAE değerinin 0.0509 ve RMSE değerinin ise 0.0614 olduğu görülmektedir. Verilerin %75'i eğitim verisi %15'i ise test verisi olarak kullanılarak tahmin gerçekleştirilmiştir. Yapay Sinir Ağları modeli ile regresyon sonuçları kıyaslandığında, YSA'nın hata oranlarının regresyon modeline göre daha iyi sonuç verdiği söylenebilir.

Ay	2021 Gerçekleşen	Regresyon Tahmin	Sapma
Ocak	737.985	692.100	45.885
Şubat	665.124	708.200	43.076
Mart	665.435	731.800	66.365
Nisan	690.074	758.400	68.326
Mayıs	583.149	792.300	209.151
Haziran	588.805	818.800	229.995
Temmuz	416.467	848.800	432.333
Ağustos	650.684	877.700	227.016
Eylül	695.787	901.900	206.113
Ekim	609.367	931.700	322.333
Kasım	719.240	975.000	255.760
Aralık	725.846	1.005.900	280.054

Tablo 6 Regresyon Tahmin Sonuçları



Şekil 26 Regresyon Gerçekleşen ve Tahmin Grafiği

Grafik sonucu incelendiğinde, gerçekleşen tahmin değerinin, dalgalanmalar gösterdiği ancak çoklu doğrusal regresyon modelinin sürekli olarak artış eğiliminde olduğu görülmektedir. Bir firmada üretim esnasında beklenmeyen sebepler nedeniyle dalgalanmaların yaşanması muhtemeldir. Bu sebeple sürekli büyümeyi hedefleyen firmalarda dalgalanmalara maruz kalmaktadır. Çoklu doğrusal regresyon modeli YSA modeliyle kıyaslandığında tahmin olarak daha az başarılı sonuç verdiği görülmektedir. MAPE, MAE ve RMSE değerlerinde de YSA daha başarılı sonuçlar elde etmiştir.

7. Destek Vektörleri Makineleri Uygulaması

DVM, istatistiksel öğrenme teorisine dayanmaktadır. Yapısal risk minimizasyonunu belirlemesi DVM geleneksel sinir ağlarına göre daha fazla avantaj sunmaktadır.

DVR			
MAPE	MAE	RMSE	Korelasyon
0,1987	0,0601	0,0812	0,7741

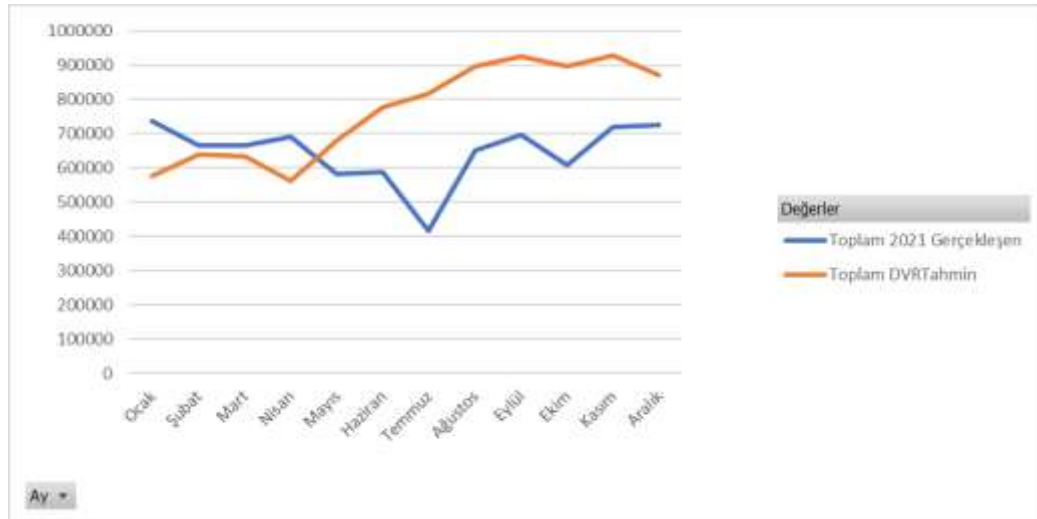
Tablo 7 DVM Hata Oranı Sonuçları

Hata oranlarının sıfıra yakın olması destek vektörleri makine uygulamasının başarılı olduğunu göstermektedir.

Ay	2021 Gerçekleşen	DVRTahmin	Sapma
Ocak	737.985	578.200	159.785
Şubat	665.124	639.700	25.424
Mart	665.435	633.900	31.535
Nisan	690.074	562.800	127.274
Mayıs	583.149	678.600	95.451
Haziran	588.805	777.800	188.995
Temmuz	416.467	818.200	401.733
Ağustos	650.684	896.200	245.516
Eylül	695.787	924.800	229.013
Ekim	609.367	897.000	287.633
Kasım	719.240	929.600	210.360
Aralık	725.846	871.300	145.454

Tablo 8 DVM Tahmin Sonuçları

Tahmin sonuçları incelendiğinde ilk dört ay tahmin başarısının düşük olduğunu, son 2 çeyrek periyotta ise tahmin başarısının arttığı söylenebilir.



Şekil 27 DVM Gerçekleşen ve Tahmin Grafiği

DVM modeli incelendiğinde; özellikle yılın ilk çeyrek periyodunda tahmin başarısının yüksek olduğunu, devam eden dönemler de ise artış eğilimi gösterdiği ve gerçekleşen değerler ile arasında fark olduğu söylenebilir. DVM modelinin, doğrusal regresyona göre daha iyi sonuç verirken, YSA modeline göre daha az başarılı tahmin sonuçlarını ürettiğini söyleyebiliriz.

Gelinen aşamanın ardından, 2022 yılı tahminleri YSA, DVM ve regresyon modelleri için uygulanmıştır.

Tablo 9 2022 Tahmin Sonuçları*

Ay/Yıl	YSA	DVR	Regresyon
Ocak_2022*	712.100	875.200	1.043.200
Şubat_2022*	712.500	933.200	1.080.600
Mart_2022*	712.800	942.500	1.117.500
Nisan_2022*	711.300	1.006.600	1.154.100
Mayıs_2022*	710.700	1.094.300	1.189.400
Haziran_2022*	711.200	1.138.000	1.226.300
Temmuz_2022*	710.300	1.198.100	1.263.400
Ağustos_2022*	709.700	1.293.300	1.301.200
Eylül_2022*	709.600	1.222.500	1.341.200
Ekim_2022*	709.000	1.258.500	1.380.800
Kasım_2022*	708.700	1.261.400	1.416.200
Aralık_2022*	708.400	1.272.900	1.454.500

2021 yılı tahmin sonuçları incelendiğinde en iyi sonuç veren modelin YSA olduğu söylenebilir. Yapılan tahmin sonuçları 2022 yılı içinde uygulanmıştır. Tablo 10'da 2022 tahmin sonuçları paylaşılmıştır. YSA modelinin ortalamaya daha yakın seyrettiği, DVM ve regresyon modellerinin ise artan bir eğilimde olduğu söylenebilir. 2022 yılı tahminlerine bakıldığında, piyasa koşulları ve yapılan yıllık anlaşmalar değerlendirildiğinde, 2021 yılı verilerine yakın sonuçlara ulaşılacağı ön görülebilir. Bu durumda 2022 yılı için de YSA modelinin gerçeğe daha yakın olduğu söylenebilir.

8. Uygulama Sonuçlarının Değerlendirilmesi

Günümüzde veri madenciliği, önemi ve yaygınlığı giderek artan bir teknoloji olmaktadır. Özellikle büyük miktardaki veri yığınlarından üretilen stratejik bilgiler nedeniyle gelecekte de yaygın olarak kullanılabilmesi öngörülmektedir. Tahmin edici veri madenciliği teknikleri ise bir model oluşturarak veri kümelerinin davranışlarını tahmin etmeye çalışır. Tahmin önemli modeller çıkartmak için veya gelecekteki trendleri tahmin etmek için kullanılabilen bir veri analiz biçimidir. Bu tür analizler hem mevcut veri setini daha iyi anlamamıza hem de veri setinin gelecekteki davranışını tahmin etmede sıklıkla kullanılmaktadır. Gelecekle ilgili sayısal veriler ise içerisinde bulunduğumuz teknoloji çağında sağlıktan eğitime, finanstan pazarlamaya kadar her alanda oldukça önemlidir.

Hızla gelişen teknoloji son yıllarda faaliyetlerin daha etkin bir biçimde yürütülmesini sağlamaktadır. Yapay zekânın gelişimi ile öğrenme anlamında da birçok yenilik yaşanmış ve çalışmaların ileriye taşınmasına sebep olmuştur. Talep değişkenliğinin olduğu koşullarda yapay zekâ uygulamalarının daha iyi sonuç verdiği son yıllarda yapılan çalışmalar incelendiğinde söylenebilir.

Bu çalışmada açık kaynak kodlu veri madenciliği programı olan WEKA üzerinde gerçekleştirilmiştir. WEKA yazılımı, açık kaynak kodlu, genel kamu lisansına sahip ve kolay uygulanabilir olmasından dolayı tercih edilmiştir. Yapay sinir ağı, doğrusal regresyon ve destek vektörleri makine öğrenmesi yöntemleri ile 2021 yılına ait satış adetleri tahmin edilmiştir. Tahmin performansları, ortalama mutlak hata (MAE), kök ortalama kare hatası (RMSE), ortalama mutlak yüzde hata (MAPE) ölçütlerine göre kıyaslanmış ve bu yöntemler performans ölçülerine göre değerlendirilmiştir. Yapılan tahmin sonuçları incelendiğinde; YSA ile yapılan tahminlerin doğruluk oranlarının DVM ve Doğrusal Regresyon sonuçlarına göre daha yüksek olduğu görülmüştür. YSA ampirik risk minimizasyonu yerine, yapısal risk minimizasyonu prensibini benimsenmesi nedeniyle, veri setine daha iyi uyum sağlayarak daha başarılı sonuçlar elde ettiği söylenebilir.

	YSA	Regresyon	DVR
MAPE	0,1285	0,2581	0,1987
MAE	0,0446	2,6200	0,0601
RMSE	0,0610	3,4500	0,0812
Korelasyon	0,808	0,9801	0,7741

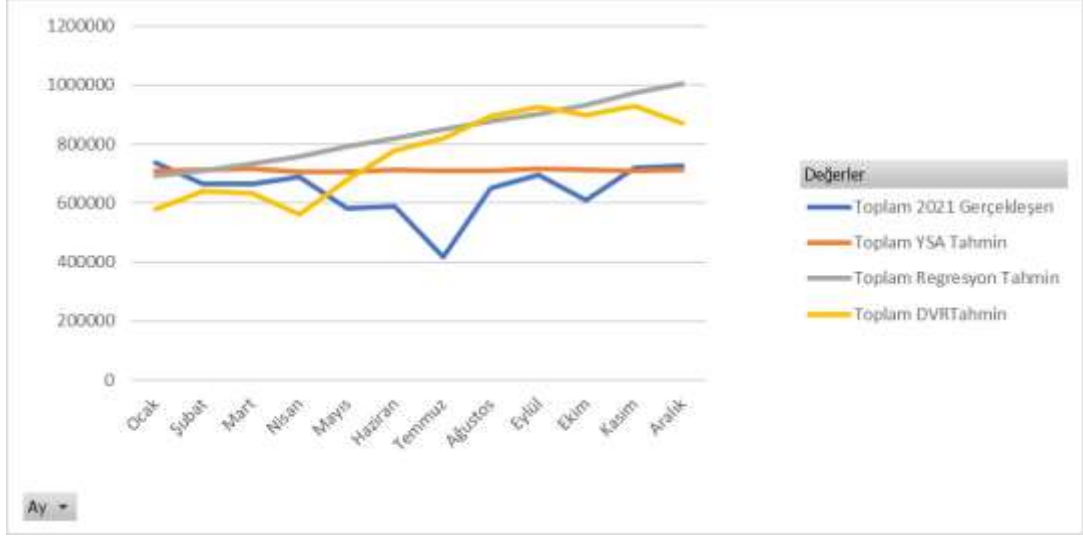
Tablo 10 Hata Oranı Karşılaştırması

Ortalama standart hatası sifira en yakın olan tahmin sonucu en başarılı tahmin olarak değerlendirilmektedir. MAPE, değerinin %10'nun altında olması tahminlerin "yüksek doğruluk" seviyesinde olduğu söylenebilir. MAE, 0'dan ∞ 'a kadar değer alabilir. Ancak, daha düşük değerlere sahip tahminleyiciler daha iyi performans gösterir. RMSE, 0'dan ∞ 'a kadar değer alabilir. Ancak, daha düşük değerlere sahip tahminleyiciler daha iyi performans gösterir. Bu durumda veri setimiz için, YSA analizinin daha başarılı bir tahmin yöntemi olduğunu söyleyebiliriz. Korelasyon katsayılarına bakıldığında ise en yüksek sonucu veren analizin doğrusal regresyon olduğu görülmektedir.

Ay	2021 Gerçekleşen	YSA Tahmin	Regresyon Tahmin	DVRTahmin
Ocak	737.985	708.200	692.100	578.200
Şubat	665.124	714.400	708.200	639.700
Mart	665.435	715.900	731.800	633.900
Nisan	690.074	706.900	758.400	562.800
Mayıs	583.149	706.800	792.300	678.600
Haziran	588.805	714.100	818.800	777.800
Temmuz	416.467	710.500	848.800	818.200
Ağustos	650.684	707.900	877.700	896.200
Eylül	695.787	714.800	901.900	924.800
Ekim	609.367	713.700	931.700	897.000
Kasım	719.240	710.500	975.000	929.600
Aralık	725.846	712.700	1.005.900	871.300

Tablo 11 Tahmin Karşılaştırmaları

Test verisi ile gerçekleşen değerler Tablo 12'de gösterilmiştir. Gerçekleşen değerler ile tahmin değerleri kıyaslandığında en az sapmaya sahip ve hata oranı en düşük olan YSA'nın verdiği söylenebilir. Tablo sonucuna göre, veri seti arttıkça YSA'nın tahmin başarısının daha fazla yükseleceği söylenebilir.



Şekil 28 Tahmin Sonuç Grafiği

Grafik incelendiğinde, YSA tahminin Kasım ve Aralık ayında gerçekleşen değerler ile neredeyse aynı olduğu görülmektedir. Bu durumda, 2022 için yapılan tahminin doğru olacağını bizlere ispatlar niteliktedir.

Tahmin geleceğin verilerinin bugünden kestirilmesi işlemidir. Firmaların hepsi geleceği önden görerek planlamalarını bu doğrultuda yapmak istemektedirler. Bir firmaya ait satış adetleri tahmin edildiğinde, gelecekte yapılacak müşteri talepleri tahmin edilmiş olmaktadır. Yapılan analiz baz alınarak şirket bir sonraki yıl için kendini hazırlayabilecek ve bu doğrultuda makine ekipman yatırımı, hammadde siparişleri, personel ihtiyacı vb. konularda yatırım gerçekleştirme fırsatı bulabilecektir. Veri seti daha fazla genişletilerek ve farklı analizlerle tahmin performansı tekrar ölçülerek modeller genişletilebilir.

SONUÇ

Ekonominin deęişken yapısına ayak uydurabilmek için istatistiksel yöntemler kullanılarak tahmin etme günümüzde gittikçe önem kazanan bir yaklaşımdır. Çalışmada tahmin yöntemi olarak Yapay Sinir Ağları, çoklu doğrusal regresyon ve Destek Vektör Makinaları tercih edilmiştir. Yapılan analizin sonucunda en başarılı tahmin yönteminin Yapay Sinir Ağları olduğu ortaya konmuştur. YSA modellerinin kullanımının kolay ve yapısının esnek olması hatalar karşısında performans düşüklüğü göstermemesi sebebi ile, iyi bir model olduğu ortaya çıkmaktadır. Ülkemizde yapılan Yapay Sinir Ağları çalışmaları her geçen gün artıyor olsa da, hala klasik yöntemlere göre daha az tercih edilmektedir. Uygulamada kullanılan klasik regresyon modeli, hata oranlarında düşük sonuçlar gösterse de tahmin edilen değerler konusunda büyük sapmalarla sonuçlanmıştır. Gelişen teknolojiyle birlikte ortaya çıkan yeni modellerin, klasik yöntemlere göre daha iyi sonuçlar elde ettiği yapılan çalışma ile bir kez daha kanıtlanmıştır. Destek vektörleri makinaları ile yapılan tahmin hata oranlarıyla ve tahmin edilen değerler ile, klasik yöntem olan çoklu doğrusal regresyona göre daha iyi sonuçlar ortaya koymuştur. Ancak araştırılan üç yöntem içerisinde en başarılı hata oranlarını veren model Yapay Sinir Ağlarıdır. Ayrıca YSA modeliyle birlikte yapılan tahmin, gerçek değerler ile çok az sapma göstermektedir. Yapılan çalışmayla birlikte elde edilen sonuçlar, Yapay Sinir Ağlarının iyi bir tahminci olduğunu bir kez daha kanıtlar niteliktedir. Yapılan çalışma ile YSA kullanım konusunda literatüre bir katkı sağlanmış olmaktadır.

Çalışmanın başlıca amacı; firmanın üretim planlaması yaparken, stratejik olarak düşünmesini sağlamaktır. Yapılan çalışma firmanın; insan, ekipman ve kaynak önceden oluşturmasına yardımcı olmaktadır. Yapılan tahmin sonucunda elde edilen bilgiler; üreticiler, araştırmacılar ve pazarlamacılara yol gösterici olacaktır. Satış adetlerinin tahmin edilmesiyle birlikte firma, üretim maliyetlerini düşürmek için düşük hammadde maliyeti ile üretim sağlayarak stoklu ilerleme imkanına sahip olacaktır. Böylece ülkemizde örnekleri görülmeye başlayan lisanslı depoculuk faaliyetlerine geçiş yapılması da sağlanabilecektir. Ortaya konulan sonuçlarla birlikte, rekabetçi ortamlar yaratılabilecek ve müşteri sayısını artırma yolunu seçebileceklerdir. Elde edilen sonuçlar

ışığında üretim kapasitesinin ne kadar arttırılabileceği değerlendirilebilecektir. Bu da farklı müşteri arayışlarını beraberinde getirerek üretim sayısının ve dolayısıyla da satış adetlerinin artmasını beraberinde getirecektir. Tüm bu sonuçları değerlendirdiğimizde; firma minimum üretim maliyeti ve maksimum kâr hedefleyerek varlığını sürdürebilme imkânı yakalayabilecektir.

- 1- Ekonominin değişken yapısına ayak uydurabilmek için istatistiksel yöntemler kullanılarak tahmin etme günümüzde gittikçe önem kazanan bir yaklaşımdır. Çalışmada tahmin yöntemi olarak yapay sinir ağları, çoklu doğrusal regresyon ve destek vektör makinaları tercih edilmiştir. Yapılan analizin sonucunda en başarılı tahmin yönteminin yapay sinir ağları olduğu ortaya konmuştur. YSA modellerinin kullanımının kolay ve yapısının esnek olması hatalar karşısında performans düşüklüğü göstermemesi sebebi ile, iyi bir model olduğu ortaya çıkmaktadır. Uygulamada kullanılan klasik regresyon modeli, hata oranlarında düşük sonuçlar gösterse de tahmin edilen değerler konusunda büyük sapmalarla sonuçlanmıştır. Destek vektörleri makinaları ile yapılan tahmin hata oranlarıyla ve tahmin edilen değerler ile, klasik yöntem olan çoklu doğrusal regresyona göre daha iyi sonuçlar ortaya koymuştur. Gelişen teknolojiyle birlikte ortaya çıkan yeni modellerin, klasik yöntemlere göre daha iyi sonuçlar elde ettiği yapılan çalışma ile bir kez daha kanıtlanmıştır.
- 2- Ülkemizde yapılan yapay sinir ağları çalışmaları her geçen gün artıyor olsa da hala klasik yöntemlere göre daha az tercih edilmektedir. YSA modeliyle birlikte yapılan tahmin, gerçek değerler ile çok az sapma göstermektedir. Yapılan çalışmayla birlikte elde edilen sonuçlar, yapay sinir ağlarının iyi bir tahminci olduğunu bir kez daha kanıtlar niteliktedir. Yapılan çalışma ile YSA kullanım konusunda literatüre bir katkı sağlanmış olmaktadır.
- 3- Çalışmanın başlıca amacı; firmanın üretim planlaması yaparken, stratejik olarak düşünmesini sağlamaktır. Yapılan çalışma firmanın; insan, ekipman ve kaynak önceden oluşturmasına yardımcı olmaktadır. Yapılan tahmin sonucunda elde

edilen bilgiler; üreticiler, arařtırmacılar ve pazarlamacılaraya yol gösterici olacaktır. Satıř adetlerinin tahmin edilmesiyle birlikte firma, üretim maliyetlerini düşürmek için düşük hammadde maliyeti ile üretim sağlayarak stoklu ilerleme imkanına sahip olacaktır. Böylece ülkemizde örnekleri görülmeye başlayan lisanslı depoculuk faaliyetlerine geçiř yapılması da sağlanabilecektir.

- 4- Ortaya konulan sonuçlarla birlikte, rekabetçi ortamlar yaratılabilecek ve müşteri sayısını arttırma yolunu seçebileceklerdir. Elde edilen sonuçlar ışığında üretim kapasitesinin ne kadar arttırılabileceği değerlendirilebilecektir. Bu da farklı müşteri arayışlarını beraberinde getirerek üretim sayısının ve dolayısıyla da satıř adetlerinin artmasını beraberinde getirecektir.
- 5- Yapılan çalışma, beyaz eşya sektöründe faaliyet gösteren diğer firmalara da kaynak niteliğinde olacaktır. Bu sayede firmalar minimum üretim maliyeti ve maksimum kâr hedefleyerek varlığını sürdürebilme imkânı yakalayabilecektir.

EKLER

Yıl	Sevkedeti	RKGE	GSYH	SÜE	Evlenmeorani	TGE	EGE
1.01.2016	251801	106,50	628838414	102,6	28552	92,20	127,0
1.02.2016	250134	105,10	628838414	103,7	30163	90,00	124,0
1.03.2016	270707	104,60	628838414	102,6	33464	90,10	126,0
1.04.2016	275353	105,30	647310685	102,8	16657	91,30	120,0
1.05.2016	285102	105,00	647310685	104,5	11666	91,50	121,0
1.06.2016	294857	104,10	647310685	104,0	46387	91,60	119,0
1.07.2016	168289	105,70	635092117	97,7	72402	88,90	121,0
1.08.2016	291938	101,70	635092117	105,0	77086	94,70	116,0
1.09.2016	286570	106,60	635092117	102,0	62256	94,80	119,0
1.10.2016	263464	104,90	715318494	106,5	45425	95,30	119,0
1.11.2016	353122	108,20	715318494	105,7	35476	91,60	119,0
1.12.2016	222326	103,90	715318494	105,0	27736	87,40	120,0
1.01.2017	293219	100,60	726280402	105,9	27993	88,70	115,0
1.02.2017	296781	106,20	726280402	107,5	27237	88,00	117,0
1.03.2017	326307	106,50	726280402	108,9	33770	90,10	117,0
1.04.2017	315711	106,30	756765449	111,9	53155	92,00	120,0
1.05.2017	238141	104,80	756765449	110,5	30376	94,00	122,0
1.06.2017	275594	108,80	756765449	111,2	55453	92,30	121,0
1.07.2017	191972	107,70	797966523	110,9	66932	92,40	126,0
1.08.2017	289890	110,20	797966523	114,4	66987	93,00	125,0
1.09.2017	346128	111,20	797966523	115,8	66783	92,00	124,0
1.10.2017	357176	112,20	852691894	116,3	47467	89,90	123,0
1.11.2017	367285	109,80	852691894	117,0	36485	87,60	119,0
1.12.2017	242970	109,20	852691894	120,1	29676	88,20	120,0
1.01.2018	397423	110,90	883302215	117,8	29688	92,70	124,0
1.02.2018	407965	110,80	883302215	117,1	28817	93,40	123,0
1.03.2018	442885	109,50	883302215	117,1	35854	92,50	121,0
1.04.2018	302219	106,80	918474657	118,5	53371	91,70	119,0
1.05.2018	320378	106,70	918474657	116,8	43025	90,70	116,0
1.06.2018	205453	102,50	918474657	113,5	45394	90,60	113,0
1.07.2018	220248	101,50	985956214	116,2	70261	92,10	115,0
1.08.2018	315529	96,30	985956214	115,8	66131	88,20	109,0
1.09.2018	360383	90,40	985956214	112,1	69761	81,10	104,0
1.10.2018	429209	91,10	971040640	109,8	49017	78,80	102,0
1.11.2018	487459	96,80	971040640	109,7	35191	81,20	107,0
1.12.2018	302769	97,70	971040640	108,1	27879	80,10	115,0

1.01.2019	402442	95,40	1024701223	109,7	28907	80,50	111,0
1.02.2019	381039	96,90	1024701223	111,6	28726	79,20	109,0
1.03.2019	370104	99,30	1024701223	114,9	37376	81,30	109,0
1.04.2019	360500	100,00	1062351222	114,1	50139	83,60	109,0
1.05.2019	429611	94,70	1062351222	115,3	58513	76,90	109,0
1.06.2019	281901	99,60	1062351222	109,7	31119	79,80	113,0
1.07.2019	345221	96,60	1094399041	114,8	81392	78,30	110,0
1.08.2019	339585	102,10	1094399041	111,1	67723	79,10	113,0
1.09.2019	430147	99,70	1094399041	115,7	66392	77,70	117,0
1.10.2019	365925	104,20	1136335423	114,2	52963	78,50	122,0
1.11.2019	453812	105,90	1136335423	115,0	36949	81,30	120,0
1.12.2019	190614	108,70	1136335423	117,7	29260	80,70	123,0
1.01.2020	512190	106,40	1183461557	117,3	29668	81,40	125,0
1.02.2020	494562	106,70	1183461557	120,4	33369	79,60	122,0
1.03.2020	493764	98,60	1183461557	113,2	38259	81,10	121,0
1.04.2020	250918	62,30	1081436357	79,1	50778	78,10	78,0
1.05.2020	210352	73,50	1081436357	94,3	64742	82,70	88,0
1.06.2020	533711	89,80	1081436357	109,8	40840	82,70	104,0
1.07.2020	546963	99,40	1344106482	120,6	67398	82,30	116,0
1.08.2020	600942	105,20	1344106482	122,6	77027	79,40	115,0
1.09.2020	697508	105,70	1344106482	125,4	63643	82,00	115,0
1.10.2020	705419	109,70	1437878911	126,1	57559	81,90	121,0
1.11.2020	625974	107,40	1437878911	127,3	39919	80,10	122,0
1.12.2020	695779	110,40	1437878911	129,0	31291	80,10	109,0

KAYNAKÇA

- Aksel, F. (2000). Regresyon Analizi ve Yapay Sinir Ağı Yöntemleri ile Uzun Dönem Yük Tahmini, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Arı, Ayşe & Berberler Murat E. (2017). Yapay Sinir Ağları ile Tahmin ve Sınıflandırma Problemlerinin Çözümü İçin Arayüz Tasarımı, Dokuz Eylül Üniversitesi, İzmir.
- Azadeh, A. & Neshat, N. & Mardan, E. & Saberi, M. (2013). Optimization of steel demand forecasting with complex and uncertain economic inputs by an integrated neural network–fuzzy mathematical programming approach. The International Journal of Advanced Manufacturing Technology.
- Bidgoli, B. M. (2004). Data Mining For a Web Based Educational System, Ph. D. Thesis, Department of Computer Science and Engineering, Michigan State University.
- Carlson, RL and Umble, M., (1980), "Statistical demand functions for automobiles and their use for forecasting in an energy crisis" The Journal of Business.
- Catal, C. & Ece, K. & Arslan, B. & Akbulut, A. (2019). Benchmarking of Regression Algorithms and Time Series Analysis Techniques for Sales Forecasting. Balkan Journal of Electrical and Computer Engineering.
- Coenen, F. (2004). Data Mining: Past, Present and Future, The Knowledge Engineering Review Vol. 00:0 1-24 Cambridge University Press.

Çayırođlu, İ. (2015). Yapay Sinir Ağları, Yüksek Lisans Tezi, Karabük Üniversitesi, Karabük.

Çelik, M. (2009). Veri Madenciliğinde Kullanılan Sınıflandırma Yöntemleri ve Bir Uygulama, Yüksek Lisans Tezi, İstanbul.

Çomak, E. (2008). Destek Vektör Makinelerinin Etkin Eğitimi için Yeni Yaklaşımlar, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.

Demirbaş F. Pınar (2011). Kombi Üretiminde Talep Tahmin Yöntemlerinin Uygulanması, Yüksek Lisans Tezi, Kocaeli Üniversitesi, Kocaeli.

Demirel, Ö. (2009). ANFIS ve ARMA modelleri ile elektrik enerjisi yük tahmini, Yüksek Lisans Tezi, Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 1-26.

Dunham, (2003). Margaret H. Data Mining Introductory and Advanced Topics. New Jersey.

Durmaz, C. & Kocamış, M. (2008). Oracle Data Miner ile Öğrenci Kayıtları Üzerine Bir Veri Madenciliği Uygulaması, İstanbul Teknik Üniversitesi, İstanbul.

Ecemiş, O. (2018). Model ağaç yöntemiyle satış tahmini: Paslanmaz çelik sektöründe bir uygulama. Akademik Sosyal Araştırmalar Dergisi, 6(84), 336-350.

Efendigil, T. & Önüt, S. & Kahraman, C. (2009). A decision support system for demand

forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications*, 36(3), 6697-6707.

Elmas, Çetin (2003). *Yapay Sinir Ağları (Kuram, Mimari, Eğitim, Uygulama)*, Seçkin Yayıncılık, Ankara.

Fausett, (1994). Laurene, *Fundamentals of Neural Networks*, Prentice Hall International, London.

Gavcar, E. & Sen, S. & Aytakin, A. (1999). Türkiye’de Kullanılan Kağıt- Karton Türlerinin Talep Tahminlerinin Belirlenmesi, *Tr. Journal of Agriculture and Forestry*, Tübitak.

Gorunescu, F. (2011). *Data Mining Concepts, Models And Techniques*, Berlin, Almanya.

Güriş, S. & Çağlayan, E. (2010). *Ekonometri Temel Kavramlar*. Der Yayıncılık, İstanbul.

Güvenç, E. (2001) *Yüksek Öğretimde Öğrenci Performansının Veri Madenciliği Teknikleri ile Belirlenmesi*, Yüksek Lisans Tezi, Boğaziçi Üniversitesi, İstanbul

Hamzaçelebi, C. (2011). *Yapay Sinir Ağları*. Ekin Kitapevi, İstanbul.

Han, J. & M. Kamber.(2001). *Data Mining*. San Francisco:Morgan Kaufmann Publishers.

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation [Elektronik Sürüm], Prentice Hall Inc, New Jersey.

Hicham, A. & Bouhorma, M. (2012). Hybrid intelligent system for sale forecasting using delphi and adaptive fuzzy back-propagation neural networks. (IJACSA) International Journal of Advanced Computer Science and Applications, 3(11), 122-130.

Jordan, M., Kleinberg, J., & Scholkopf, B. (2008). Support Vector Machines. Information Science and Statistics. New York: Springer.

Kabalıcı, E. (2014). Yapay Sinir Ağları. Ders Notlar
<https://ekblc.files.wordpress.com/2013/09/ysa.pdf>

Kalaycı, Ş. (2010), SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil Yayın Dağıtım Ltd. Şti., Ankara.

Karaatlı Meltem & Helcacioğlu Ceyda & Ömürbek Nuri & Toköz Gönül (2012).
Yapay Sinir Ağları ile Otomobil Satış Tahmini, Uluslararası Yönetim İktisat ve İşletme Dergisi, Cilt 8, Sayı 17, 2012.

Karahan, Mehmet (2011). İstatistiksel Tahmin Yöntemleri: Yapay Sinir Ağları Metodu ile Ürün Talep Tahmini Uygulaması, Doktora Tezi, Selçuk Üniversitesi, Konya.

Kargı, V. Sinem, (2013). Yapay Sinir Ağ Modelleri ve Bir Tekstil Firmasında Uygulama, Doktora Tezi, Uludağ Üniversitesi, Bursa.

Kılıç, Fatih vd. (2018), Yemekhane için Yapay Zeka Teknikleri Kullanımı ile Günlük

Talep Tahmini, Adana.

Koochakpour, K. & Tarokh, M. J. (2016). Sales budget forecasting and revision by adaptive network fuzzy base inference system and optimization methods. *Journal of Computer & Robotics*, 9(1), 25-38.

Kurnaz, K. (2014). Yapay Sinir Ağları İle Makine Öğrenmesi Uygulaması, Bitirme Tezi

Larose TD. 2005. *Discovering knowledge in data*. Wiley Interscience.

Lewis & Colin D. (1982), *Industrial and Business Forecasting Methods*. Butterworths Publishing: London.

Lin, C. F. & Wang, S. D. (2002). Fuzzy Support Vector Machines. *IEEE Transactions on Neural Networks*, 13(2), 464-471.

Lu, C. J. & Lee, T. S. & Chiu, C. C. (2009). Financial Time Series Forecasting Using Independent Component Analysis And Support Vector Regression. *Decision Support Systems*, 47(2), 115-125.

Oğuzlar, Ayşe (2003). “Veri Ön İşleme”, Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Sayı: 21.

Oğuzlar, A. (2004) *Veri Madenciliğine Giriş*, Ekin Kitabevi, Bursa.

Öztemel, E. (2012). *Yapay Sinir Ağları*, Papatya Yayıncılık, İstanbul.,

Pavlyshenko, B. M. (2019). Machine-Learning Models for Sales Time Series Forecasting. *Data*, 4(1), 1–11. doi: <https://doi.org/10.3390/data4010015>

Poyraz, İlker (2020). Makine Öğrenmesi Algoritmaları ile Eczaneler için İlaç Talep Tahmini, Yüksek Lisans Tezi, İstanbul Aydın Üniversitesi, İstanbul.

Reynolds, G.W. (1992). *Information System for Managers*. West Publishing Company, St. Paul.

Richard & P. Lippmann, (1987). An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine*.

Rincon-Patino, J. & Lasso, E. & Corrales, J. C. (2018). Estimating Avocado Sales Using Machine Learning Algorithms and Weather Data. *Sustainability*, 10(10), 12. doi: <https://doi.org/10.3390/su10103498>

Sağıroğlu, Ş. & Beşdok, E. & Erler, M. (2004). Mühendislikte Yapay Zeka uygulamaları Yapay Sinir Ağları , *Ufuk Kitap Kirtasiye*, Kayseri.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, <https://doi.org/10.1147/rd.441.0206>

Saygılı, Y. S. (2008). İstatistiksel Yöntemlerle Yapay Sinir Ağları Uygulamalarının Karşılaştırılması: Millî Savunma Bakanlığı Bütçesinin Öngörülenmesi, Yüksek Lisans Tezi, Kara Harp Okulu Savunma Bilimleri Enstitüsü, Ankara.

Sevüktekin, Mustafa (2013). Ekonometriye Giriş, Dora Basım Yayın, Bursa.

Silahtaroglu, G. (2013). Veri Madenciliği Kavram ve Algoritmaları. 2.Baskı. İstanbul: Papatya Yayıncılık Eğitim A.Ş.

Stephan Kudyaba & R. Hoptroff, Data Mining and Business Intelligence (USA: Idea Group, 2001, s.8.)

Suykens, J. A. & Vandewalle, J. (1999). Least Squares Support Vector Machine Classifiers. Neural processing letters.

Tekin, M. (2009). Üretim Yönetimi Cilt 1, 6. Baskı, Günay Ofset, Konya

Tozak, Esra Y. (2021). Veri Madenciliği Programları Kullanılarak Bir Tekstil Firmasının Satış Verilerinin Değerlendirilmesi, Yüksek Lisans Tezi, Karabük.

Türk, Emre & Kıani Farzad (2019). Yapay Sinir Ağları ile Talep Tahmini Yapma: Beyaz Eşya Üretim Planlaması için YSA Uygulaması, İstanbul.

Tütek H. & Gümüsoğlu, Ş. (2008). İşletme İstatistiği. Beta Basım Yayım Dağıtım A.Ş. İstanbul.

Tüzüntürk, Selim (2010). Veri Madenciliği ve İstatistik, İktisadi ve İdari Bilimler Fakültesi Dergisi, Uludağ Üniversitesi, Bursa.

- Tüzüntürk, Selim & Eteman S. Fatma & Sezen, Kemal (2016). Yapay Sinir Ağı Yöntemi ile Damacana Su Satış Miktarlarının Tahmini, Uluslararası Hakemli Sosyal Bilimler E-Dergisi, İktisat ve Girişimcilik Üniversitesi, Kırgızistan.
- Venkatadri M. & Reddy L. C. (2011). “A Review on Data Mining from Past to Future”, International Journal of Computer Applications, Volume 15 No. 7, 19 – 22.
- Yeğen, N. (2020). Perakende Sektöründe Veri Madenciliği ile Satış Tahmini, Bahçeşehir Üniversitesi, İstanbul
- Yılmaz, Hilal (2018). Antalya ili Konut Satışının Çoklu Doğrusal Regresyon ve Yapay Sinir Ağları ile Tahmini, Yüksek Lisans Tezi, Akdeniz Üniversitesi, Antalya.
- Yücesoy, Mihriban (2011), Temizlik Kağıtları Sektöründe Yapay Sinir Ağları ile Talep Tahmini, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, İstanbul.
- Zhang, N. & Xiong, J. & Zhong, J. & Leatham, K. (2018). Gaussian process regression method for classification for high- dimensional data with limited samples. In 2018 Eighth International Conference on Information Science and Technology (ICIST) (pp. 358-363). IEEE