T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞLARI İLE BİR TELEKOM FİRMASI MÜŞTERİLERİNE AİT MÜŞTERİ KAYBI ANALİZİ

Pelin Su TÜFEKÇİ

YÜKSEK LİSANS TEZİ

İstatistik Anabilim Dalı

İstatistik Bölümü

Danışman

Prof. Dr. Ali Hakan BÜYÜKLÜ

Haziran, 2021

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞLARI İLE BİR TELEKOM FİRMASI MÜŞTERİLERİNE AİT MÜŞTERİ KAYBI ANALİZİ

Pelin Su TÜFEKÇİ tarafından hazırlanan tez çalışması 07.06.2021 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstatistik Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof. Dr. Ali Hakan BÜYÜKLÜ
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof. Dr. Ali Hakan BÜYÜKLÜ
Yıldız Teknik Üniversitesi _____

Prof. Dr. Semra ERBOLAT TAŞABAT
Mimar Sinan Güzel Sanatlar Üniversitesi _____

Doç. Dr. Öyküm Esra YİĞİT
Yıldız Teknik Üniversitesi _____

Danışmanım Prof. Dr. Ali Hakan BÜYÜKLÜ sorumluluğunda tarafımca hazırlanan Yapay sinir ağları ile bir telekom firması müşterilerine ait müşteri kaybı analizi başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Pelin Su TÜFEKÇİ

Aileme

TEŞEKKÜR

Yüksek lisans tez çalışmam boyunca beni aydınlatan tez danışmanı hocam sayın Prof. Dr. Ali Hakan Büyüklü'ye, hem tez çalışmam boyunca hem de yüksek lisans eğitimim boyunca her konuda desteğini esirgemeyen sayın hocam Coşkun Parim'e, hayatım boyunca tüm kararlarım ve çalışmalarımda beni her zaman destekleyen sevgili anneme, yanımızda olmasa da her zaman varlığını hissettiğim sevgili babama ve her zaman yanımda olan değerli dostlarıma teşekkürü borç bilirim.

Pelin Su TÜFEKÇİ

İÇİNDEKİLER

SİMGE LİSTESİ vii

ŞEKİL LİSTESİ viii

TABLO LİSTESİ ix

ÖZET x

ABSTRACT xi

1 GİRİŞ 1

1.1 Literatür Özeti 1

1.2 Tezin Amacı 2

1.3 Hipotez 3

2 ARAŞTIRMA 4

2.1 Veri Madenciliği 4

2.1.1 Veri Madenciliği Adımları 4

2.2 Müşteri Geri Kazanımı 6

2.2.1 Müşteri Geri Kazanımı: Tanım ve Kapsamı 7

2.2.2 Müşteri Geri Kazanımı: Önem 8

2.3 Yapay Sinir Ağları 9

2.3.1 Yapay Sinir Ağlarının Tarihçesi 9

2.3.2 Yapay Sinir Ağlarının Tanımı ve Özellikleri 11

2.3.3. Yapay Sinir Ağlarının Formülizasyonu 16

2.4 Sınıflandırma Performanslarının Ölçülmesi 20

2.4.1 Akaike Bilgi Kriteri ve Bayesci Bilgi Kriteri 22

2.4.2 Ham Akaike Bilgi Kriteri (AIC) 22

2.4.3 Normalleştirilmiş Akaike Bilgi Kriteri (AICc) 22

2.4.4 Bayesci Bilgi Kriteri (BIC) 22

2.5 Öznitelik Seçimi 22

2.5.1 T-testi ile öznitelik seçimi 23
2.6 Lojistik Regresyon 23
2.6.1 İkili (Binary) Lojistik Regresyon 24
2.6.2 Basit Lojistik Regresyon 24
2.6.3 Lojistik Regresyon Modeli 25
3 UYGULAMA 27
3.1 Uygulama Yöntemi ve Kullanılan Modeller 27
3.1.1 Trainscg - Ölçekli Eşlenik Gradyan Geri Yayılım 27
3.1.2 Trainbfg - BFGS Yarı-Newton Geri Yayılımı 28
3.1.3 Traingd - Gradyan Azalan Geri Yayılım 28
3.1.4 Traingda - Uyarlanabilir öğrenme Oranı ile Geri Yayılımlı Gradyan Azalışı 29
3.1.5 Traingdx - Momentum ve Uyarlanabilir Öğrenme Oranı Geri Yayılımı ile Gradyan Azalışı 30
3.1.6 Trainr - Öğrenme İşlevleriyle Rastgele Sıralı Artımlı Eğitim 31
3.1.7 Trainb - Ağırlık ve Sapma Öğrenme Kurallarıyla Toplu Eğitim 31
3.2 Kullanılan Değişkenler 31
3.3 T-Testi ile Değişken Seçimi 32
3.4 Yapay Sinir Ağları Modellerinin Uygulanması 36
3.5 Lojistik Regresyon Modelinin Uygulanması 39
3.6 En İyi Modele Ait Detaylar 39
4 SONUÇ VE ÖNERİLER 43
KAYNAKÇA 45
TEZDEN ÜRETİLMİŞ YAYINLAR 48

SİMGE LİSTESİ

w_{ij}	Ağırlık katsayıları
φ	Aktivasyon fonksiyonu
x_i	Girdi değeri
O_i	Gözlenen değer
p_o	Gözlenen uyum oranı
κ	Kappa istatistiği
p_e	Rastgele beklenen uyumluluk oranı
$\beta_1, \beta_2, \dots, \beta_n$	Regresyon katsayıları
β_0	Regresyon Sabiti
σ	Standard sapma
p_i	Tahminlenen değer

ŞEKİL LİSTESİ

Şekil 2.1 Biyolojik Sinir Hücresi ve Yapay Sinir Ağı 11
Şekil 2.2 Yapay Sinir Hücresi 12
Şekil 2.3 Tek Katmanlı Algılayıcı Modeli 15
Şekil 2.4 Çok Katmanlı Algılayıcı Modeli 15
Şekil 2.5 İleri Beslemeli Ağ Yapısı 16
Şekil 2.6 Geri Beslemeli Ağ Yapısı 16
Şekil 2.7 Aktivasyon Fonksiyonu Örneklerinin Grafikleri 19
Şekil 3.1 Hesap Uzunluğu ve Gün İçi Konuşma Kutu Grafikleri 34
Şekil 3.2 Akşam Konuşma ve Gece Konuşma Kutu Grafikleri 34
Şekil 3.3 Uluslararası Plan, Dakika ve Konuşma Kutu Grafikleri 34
Şekil 3.4 Sesli Mesaj Plan, Mesaj Sayısı ve Gün İçi Dakika Kutu Grafikleri 35
Şekil 3.5 Gün İçi Konuşma, Akşam Konuşma ve Akşam Ücret Kutu Grafikleri 35
Şekil 3.6 Gece Konuşma Dakika, Ücret ve Müşteri Hizmetleri Arama Kutu Grafikleri 35
Şekil 3.7 Uluslararası Ücret Kutu Grafiği 35
Şekil 3.8 En İyi Modelin Hata Matrisleri 40
Şekil 3.9 En İyi Modelin ROC Eğrileri 41
Şekil 3.10 En İyi Modelin Hata Histogramı 42

TABLO LİSTESİ

Tablo 2.1 Biyolojik Sinir Sistemi Kısımlarının Yapay Sinir Sistemindeki Karşılığı 11
Tablo 3.1 Kullanılan Değişkenler ve Açıklamaları 31
Tablo 3.2 T-testi Sonuçları 32
Tablo 3.3 Yapay Sinir Ağları Model Çıktıları 37
Tablo 3.4 Lojistik Regresyon Model Çıktıları 39

ÖZET

Yapay Sinir Ağları ile Bir Telekom Firması Müşterilerine Ait Müşteri Kaybı Analizi

Pelin Su TÜFEKÇİ

İstatistik Anabilim Dalı

Yüksek Lisans Tezi

Danışman: Prof. Dr. Ali Hakan BÜYÜKLÜ

Günümüzde gelişen teknoloji ile firmalar müşterilerinin verilerini toplayarak ve analiz ederek onlar hakkında çok daha fazla bilgiye sahip olmaya başlamışlardır. Verilerini analiz ettikleri insanları hem potansiyel müşteri olarak belirleyebilir, onlara özel ürünler önerebilir hem de mevcut müşterilerinin sadakatini arttırıcı çalışmalar yapabilirler. Müşterilerinin verilerini doğru modelleyebilen firmalar, üyeliklerini iptal edip artık onları kullanmayacak müşterileri de önceden tespit ederek bu kişilere özel çalışmalar yapabilirler.

Bugün herkes en çok vakit geçirdiği ve neredeyse tüm verilerini barındırdığı cep telefonları için bir operatör kullanmak zorundadır. Bu yüzden de telekom firmaları en fazla müşteri verisine sahip firmaların başında gelirler.

Bu tez çalışmasında, veri madenciliği kullanılarak bir telekom firmasına ait müşterilerin kayıp analizleri yapılmıştır. Yapay sinir ağları yöntemi ile yapılan bu çalışmada telekom firmasının hangi müşterilerini kaybedeceğinin tahmin edilebilmesi hedeflenmiştir.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ABSTRACT

Customer Churn Analysis of a Telecom Company By Using Artificial Neural Network

Pelin Su TÜFEKÇİ

Department of Statistics

Master of Science Thesis

Supervisor: Prof. Dr. Ali Hakan BÜYÜKLÜ

With today's developing technology, companies began to learn more about their customers by collecting and analyzing their data. They can identify the people they analyze data as potential customers, offer them special products, and work to increase the loyalty of their existing customers. Companies that can model their customers' data correctly can identify customers who will cancel their membership and no longer use them, and can work specifically for them. Today, everyone should use an operator for mobile phones. Therefore, telecom companies are among the companies with the most customer data. In this thesis, churn analyzes of customers of a telecom company were made using data mining. In the study conducted with artificial neural networks method, it was aimed to predict which customers of the telecom company will be lost.

Keywords: Churn analysis, artificial neural network, data mining.

1
GİRİŞ

Geçmişten günümüze tüm zamanlarda ticari işletmelerin ana amacı her zaman karlarını artırmak olmuştur. Karlarını artırmanın en temel yöntemi müşteri sayısını artırmaktan geçmektedir. İşletmeler bunun için yıllardır farklı kampanyalar ve reklamlar ile markalarının seslerini duyurarak yeni müşteriler çekmek adına yüklü pazarlama bütçeleri ayırmışlardır. Pazarlama alanının gelişmesiyle birlikte aslında eldeki müşteriyi tutundurmanın yeni müşteri elde etmekten daha az maliyetli olduğu görülmeye başlanmıştır. Yeni müşterileri elde etmek için hem pazarlama bütçelerine hem de potansiyel müşteri belirleyebilmek adına danışmanlık firmalarına yüklü miktarlar ayırmak gerekmektedir. Oysa ki eldeki müşteriyi tutundurmak için müşteri ile iletişime geçmek ya da ona özel kampanya hazırlamak gibi daha düşük maliyetli çalışmalar yapılabilmektedir.

Küçük işletmeler az sayıda müşteriye sahip olduğundan her bir müşteri ile temas halinde olarak müşterilerinin problemlerini önceden anlayabilir ve bu müşterileri kaybetmeyi önleyebilirler. Büyük işletmelerde ise çok fazla sayıda müşteri olduğundan ve çoğu müşteri ile birebir temas her zaman mümkün olmadığından bu müşterilerin kaybedilip edilmeyeceğini hareketlerinden anlamak gerekecektir. Bunun için günümüzde veri biliminden faydalanılarak istatistiksel modeller oluşturulmaktadır ve kaybedilebilecek müşteriler önceden tespit edilerek önleme çalışmaları yapılmaktadır.

1.1 Literatür Özeti

Müşteri kaybını tahminlemek için kullanılan farklı farklı yöntemler bulunmaktadır. Bu yöntemler ile ilgili farklı akademik çalışmalar yer almaktadır. Bu çalışmalardan bazıları aşağıdaki gibidir:

Abdelrahim Kased Ahmad ve diğerleri 2019 yılında yaptıkları çalışmada SyriaTel telekomünikasyon şirketinin 9 aylık müşterilerinin verilerinden Decision Tree, Random Forest, Gradient Boosted Machine Tree “gbm” ve Extreme Gradient Boosting “xgboost” algoritmalarını kullanarak müşteri kaybını tahminlemişlerdir. En iyi sonucu %93,3 başarı ile extreme gradient boosting algoritmasında yakalamışlardır (Ahmad vd., 2019).

Gürpınar, 2015 yılında yaptığı çalışmada Türkiye pazarında telekomünikasyon şirketine ait müşterilerin müşteri kaybı analizini yapay sinir ağları ve sosyal ağ analizi kullanarak tamamlamıştır. Gürpınar, çalışmasında müşterilerin sosyal ağlarının da kayıp analizleri için önemli bir etkisi olduğuna değinmiştir (Gürpınar, 2015).

Hung ve diğerleri, churn tahminlemesi yapabilmek için kümeleme, karar ağaçları ve yapay sinir ağlarından faydalanmışlardır. Yapay sinir ağlarının diğer yöntemlere göre daha iyi sonuç verdiği sonucuna ulaşmışlardır. (Hung vd., 2006).

Zhang ve diğerleri, sigortacılık sektöründe müşteri kaybı tahminlemesi yapmak için derin ve sığ modeli önerdiklerini çalışmalarında söylemişlerdir. Sigortacılık sektörüne ait verileri hem derin ve sığ model hem de stokastik gradient descent, Lineer Discriminant Analysis, Gaussian Naive Bayes, Random Forest ve Gradient Tree Boosting gibi farklı modeller ile tahminlemeye çalıştıklarında en iyi sonucu derin ve sığ model ile elde ettiklerini paylaşmışlardır.(Zhang vd., 2017)

Isphreet Kaur ve Jasleen Kaur ise 2020 yılında yaptıkları çalışmada bankacılık sektörüne ait verileri inceleyerek bu sektör üzerinde bir kayıp analizi tahminlemesi yapmayı hedeflemişlerdir. Çalışmalarında lojistik regresyon, decision tree, K-nearest neighbor ve random forest kullanarak modellerin başarılarını karşılaştırmışlardır. (Kaur & Kaur, 2020)

1.2 Tezin Amacı

Günümüzde özellikle büyük şirketlerin müşterilerini elinde tutmak adına çok fazla pazarlama ve analiz çalışması olmaktadır. Gün geçtikçe rekabetin arttığı telekomünikasyon sektöründe ise ülkemizde üç büyük şirketin pazarda söz sahibi olduğunu söylemek doğrudur. Bu şirketler rakip şirketlere müşterilerini kaptırmamak adına çok sert çalışmalar içerisindedir. Bu tez kapsamında kaggle.com üzerinden alınan yabancı bir firmaya ait müşteri verilerinden müşteri churn tahmini yapılması hedeflenmiştir. Çalışma kapsamında elde edilen veri setinden öz nitelik seçimi yardımıyla değişkenler seçilerek öncelikle farklı yapay sinir ağları algoritmaları 4 farklı nöron ve her bir nöron için iki farklı veri bölümlemesi kullanılarak çalıştırılmıştır. Elde edilen sonuçlar tekrardan 4 farklı veri bölümlemesi

yapılarak çalıştırılan lojistik regresyon modelleri ile karşılaştırılmıştır. Bu tez çalışması kapsamında telekomünikasyon şirketlerinin müşteri kaybı analizlerinde faydalanabilecekleri başarılı bir model sonucu ortaya konması hedeflenmiştir.

1.3 Hipotez

Bu tez kapsamında aşağıda belirtilen hipoteze cevap aranmıştır:

1) Bir telekomünikasyon şirketinin verilerinden yola çıkarak yapılan müşteri kaybı tahmininde yapay sinir ağları ya da lojistik regresyon modellerinden hangisi daha başarılı sonuç vermiştir?

2
ARAŞTIRMA

2.1 Veri Madenciliği

Bilgi madenciliği, el ile işlemesi zor olan büyük veri kütlelerinden, bilgisayarlar yardımıyla yarı otomatize bir şekilde faydalı sonuçlar çıkarma sürecidir. Süreç tamamen otomatik değildir. Bu noktada bilgisayar yalnızca analistin karar verme ve sonuca ulaşma süresini kısaltacak şekilde kullanılır. Bilgi keşfinin bir diğer tanımı "Bilgi keşfi veri hakkındaki daha önceden fark edilmemiş üstü kapalı fakat faydalı olma potansiyeli olan bilginin çıkarımıdır" (Frawley vd., 1992) olarak ifade edilebilir. Bilgi keşfi sürecinde veriyi açıklayan, özetleyen ve gereksiz detayları göz ardı eden yapılara "model" adı verilir. Modeller çok büyük boyuttaki ham verileri özetleyerek, fark edilmesi güç ilişkileri gün yüzüne çıkarabilir. Veriyi benzerliklerine göre ayırmaya "sınıflandırma problemi" denir ve bu problemin çözümünde bayesyen sınıflandırma, karar ağaçları, kümeleme analizleri, destek vektör makineleri ve lojistik regresyon gibi modeller kullanılır. Veri madenciliği net bir amaç olmaksızın büyük boyutlu ham veriden anlamlı ve faydalı sonuçlar çıkarmayı, yani bilgi keşfini gerçekleştirmeyi hedefler. Bu bağlamda veri madenciliği çoğunlukla veri tabanların üzerinden bilgi keşif süreci olarak tanımlanır (Hamel, 2009).

2.1.1 Veri Madenciliği Adımları

Her ne kadar bilgi keşfi süreci için geliştirilmiş yöntemler olsa da, sürecin etkili sonuçlar doğurabilmesi için incelenen veri üzerinde kimi manipülasyonlar yapılarak verinin sürece uyarlanması gerekir. Bu noktada, verinin incelenmesi ve bilginin keşfedilmesi süreci şu şekilde adımlandırılabilir:

- Problemin tanımlanması
- Verinin hazırlanması
    - Veri temizleme
    - Veriyi standardize etme
    - Veri indirgeme
    - Veri dönüşümü ve entegrasyonu
- Modelleme

    - Kullanılacak tekniğin seçimi
    - Test ortamının oluşturulması
    - Modelin kurulması
    - Sonuçların değerlendirilmesi
- Modelin kullanılması

Problemin Tanımlanması: Veri madenciliği sürecinin tabiatı gereği net çizgilerle belirlenmiş bir problemi yoktur. Problem özelleştirilmeden inceleme adımına geçilebilir fakat incelenecek verinin yapısı ve kullanılacak modelin belirlenmesi gereklidir. Bu aşamada incelemeyi gerçekleştirecek analist ilgili veride odaklanması gereken noktaları ve değerli olabilecek kısımları tespit etmek için o alandaki bir uzmanla işbirliğine gitmelidir. Böylelikle varılmak istenen sonuçlar ve kullanılacak yöntemler belirlenebilir.

Verilerin Hazırlanması: Veri madenciliği büyük çaplı, ham veri yığınları üzerinde çalıştığı için kullanılan veriler araştırma yapmaya elverişli olmayan pek çok veriyi özünde barındırabilir. Bu yüzden, yapılan çalışmadan elde edilecek sonuçların kalitesini arttırmak adına bu veriler kimi işlemlere tabi tutulabilir. Bu süreç verinin tamamını kapsadığı için verinin boyutuna da bağlı olarak çok fazla zaman alabilir (Sund, 2002).

Veri madenciliği uygulamalarını gerçekleştirmek adına veri ambarları ile farklı hedeflerden gelen düzensiz veriler, düzenlenip bir araya getirilebileceği gibi veri tabanlarını veri madenciliğine uygun hale getirilerek de gerçekleştirilebilir (Chapman vd., 2000).

Veri Temizleme: Bu aşama, incelenen verideki sapan veya eksik verilerin sistemden ayıklanmasını içerir. Verideki eksik veriler belli bir düzeyde kalıyorsa bu noktalar tahmin edilerek doldurulabilir. Ancak çok fazla eksik veri varsa tutarlı bir tahmin yapılamayacağından bu değerler göz ardı edilir. Aşırı uçta bulunan ve sistemde sapmalara neden olan verilerin de sistemin dışında tutulması gerekmektedir (Adriaans, Zantinge, 1997).

Veriyi Standardize Etme: Verideki nümerik değişkenler farklı değişken aralıklarına sahip olmaya eğilimlidir. Her değişkenin etki aralığını standart bir forma sokmak için Min – Maks Normalizasyonu, Z – Skor Standardizasyonu gibi yöntemler sıklıkla kullanılır.

Veri İndirgeme: Veri indirgeme, verinin bütünü ile aynı veya kısmen aynı analitik sonuçları üreten bir alt küme elde etme işlemidir. Boyut indirgeme ve sayısal indirgeme bu amaçla sıkça kullanılan metotlardır.

Veri Dönüştürme: Veri entegrasyonu farklı kaynaklardan elde edilen farklı isimli fakat aynı değişkeni ifade eden değişkenlerin ortadan kaldırılmasını tanımlar.

Modelleme: Modelleme; kullanılacak modellerin seçilmesini ve modellerin optimum sonuçlara ulaşması için parametrelerin ayarlanmasını kapsamaktadır.

Modelleme Tekniğinin Seçimi: Bu aşamada kullanılacak tekniğin varsayımları dikkate alınarak seçim yapılmalıdır. Verinin incelendiği bölümde kullanılacak modele karar verilmiş olmasına rağmen kullanılacak teknik bu aşamada karar verilir. Örnek ile açıklamak gerekirse, karar ağaçlarının kullanılacağına karar verilmişse CHAID algoritmasını J48 algoritmasına karşı tercih edilmesi aşamasıdır. Bu tercih işlemi efektif tekniğin kullanılması için önemlidir.

Test Tasarımının Oluşturulması: Bu aşamada veri öğrenme ve test olmak üzere iki gruba ayrılır. Model test grubu üzerinden kurup optimize edilerek öğrenme grubu üzerinden başarısı ölçülür.

Modelin Kurulması: Modelin kurulması aşamasında model farklı ana veriden türetilen farklı öğrenme ve test verileri üzerinde en optimal sonucu üreten modele ve parametrelere karar verilene kadar kendinin tekrarlayan bir süreçtir.

Modelin Değerlendirilmesi: Elde edilen sonuçların araştırılmak istenen konu bazındaki yeterliliğinin ölçülmesi aşamasıdır. Modelin türettiği sonuçların faydalı bulunmaması durumunda farklı bir model denenebilir veya araştırma en başından alternatif modelleri de kapsayacak şekilde kurgulanarak en uygun modelin belirlenmesi bu aşamada gerçekleştirilebilir.

Modelin Kullanılması: Önceki adımlardan geçen modelin kullanıldığı aşamadır.

2.2 Müşteri Geri Kazanımı

Müşteri geri kazanımının tahmini ve yönetimi, hücresel pazarın serbestleştirilmesi nedeniyle daha önemli bir görev haline gelmiştir. Şirketten ayrılmayı amaçlayan sadık müşterilerin zamanında tahmini, kimliklerini korumaya yardımcı olabilir ve onları korumak için proaktif eyleme tabi olabilir. Güçlü müşteri tabanını korumak

için bu analizle bir model kurmak erken bir uyarı sistemi oluşturmaktadır. Yapay Sinir Ağını diğer sınıflandırma tekniklerinden farklı olarak müşteri geri kazanımı analizinde faydalı olduğu literatürde sıkça yer almıştır. Halihazırda mevcut olan geri yayılma (beslenme) algoritması tekniği çoğu sektördeki müşteri geri kazanımının tahmini için modeli eğitmek için kullanılır.

Çalışmanın bu kısmında müşteri geri kazanımının tanım ve kapsamı verilerek neden önemli olduğu literatürden örneklerle ortaya konacaktır.

2.2.1 Müşteri Geri Kazanımı: Tanım ve Kapsamı

Veri madenciliğinin bir parçası haline gelen Müşteri Geri Kazanımı (Churn Analysis) herhangi bir şirketin müşteri tabanındaki yıpranma oranının hesaplanmasıdır. Bir hizmet veya ürün kullanmayı bırakma olasılığı en yüksek olan tüketicilerin tanımlanmasını içerir. Bu analiz, bir şirkette müşteriyi elde tutmak için sürdürülebilir ve sağlam bir strateji geliştirmede son derece yararlıdır.

Bir şirket, belirli bir zaman diliminde onlarla ilişkilerini sona erdiren müşterilerin yüzdesinin farkında olduğunda, karmaşa oranının nedenlerinin ayrıntılı bir analizini kolayca oluşturabilir. Uzun mesafeli telefon hizmeti veya dergiler gibi abonelik hizmetleri. Churn analizi, aboneliklerini iptal eden ve işlerini bir rakibe taşıyan müşterilerin davranışlarının anlaşılmasına ve bu olayın gerçekleşme olasılığının öngörülmesine yardımcı olur (Forhad vd., 2014).

Müşteri Geri Kazanımı belirli bir süre içinde dışarıya giren müşteri sayısının ölçümüdür. Müşteriyi kaybetmenin nedeni biliniyorsa, sağlayıcılar müşterilerin ihtiyaçlarını karşılamak için hizmetlerini geliştirebilirler. Potansiyel müşterinin de geçmişini sistematik olarak analiz ederek doğacak kayıplar azaltılıp müşteri tekrar kazanılabilir (Dahiya & Bhatia, 2015). Belirsizliğin yüksek olduğu sektörler rekabetçi bir ortam nedeniyle hızla değişmeye devam eden müşterilerinin her biri için büyük miktarda bilgi sağlanmaktadır. Bu bilgiler her sektörün sürecine yönelik ayrıntıları içerir. Örneğin; bir telekomünikasyon firmasında bu veriler faturalandırma, ağ verileri, aramalar gibi detayları içermektedir. Mevcut bilgiler, operatörlerin kayıp oranlarını tahmin etmeleri ve azaltmaları için çeşitli yollar sağlamak amacıyla farklı perspektiflerde analiz edilebilir. Verilerdeki ilginç kalıpları keşfetmek için kullanılan yöntemler ve müşterinin veri tabanında depolanan verilere dayanarak titreme olup olmayacağını tahmin etmeyi öğrenmeye yardımcı

olan bu analiz kapsamıyla her sektöre her yapay sinir ağlarına uyarlanabilecek niteliktedir.

Müşteri geri kazanımının tahmini ve analizi farklı tekniklerle yapılır ve çoğunlukla veri madenciliği araçlarıyla kaplanır. Araştırmacılar tarafından farklı veri setlerinde değişen doğruluk ve hassasiyetle karma tahmin modelleri oluşturmak için birçok farklı çalışma yürütülmektedir. Tayvan'daki bir lojistik şirketinin veri kümesinden model oluşturmak için Destek Vektör Makinesi (Support Vector Machine - SVM), Lojistik Regresyon (Logistic Regression - LR) ve Çok Katmanlı Algılayıcı (Multilayer Perceptron) kullanılmaktadır (Chen vd., 2015). Benzer şekilde, karar ağacı algoritması Malezya Telekomünikasyon Şirketi için tahmin modelinin oluşturulmasında kullanılmaktadır (Arsan & Çimenli, 2016). Karar Ağacı temelinde yapılan bir başka çalışma karar ağacı ve lojistik regresyon ile sağkalım analizinin yapılması üzerinedir (Lu, 2002). Araştırma, Oracle'ın veri setinden elde edilen veriler kullanılarak yürütülmekte ve müşteri geri kazanımı tahmin modelinin oluşturulması için Naïve Bayesian Teoremi teknikleri kullanılmaktadır (Nath & Behara, 2003). Çin Telekomünikasyon Şirketi'nde veriler Karar Ağacı, C5.0 algoritması geri yayılımlı yapay sinir ağı (BPN), ve yapay sinir ağı kullanılarak bu analiz gerçekleştirilmiştir (Liu vd., 2006)

2.2.2 Müşteri Geri Kazanımı: Önem

Müşteri Geri Kazanı Oranı, müşterilerin bir ürün veya hizmet kullanarak durma olasılığı olarak tanımlanabilir. Müşteri kaybı, telekom, bankacılık, sigorta gibi müşteri sürekliliği sektörlerinde kritik bir konudur, çünkü mevcut müşterilerin çoğunu tutmak genellikle yeni müşteriler kazanmaktan daha düşük maliyetli işlemler gerektirir (Reichheld & Sasser, 1990). Bu tahminlerle müşteri kaybını önlemek için Müşteri İlişkileri Yönetimi çözümleri üretilebilir. Bu nedenle şirketler, ürünü veya hizmeti kullanmayı bırakma ve bu müşterileri koruma şansı yüksek müşterileri belirlemek için özel kampanyalar düzenleyebilir. Ayrıca, bu işletmelerin değeri aktif müşteri sayısıyla doğru orantılıdır. Bu nedenle, işletmelerin maliyetleri, karlılığı, büyüklüğü, yatırım kapasitesi ve nakit akışı gibi parametrelerin sayısı, müşteri sayısına ve müşteri sadakatine bağlıdır. Buna ek olarak, araştırmalar uzun vadeli müşterilerin kârlılığının daha yüksek olduğunu göstermiştir. Bu karlılığı

anlamak için müşteri yaşam boyu değeri gibi yöntemler hesaplanır (Van den Poel & Lariviere, 2004).

Müşteri Geri Kazanımı Yönetimi, yeni bir müşteri elde etmek mevcut müşterileri elde tutmaktan daha pahalı olduğu için bu kaybı azaltmak için çok önemlidir (Jadhav & Pawar, 2011). Tüketicilerin mevcut ve yeni ürünlere tepkisi, ürün ve hizmet kalitesindeki iyileşmeler ve rekabetçi davranışlar bu analizin doğru bir teknikle uygulanması ile öngörülebilir. Örneğin; geçmişteki eğilimleri analiz ederek, banka analistleri mevcut talebi belirleyebilir, tüketicilerin ürün ve hizmet bazında gelecekteki davranışlarını tahmin edebilir ve daha fazla fırsat yaratabilir. Veri madenciliği tekniği aynı zamanda karlı müşterilerin ve işletmelerin kârsız olanlardan ayırt edilmesini sağlar. "Hangi müşteriler rakiplere gitmeye daha meyilli?" gibi bir sorusunun cevabı öğretilmiş bir sistemin müşteri kazanımı analizi yapması ile verilebilir. Bu analiz şirketlerin müşterilerinin neden rakiplere yöneldiğini anlamalarını sağlar. Bu şekilde şirket, müşteri ilişkilerini yeniden düzenler ve müşteri sadakatini arttırır (Dahiya& Bhatia, 2015).

2.3 Yapay Sinir Ağları

İnsana ait özelliklerden biri olan "düşünebilme" ve "öğrenebilme" yetenekleri yapay zekâ ile araştırma konusu haline gelmiştir. Bilgisayar kullanımını artmasıyla bu alandaki çalışmalar her geçen gün boyut değiştirmiştir. Bu çalışma kapsamında ortaya çıkan ve aynı zamanda yapay zekanın destekleyicisi bir alan olan Yapay Sinir Ağları teknolojisi öğrenebilme sistemlerinin temelini de oluşturmaktadır. Bu kavram temelde beynin bir işlevi yerine getirmek yöntemini modellemek üzere tasarlanan bir sistem olarak tanımlansa da verinin içerdiği yapının tanımlanmasını gerektiren her alanda kullanılmaktadır. Bu çalışmanın da önemli bir parçası olan yapay sinir ağlarını anlayabilmek adına çalışmanın bu kısmında bu kavram literatürdeki çeşitli tanım ve formülleri yaşadığı gelişimle birlikte açıklanacaktır.

2.3.1 Yapay Sinir Ağlarının Tarihçesi

Yaklaşık 4000 yıl önce başlayan veri depolama merakı günümüzde yapay zekâ ve yapay sinir ağlarının gelişmesine katkı sağlamıştır. 1956 yılına kadar klasik filozofların düşünce sistematiğini tanımlama girişimi olsa da "yapay zekâ" ilk defa Dartmouth College (Hannover)'da yapılan bir konferansta kullanılmıştır (Lewis,

2014). Bununla beraber İngiliz matematikçi tarafından yapılan Turing Makinesi makinelerin de insan beyni gibi düşünüp bazı işlemleri yapabildiğini göstermişti (Crone, 2004).

Turing Makinesinde kullanılan ağ üzerinden ilk yapay sinir ağı modeli ise bir sinir doktoru Warren McCulloch ve matematikçi Walter Pitts ile 1943 yılında ortaya çıkmış ve "Sinir Aktivitesinde Düşüncelere Ait Bir Mantıksal Hesap (A Logical Calculus of Ideas Immanent in Nervous Activity)" başlıklı makale ile de bu model literatüre geçmiştir. Çalışmalarında fonksiyonları "ve" ya da "veya" gibi mantıksal operatörlerle modelleyerek nöronun mantığı ile yapay sinir hücrelerinin öğrenme yeteneği kazanabileceğini öne sürmüşlerdir (McCulloch& Pitts, 1943).

Bu ilk başlangıçtan sonra çeşitli bilim adamları Yapay Sinir Ağları ile ilgili çalışmalara devam etmiştir. Örneğin Hebb (1949), Newell ve Simon (1955), Rosenblatt (1959) bu örneklerden sadece birkaçıdır. Özellikle Newell ve Simon tarafından öne sürülen "Fiziksel Simge Varsayımı" insandan bağımsız bir zekâ sisteminde başlangıç kabul edilmektedir (Saraç , 2004).

Ancak Minsky ve Papert (1969) tarafından yayımlanan "Perceptron" kitabında tek katmanlı yapay sinir ağlarının çok basit problemleri çözemeyeceği gösterilmiş ve bundan sonra yapay sinir ağları ile ilgili çalışmalar yavaşlamıştır (Fletcher D. & Gross D., 1993). Minsky ve Papert'in çalışmasında XOR probleminin çözülememesi örnek gösterilerek yapay sinir ağlarının doğrusal olmayan problemlerde işlevselliği olmadığı öne sürülmüştür (Öztemel, 2006).

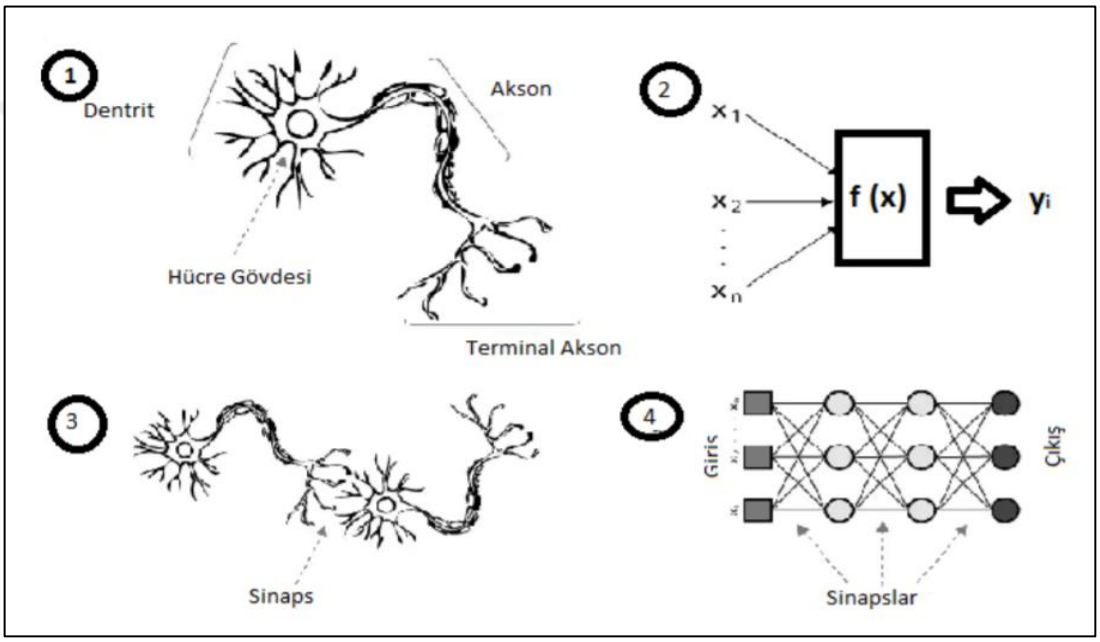
Yapay sinir ağları ile ilgili yapılan gelişmelerin günümüzdeki halini bulması ise Parker ve Werbos (1986) ile Rumelhart ve McLeland (1986)'in yapığı çalışmalardan sonra olmuştur. Rummelhart çok katmanlı algılayıcılar ile daha önce yapay sinir ağlarının çözemediği öne sürülen XOR problemini çözmüşlerdir. Çok katmanlı ağa alternatif olarak ise Broomhead ve Love (1998) tarafından Radyal Tabanlı Fonksiyonlar (RBF), Specht ile de Probabilistik Ağlar (PNN) ve Regresyon Ağları (RNN) ortaya konmuştur (Öztemel, 2006).

Özellikle Hopfield çalışmasında Geri Yayılım Algoritması yapay sinir ağlarının tekrar keşfedilmesini sağlamıştır. Hopfield "Self Organizing Maps (SOM)" u geliştirmiş ve yapay sinir ağlarının optimizasyon gibi teknik problemlere bile çözüm getirebilen yanının olduğunu ortaya koymuştur (Hopfield, 1982).

1980'li yıllarında atılan bu adımlarla yapay sinir ağlarındaki gelişmeler ivme kazanmış ve günümüzde gündelik hayatın içine bile giren bir kavram haline dönüşmüştür.

2.3.2 Yapay Sinir Ağlarının Tanımı ve Özellikleri

Yapay Sinir Ağları ilk modellendiğinde biyolojik sinir ağlarının sentetik bir kopyası (Egrioglu, 2009) iken aslında insan beyninden esinlenilmiş öğrenme sürecini matematiksel olarak modelleme uğraşıdır (Kabalcı, 2014). Şekil 1 üzerindeki hali ile biyolojik bir sinir hücresi ve yapay sinir ağının benzeşen yanları gösterilmektedir.



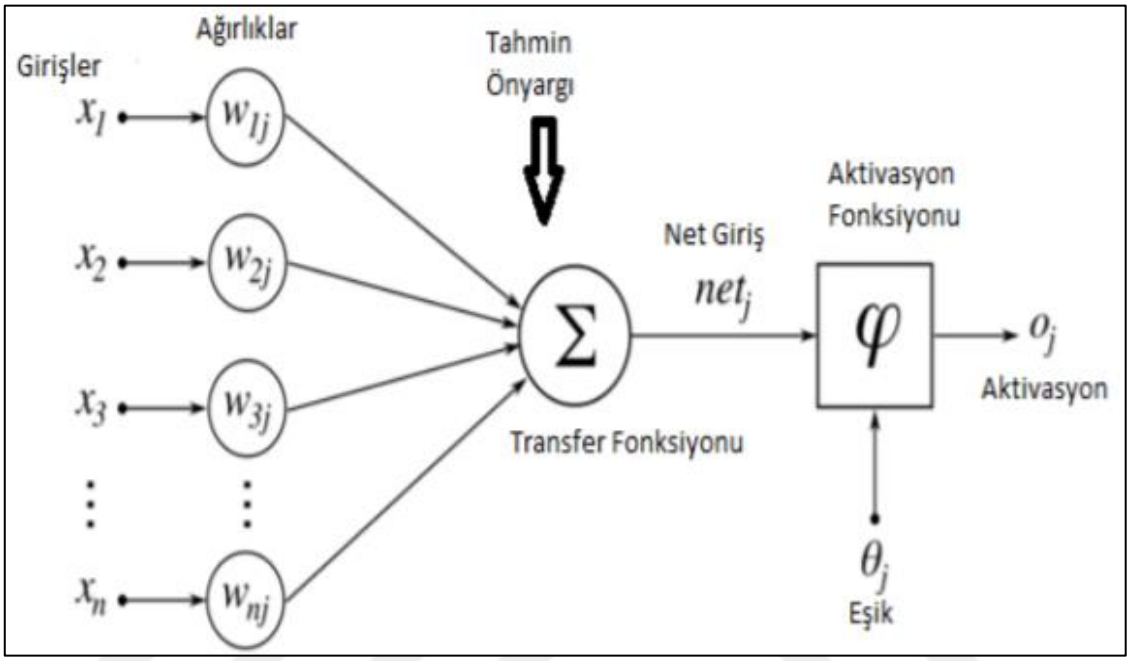
Şekil 2.1 Biyolojik Sinir Hücresi ve Yapay Sinir Ağı (Maltarollo vd., 2013)

Şekil 1 üzerinde görselleştirilmiş modellere göre bu iki sistemin kavramlarının birbirinden neye denk geldikleri ise Şekil 2 üzerinde ifade edilmiştir. Biyolojik sinir siteminde kısımlara ayrılan her parça yapay sinir ağında başka bir kavramla karşılık bulmuştur.

Tablo 2.1 Biyolojik Sinir Sistemi Kısımlarının Yapay Sinir Sistemindeki Karşılığı (Öztemel, 2006)

Biyolojik Sinir Sistemi	Yapay Sinir Sistemi
Nöron	İşlemci Elemanı
Dentrit	Toplama Fonksiyonu
Hücre Gövdesi	Transfer Fonksiyonu
Aksonlar	Yapay Nöron Çıkışı
Sinapslar	Ağırlıklar

Bu sistemin model hali ile en genel tabiri ile aşağıdaki gibi kurgulanır. Şekil 3'deki modele göre bir hücreye n tane veri girişi yapılmakta ve girilen veriler ağırlıklandırılmaktadır. Verilen ağırlıkları ile çarpıldıktan sonra toplanarak bir önyargı eklenmekte ve sonucunda ise net yargıya ulaşılmaktadır. Bu süreçten bir veri elde edilmek için ise net girdi aktivasyon fonksiyonundan geçmektedir.



Şekil 2.2 Yapay Sinir Hücresi (Kabalcı, 2014)

Günümüzde birçok alana dokunan bu kavram literatürde farklı adlarla da tanımlanmaya başlamıştır. Paralel dağıtılmış ağlar, bağlantılı ağlar ve nuromorfik ağlar bu tanımlamalardan birkaçıdır (Keskenler & Keskenler, 2017).

Literatürde farklı alanlara olan destekleyiciliği ile bilinen yapay sinir ağları aşağıdaki temel özelliklerle ilişkilendirilmektedir (Öztemel, 2006).

- **Doğrusal Olmama:** Yapay Sinir Ağları doğrusal olmama özelliği ile kompleks yapıdaki problemlere çözüm getirme becerisine sahiptir.
- **Öğrenme:** Geleneksel programlama şekillerinden farkını bu özellikle kazanan yapay sinir ağları öğretilen bilgiler ile benzer olayları açıklayabilmektedir.
- **Hata Toleransı ve Esneklik:** Yapay sinir ağlarının hatalara karşı duyarlı olması dereceli bozulma yaşamasını sağlamaktadır. Ayrıca çok fazla işlem elemanından ve ayarlanabilir ağırlıklardan oluşması sebebiyle esnek olan yapay sinir ağları modelin işlevini yitirmesi de bu esnek yapı ağsa kısmi zarar yaratabilmektedir.

- **Genelleme:** Yapay sinir ağları diğer modellerin aksine sınıflandırıcı ve gidişat tanımlayıcısı özelliğine sahiptir. Başka bir ifade ile eğitilen ağ sayesinde yeni bir veri kümesi önceki deneyimlerle kolayca öğrenilmektedir.
- **Eksik Verilerle Çalışma:** Geleneksel yöntemlerin aksine yapay sinir ağları eğitildiği zaman eksik veri ile bile çıktın yaratabilmektedir. Ancak bu durum performans kaybı yaratmaması için eksik bilginin önem derecesi de yapay sinir ağlarına öğretilebilmektedir.
- **Sınırsız Sayıda Değişken ve Parametre Kullanma:** Yapay sinir ağlarının gerçeğe yakın tahminler yapıp genel çözüm üretebilme becerisi sınırsız sayıda değişken ile çalışabilme özelliğine dayanmaktadır.
- **Hafıza:** Öğretilme sayesinde yapay sinir ağlarının bu öğretileri saklama özelliği de vardır. Ancak diğer programlardaki gibi veri tabanında değil ağın üzerinde saklamaktadır.
- **Gerçekleme Kolaylığı:** Ağ üzerinde bağlantının düzgün bir şekilde kurulabilmesi için gerçekleme yönünden kolaylıklar yaratır.
- **Gerçek Zamanlı İşlem:** Yapay sinir ağlarında işlemlerin birbirine paralel bir şekilde yapılması gerçek zamanlı işlem yapılabilmesi anlamına gelir.

Uyarlanabilirliliği yüksek olan Yapay Sinir Ağlarının en çok tahmin, sınıflandırma, veriyi ilişkilendirme, yorumlama ve filtrelemede kullandığı görülmektedir. Bu sebeple Yapay Sinir Ağlarını tanımlarken ne amaçla kullanıldığını bilmek modeli anlayabilmek adına faydalıdır (Ağyar, 2015).

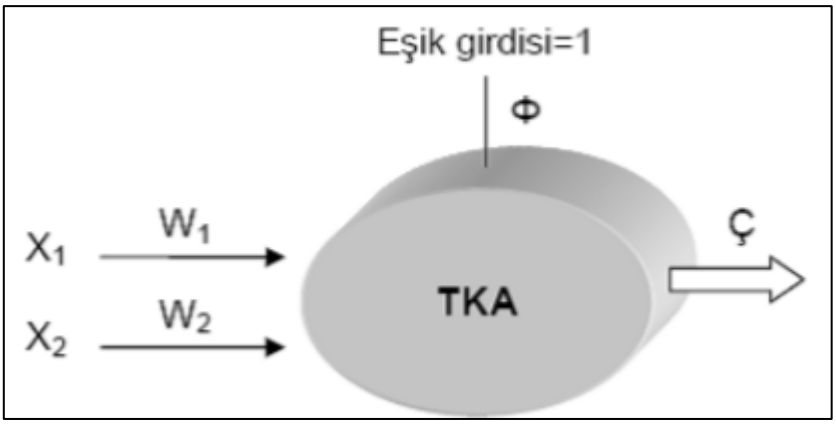
- **Tahmin:** Yapay Sinir Ağları girdi değeri ile çıktı değerini tahmin etmek için kullanılabilir. Örneğin altın ons fiyatının tahmin edilmesi gibi
- **Veri Filtreleme:** Kodlanan yapay ağların tüm veriler arasındaki işe yarayabilmesini ölçmek adına yapay sinir ağları veri filtreleme özelliği ile kullanılır.
- **Sınıflandırma:** Sistemin daha hızlı çalışması adına yapay sinir ağlarının kullanımı girdi verilerini sınıflandırarak bu amacı gerçekleştirir.
- **Veri Yorumlama:** Ağ girdilerinin eğitilip analiz edilmesiyle yapay sinir ağları yeni yorumlamalar yapılmasına olanak sağlar.

- **Veri İlişkilendirme:** Yapay sinir ağları öğrenme üzerine kurulu olduğundan öğrendiği bilgileri ilişkilendirerek eksik bilgileri tamamlama açısından faydalıdır.

Bu özellikler ve kullanım alanı çerçevesinde tanımlanan yapay sinir ağlarının "öğrenme algoritması" ve "aktivasyon fonksiyonu" olmak üzere iki bileşenden oluştuğu söylenebilir. Yapay sinir ağlarının en önemli özelliği olan "öğrenme algoritması" verinin kaynağından öğrenildiği anlamına gelir. Ağ üzerindeki veri ağırlıklarda muhafaza edilir. Bu ağırlık verileri en uygun veriyi bulmaya odaklıdır ki, bu da ağın eğitilebilir olduğunu gösterir. Yapay sinir ağlarının ikinci bileşeni olan aktivasyon fonksiyonunda ise yapay sinir hücresi girdideki veriler sayesinde işleme girerek bu veriyi karşılayan net çıktı sonuçlarının bulunmasını sağlamaktadır. Genellikle doğrusallık göstermeyen bu fonksiyonun sonuç performansını etkilediğini düşünürsek doğru seçilmesi oldukça önemlidir. Burada fonksiyon tek ve çift olmak üzere iki kutup halinde de olabilir. Genel anlamıyla aktivasyon fonksiyonları parçalı doğrusal fonksiyon, hiperbolik tanjant aktivasyon fonksiyon ve adımsal aktivasyon fonksiyonu gibi isimlerle karşılık bulmaktadır.

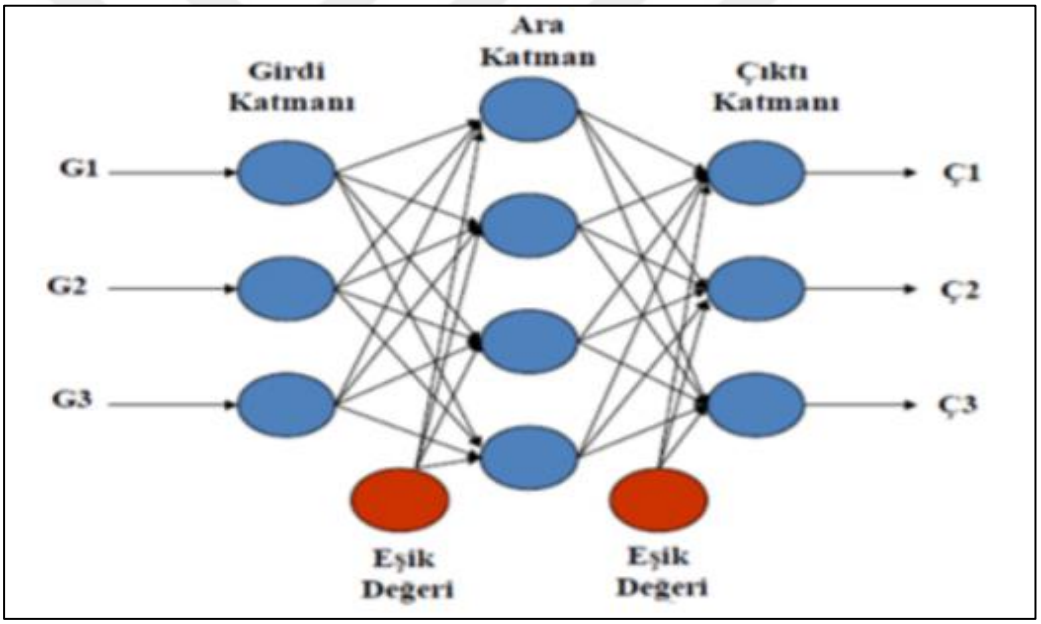
İşlevleri, bileşenleri ve özelliklerinin yanı sıra yapay sinir ağlarının modelleri de tanımlarken kapsamını belirlemede oldukça önemlidir. Genel anlamıyla tek katmanlı algılayıcılar, çok katmanlı algılayıcılar, ileri beslemeli yapay sinir ağları ve geri beslemeli yapay sinir ağları olmak üzere dört grupta incelenen yapay sinir ağları her bir model için görselleştirilmiş bir şemaya da sahiptir.

Şekil 4 üzerinde görselleştirilmiş modelde sadece girdi ve çıktıdan oluşan yapay sinir ağları "tek katmanlı ağlar" (TKA) olarak ifade edilmektedir. Bu algılayıcıda doğrusal özellikteki çıktı fonksiyonu +1 veya -1 değerini almaktadır. Çıktının +1 olması birinci sınıfta, -1 de olması ise ikinci sınıfta olduğu anlamına gelmektedir (Öztemel, 2006).



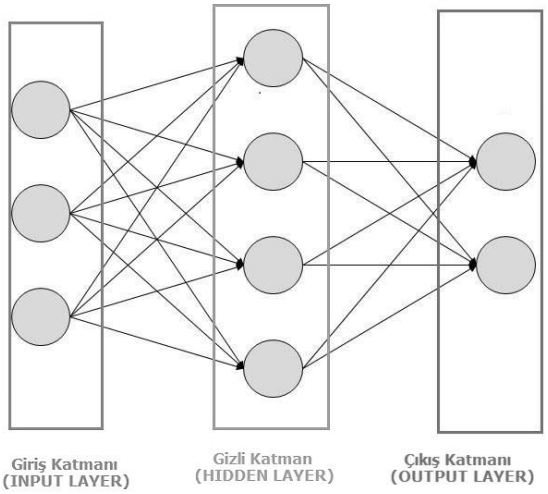
Şekil 2.3 Tek Katmanlı Algılayıcı Modeli (Öztemel, 2006)

İkinci olarak; Şekil 5 üzerinde de görselleştirilmiş hali ile doğrusal olmayan yapay sinir ağları da vardır. Birçok nöronun belirli bir üstünlükle bağlandığı bu yapı "Çok Katmanlı Algılayıcılar" olarak ifade edilir (Kabalcı, 2014).



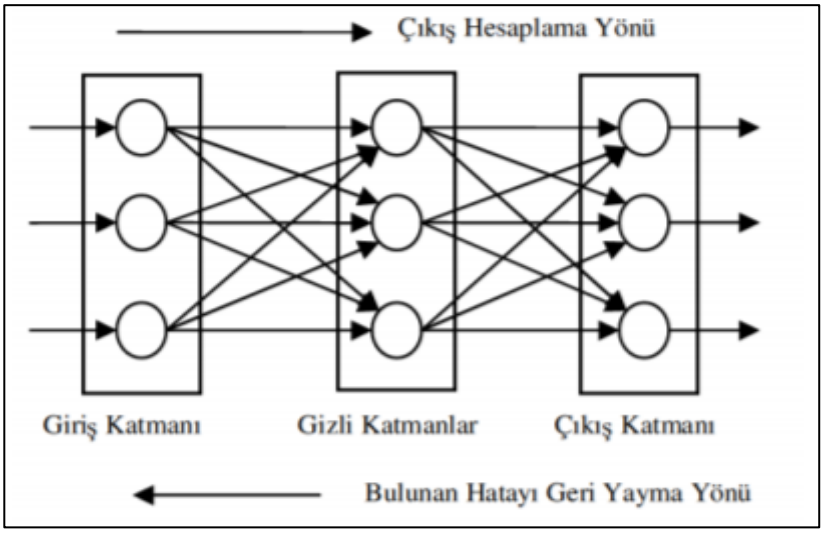
Şekil 2.4 Çok Katmanlı Algılayıcı Modeli (Kabalcı, 2014)

Üçüncü grup ise İleri Beslemeli Yapay Sinir Ağları'dır. Bu grupta nöronların yapısı girişten çıkışa kadar düzenli katmanlar halindedir. Şekil 6 üzerinde de görüldüğü gibi bir katmanın sadece kendinden sonraki katmanla bağı vardır. Ayrıca bu ağın girişine gelen bilgilerde bir değişim yaşanmadan gizli katmandaki hücreye iletilmesi söz konusudur. Buradan sonra ise veri sırasıyla çıkış katmanından işlenir ve dış ortama geçer (Öztemel, 2006).



Şekil 2.5 İleri Beslemeli Ağ Yapısı (Kabalcı, 2014)

Son olarak Geri Beslemeli Yapay Sinir Ağlarında ise bir önceki sistemden farklı olarak nöronun çıktısı kendinden önceki katmanda veya kendi katmanındaki farklı bir nörona girdi olarak verilebilir. Bu yönüyle ileri beslemeli bir yapay sinir ağının doğrusal olmayan bir dinamikteki hali gibidir. İleri beslemeden farklı olarak geri besleme özelliğinin olması bağlantıların davranış şekline göredir. Literatürde Elman Ağı olarak da geçen bu sitemi Elman (1990) zaman serisi öngörü probleminde kullanmıştı (Erilli, 2010).



Şekil 2.6 Geri Beslemeli Ağ Yapısı (Kabalcı, 2014)

2.3.3. Yapay Sinir Ağlarının Formülizasyonu

Şekil 3 üzerinde yapay sinir hücresi olarak gösterilen modelde yapay sinir ağlarının temel bileşenleri gösterilmekte ve bu ağ üzerinden yapılan fonksiyonların temelini

oluşturmaktadır. Yapay sinir ağlarının girdiler, ağırlıklar, toplam fonksiyonu ve aktivasyon fonksiyonu ve hücre çıktısı olmak üzere beş bileşeni vardır.

Girdiler bir ağa kendisinden ya da dış ortamdan gelen bilgiye karşılık gelirken; ağırlıklar bu girdilerin çıktı üzerindeki etkisini ortaya koymaktadır (Kubat C, 2015). Ayrıca bir ağırlığın değeri sıfır ise girdinin çıktı üzerinde bir etkisi olmadığını ifade etmekte ve değerin sıfırdan farklı olduğu durumda da artı veya eksi değerde olması sırasıyla pozitif ve negatif etki yaratacağı anlamına gelmektedir (Öztemel, 2006)

Bu ağın toplama fonksiyonu (birleştirme fonksiyonu) bilgilerin ağırlıklar ile çarpımından elde edilen girdi toplamıdır. Net girdi toplamı fonksiyonda girdi değeri "x_i", ağırlık değeri "w_i" ve eşik değeri "b" ile gösterilmektedir. Ayıca "n" toplam girdiyi ifade etmektedir. Bu bağlamda bir hücredeki net girdi toplamının formülü (2.1) aşağıdaki gibidir (Kaynar, 2010).

$$NET = \Sigma_{i=1}^{n} x_i w_i + b \tag{2.1}$$

Bu net giriş fonksiyonu ile ise yapılan çeşitli hesaplamalar aşağıdaki gibidir. Buna göre (2.2) üzerinde ağırlık değeri ve girdilerin çarpılıp; bulunan değerlerin birbiri ile çarpımından net girdi hesaplanmaktadır. Ayrıca (2.3) üzerinde ağırlıkların çarpıldıktan sonra N tane girdi içinden en büyük olanı net girdisi olarak ifade edilmektedir. Benzer şekilde en küçük olanın seçilmesiyle de (2.4) üzerinde ifade edilen net girdiye ulaşılmaktadır (Öztemel, 2006).

$$Net\ Girdi = \prod_{i} x_i w_i \tag{2.2}$$

$$Net\ Girdi = \max(x_i \omega_i)\,, i = 1 \dots \dots N \tag{2.3}$$

$$Net\ Girdi = \min(x_i \omega_i)\,, i = 1 \dots \dots N \tag{2.4}$$

Ağırlıklar ile çarpıldıktan N tane girdiğin içerisindeki pozitif ve negatif değerlerin bulunup büyük olan sayı hücresinin net girdi olarak kabul edilmesi ise (2.5) üzerinde formülize edilmiştir. Ayrıca (2.6)'daki formülde gelen bilgilerin ağırlıklandırılarak toplanıp öncesinde gelen bilgiler ile net girdi toplanması görülmektedir (Öztemel, 2006).

$$Net\ Girdi = \Sigma_i \operatorname{sgn}(x_i w_i) \tag{2.5}$$

$$Net\ Girdi = Net(eski) + \Sigma_i (x_i w_i) \tag{2.6}$$

Yapay sinir ağlarının bir diğer bileşeni olan aktivasyon fonksiyonu isenet girdini işlenip ağın gelen girdilere karşı ne kadar çıktı üreteceğini belirlemektedir. Aktivasyon formülü seçilirken önemli olan türevinin kolay hesaplanabilmesidir. Örneğin geri beslemeli bir ağda aktivasyon fonksiyonun türevi kullanıldığından seçim sırasında kolay hesaplanabilir türev prensibi zaman açısından fayda sağlar (Kubat, 2015). Yukarıdaki toplama fonksiyonuna benzer şekilde aktivasyon (transfer) fonksiyonu için de aşağıdaki formüller ifade edilmektedir. Buna göre aktivasyon fonksiyon örnekleri şu şekilde ifade edilir (Öztemel, 2006).

- **Lineer Fonksiyon:** Gelen girdi hücrenin çıktısı olarak kabul edilmektedir.

$$f(NET) = NET \tag{2.7}$$

- **Step (İşaret) Fonksiyonu:** Net girdi değerini belirleyen bir eşik değeri "b" üzerinden kurulmaktadır. Hücre çıktısı eşik değerinin altında olursa 0, üzerinde olursa 1 değerini almaktadır.

$$f(NET) = \begin{cases} 1, eğer\ NET > b \\ 0, \text{eğer N}ET \leq b \end{cases} \tag{2.8}$$

- **Sinüs (Sigmoid) Fonksiyonu:** Bu fonksiyon uygun dağılımı olan ve öğrenilmesi düşünülen durumlar için kullanılmaktadır.

$$f(NET)=\text{Sin(NET)} \tag{2.9}$$

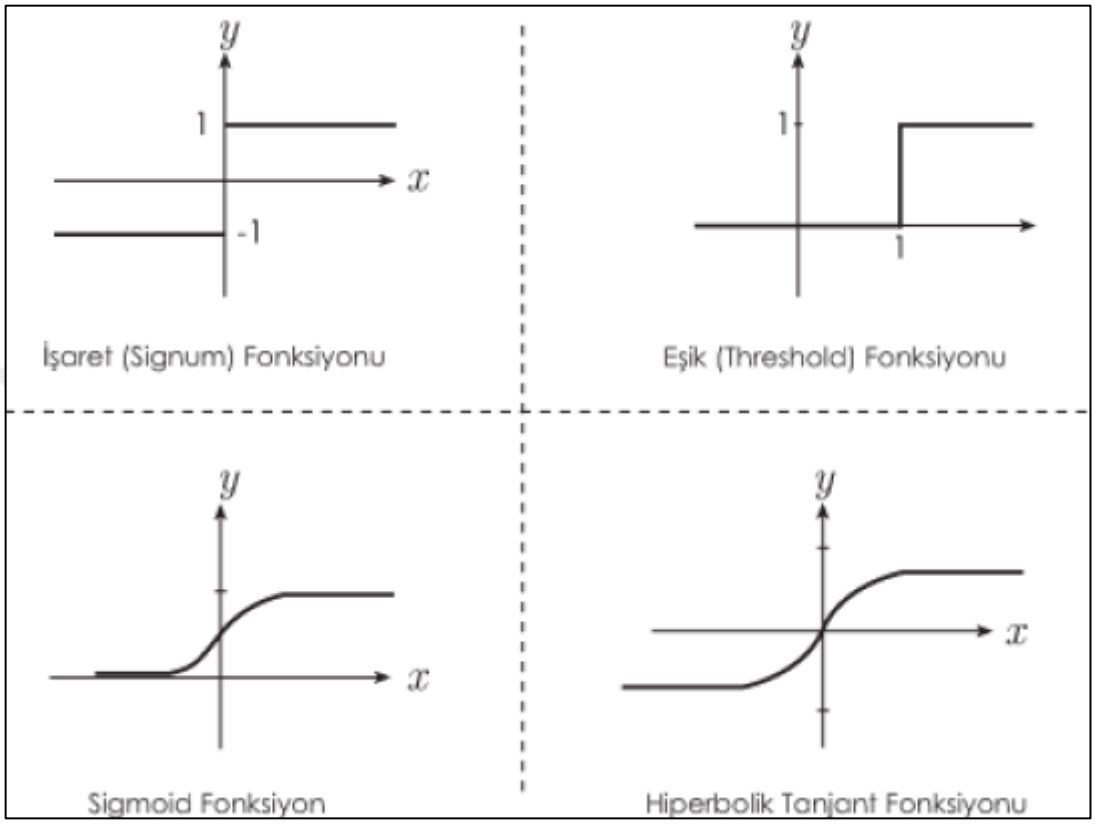
- **Eşik Değer (Treshold) Fonksiyonu:** Gelen bilginin 0 olması ya da 1'den büyük ve küçük olmasına göre çıktının değer aldığı fonksiyonudur. Ayrıca bu değer sadece o veya 1 değerini alabilir.

$$f(NET) = \begin{cases} 0, eğer & NET \leq 0 \\ NET, eğer & 0 < NET < 1 \\ 1, \text{eğer} & \text{N}ET \geq 1 \end{cases} \tag{2.10}$$

- **Hiperbolik Tanjant Fonksiyonu:** NET girdi değerinin tanjant fonksiyonundan geçirilmesi üzerine kuruludur.

$$f(NET) = (e^{NET} + e^{-NET}) \,/\, (e^{NET} - e^{-NET}) \qquad (2.11)$$

Yukarıda bahsedilen aktivasyon fonksiyon örneklerinin grafikleri Şekil 8 üzerinde görselleştirilmiştir.



Şekil 2.7 Aktivasyon Fonksiyonu Örneklerinin Grafikleri (Öztemel, 2006)

Aktivasyon fonksiyonu aldığı değeri bir algoritma ile çıktıya dönüştürmektedir. Kullanılan transfer fonksiyonunun türüne göre çıktı değerinin aralığı [-1,1] veya [0,1]'dir. Bu fonksiyonlar genellikle doğrusal değildir. Doğrusal olmayan bir aktivasyon fonksiyonunu kullanılması ise yapay sinir ağlarının hem karmaşık hem de farklı alanlardaki problemleri çözümlemesine fayda sağlamaktadır.

Son olarak aşağıda işlem elemanı çıktı formülasyonu da verilen modelde bu değer doğrusal ya da doğrusal olmayan türevlenebilir bir transfer fonksiyonundan geçirilerek oluşturulmaktadır (Kaynar, 2008).

$$y = f(NET) = f\left(\sum_{i=1}^{n} x_i \omega_i + b\right) \qquad (2.12)$$

Bu değer dışa aktarışsa da ağın içerisinde tekrar kullanılabilmektedir. Ağ şeklinde görülen süreçte birden çok çıktı üretiliyor gibi görünse de tek çıktı değeri bulunur. Bu çıktı değeri birden fazla süreç elemanının girdisi olabilmektedir (Öztemel, 2006).

2.4 Sınıflandırma Performanslarının Ölçülmesi

Sınıflandırma algoritmasının performansını ölçmek için farklı ölçütler kullanılır ve bu ölçütler daha doğru sonuçlar üretmek için birlikte değerlendirilir.

Kappa İstatistiği: Kapa istatistiği (κ), ki-kare tablosuna dayanarak sınıflandırma modelinin tahmin performansını ölçer. Diğer bir değişle kategorik değişkenler arasındaki dereceler arası uyuşmayı ölçmek için kullanılan bir uygunluk istatistiğidir (Donner & Klar , 1996). κ istatistiğinin 1 değerine yaklaşması oranlar arasında yüksek uyum olduğuna işaret eder. Po iki kategorik değişken arasındaki gözlenen uyumu simgeler ve değişkenler arasındaki beklenen uyum, Pe, ölçüsünde değişir (Turanoglu-Bekar vd., 2016). Po için gözlenen uyum oranı, Pe için ise rastgele beklenen uyumluluk oranıdır diyebiliriz. κ istatisiği aşağıdaki formülle elde edilebilir:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{2.13}$$

MAE ve RMSE İstatistikleri: Ortalama mutlak hata (Mean Absolute Error - MAE) ve hata kareler ortalamasının karekökü (Root Mean Square Error - RMSE) istatistikleri modeltahminleri ve gözlenen değerler arasındaki farkı tanımlar (Willmott & Matsuura, 2005). MAE istatistiği tahmin değerleri ile gözlenen değerlerin mutlak farkının ortalaması alınarak hesaplanır. Farkların ağırlıkları eşittir, yani aritmatik ortalama ile hesaplanır. MAE ve RMSE arasındaki ilişki şu şekilde ifade edilebilir; MAE ≤ RMSE ve bu istatistikler aşağıdaki şekilde ifade edilir:

$$MAE = n^{-1}\sum_{i=1}^{n}\left|P_i - O_i\right| \tag{2.14}$$

$$RMSE = \sqrt{n^{-1}\sum_{i=1}^{n}\left|P_i - O_i\right|^2} \tag{2.15}$$

Burada Pi ve Oi (i=1,…,n) simgeleri sırasıyla tahmini ve gözlenendeğerleri ifade eder. Bu değerler arasındaki fark ise (Pi – Oi) modelin tahmin hatasını temsil eder.

TP Oranı, FP Oranı, Kesinlik ve MCC: Doğru oranı (True Positive-TP oranı), hassaslık olarak da adlandırılır, doğru olarak sınıflandırılmış pozitif değerlerin oranıdır. Yanlış oranı (True Negative-TN oranı), belirginlik olarak da adlandırılır, doğru olarak sınıflandırılan negatifdeğerleri simgeler. Yanlış pozitif oranı (False Positive - FP veya 2. tip hata) gerçekte doğru olan H_0 hipotezini reddetmemekle yapılan hatadır. Yanlış negative oranı ise (False Positive FN veya 2. tip hata) gerçekte yanlış olan H_0hipotezini reddetmemekle yapılan hata. Kesinlik (precision) pozitif olarak sınıflanan örneklerin gerçek pozitiflere oranıdır. Matthews korelasyon katsayısı (MCC) ise -1 ile 1 arasında değer alır ve hata matrisinin elemanları kullanılarak hesaplanır. Pozitif değere sahip MCC'ler doğru tahminlerin türetildiğini gösterir (MCC = 1 mükemmel kesinliği işaret eder). Formüller aşağıdaki gibidir:

$$TP_{rate} = TP / (TP + FN) \tag{2.16}$$

$$TN_{rate} = TN / (TN + FP) \tag{2.17}$$

$$precision = TP / (TP + FP) \tag{2.18}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{2.19}$$

F Kriteri: f – kriteri kesinlik ölçüsü ve TP oranının ağırlıklı ortalaması alınarak bulunur ve şu şekilde elde edilir:

$$f - measure = \frac{(1 + \beta^2) \times precision \times TP}{(\beta^2 \times precision + TP)} \tag{2.20}$$

Burada β kesinlik ve TP oranı için göreli önem değişkenidir. Genellikle bu değer 1 alınır.

ROC ve PRC Eğrisi: ROC (Receiver Operating Characteristic) eğrisinin altında kalan alan sınıflandırma algoritmasının başarısını tanımlar ve sınıflandırma algoritmasının performansını değerlendirmede yaygın bir şekilde kullanılmaktadır (Bradley, 1997). ROC alan eğrisi TP değerleri Y-ekseni'nde, (1-TN) değerleri ise X-ekseni'nde olacak şekilde bütün eşik değerleri için çizilir. Daha büyük bir alan kaplayan ROC eğrisi daha iyi bir modele işaret eder. PRC (Precision-recall curve)

ROC grafiği gibi iki boyutlu bir grafiktir fakat farklı olarak Y-ekseni'nde kesinlik (precision) oranını ve X- ekseni'nde geri çağırma oranını (rate of recall) kullanır. ROC eğrisince olduğu gibi PRC eğrisi altındaki alan sınıflandırma algoritmasının başarısı hakkında fikir verir. İki eğri arasındaki temel farklılık veri setindeki yapısal farklılıktan kaynaklanmaktadır (Davis & Goadrich, 2006). PRC eğrisinin yüksek oranda çarpık olduğu verilerde tercih edilmesi gerektiğini göstermiştir.

2.4.1 Akaike Bilgi Kriteri ve Bayesci Bilgi Kriteri

Akaike'nin Bilgi Kriteri (AIC), modelin farklı bir veri seti üzerinde test edildiği durumu simüle ederek elde edilen model kalitesinin bir ölçüsünü sağlar. Akaike'nin teorisine göre, en doğru model en küçük AIC'ye sahiptir. Hem model tahmini hem de doğrulama için aynı veri kümesini kullanırsanız, model sırasını ve dolayısıyla model yapısının esnekliğini artırdıkça uyum her zaman gelişir.(Ljung, 1999).

2.4.2 Ham Akaike Bilgi Kriteri (AIC)

$$AIC = N * log\left(det\left(\frac{1}{N}\sum_I^N \varepsilon\left(t,\widehat{\theta_N}\right)\left(\varepsilon\left(t,\widehat{\theta_N}\right)\right)^T\right)\right) + 2n_p + N * (n_y * (\log(2\pi) + 1)) \quad (2.21)$$

2.4.3 Normalleştirilmiş Akaike Bilgi Kriteri (AICc)

$$nAIC = log\left(det\left(\frac{1}{N}\sum_I^N \varepsilon\left(t,\widehat{\theta_N}\right)\left(\varepsilon\left(t,\widehat{\theta_N}\right)\right)^T\right)\right) + \frac{2n_p}{N} \quad (2.22)$$

2.4.4 Bayesci Bilgi Kriteri (BIC)

$$BIC = N * log\left(det\left(\frac{1}{N}\sum_I^N \varepsilon\left(t,\widehat{\theta_N}\right)\left(\varepsilon\left(t,\widehat{\theta_N}\right)\right)^T\right)\right) + N * \left(n_y * \log(2\pi) + 1\right) + n_p * \log(N) \quad (2.23)$$

2.5 Öznitelik Seçimi

Yapay zeka çalışmalarında makine öğrenmesi yöntemlerini kullanırken dikkat edilmesi gereken konulardan bir tanesi de modele eklenecek özniteliklerin modele katkısı ve anlamlılığıdır. Bu yüzden bir model çalıştırılmadan önce o modelde yer alacak özniteliklerin seçimine özen gösterilmeli ve modelin başarısını artıracak öznitelikler modele dahil edilmelidir. Modelde anlamsız öznitelikler kullanmak hem modelin çalışması süresini artıracaktır hem de elde etmek istediğimiz basit ama yüksek başarıdan bizi uzaklaştıracaktır. Öznitelik seçimi yapılırken kullanılabilecek birçok farklı yöntem bulunmaktadır. Bu yöntemlerden bazıları Correlation attribute, Cfs Subset, One-R, Gain Ratio ve T-testi, Fisher skor, ki-kare öznitelik

seçimi yöntemleridir. Bu tez kapsamında yapılan çalışmada t-testinden faydalanılarak öznitelik seçimi yapılmış olup uygulama kısmında bu sonuçlara detaylıca yer verilmiştir.

2.5.1 T-testi ile öznitelik seçimi

Öznitelik seçiminde t-testi yöntemi her bir özelliğe ait ortalama ve standart sapma değerlerini kullanılarak hedef iki gruba ait değerlerin farklılaşmasına bakar. Her bir öznitelik için yapılan bağımsız t-testi sonuçlarında iki sınıf için anlamlı fark bulunuduğu noktada öznitelik modele dahil edilir. İki sınıf için anlamlı fark çıkmıyorsa bu özniteliğin modele dahil edilmesine gerek bulunmaz.

T-testinin uygulanabilmesi için özniteliklerin alınmış olduğu popülasyon dağılımlarının normal olması, gözlemlerin bağımsız olması ve varyansların homojen olması gibi varsayımlar sağlanmalıdır.(Eltas, 2021)

2.6 Lojistik Regresyon

Lojistik regresyon analizi, sonuç değişkeninin kategorik olduğu araştırma durumlarında yapılan çoklu regresyon analizi tekniğidir. Pratikte, kategorik sonuçlar içeren durumlar oldukça yaygındır. Örneğin, bir eğitim programının değerlendirilmesinde, başarı / başarısızlığın ya da geliştirilmiş / geliştirilmemişin ikili sonucu için öngörüler yapılabilir. Benzer şekilde, tıbbi bir ortamda, bir sonuç varoluş / hastalık yokluğu olabilir. Lojistik regresyonda cevap değişkeninin beklenen değerleri olasılık olarak elde edilir. “Lojistik regresyon modeli logaritmik dönüşümler sonucu doğrusal hale gelen, bağımlı değişkeni kategorik, yani yapay değişken olan bir modeldir. Lojistik regresyon bağımlı ve bağımsız değişkenler arasında logit bir ilişki olduğunu varsayar; dolayısıyla lojistik regresyon doğrusal olmayan modeller üretebilir.” (Baydemir, 2014)

Lojistik regresyon, sıradan en küçük kareler regresyonunun sınırlayıcı varsayımlarının çoğunu yenmek için kullanılan popüler bir yöntemdir. Bu varsayımlar lojistik regresyon için şu şekilde sıralanmıştır:

- Lojistik regresyon bağımlı ve bağımsız değişkenler arasında doğrusal bir ilişki kurmaz.
- Bağımlı değişkenlerin normal dağılıma sahip olması gerekmez.

- Bağımlı değişkenin, bağımsızların her kategorisi için homoskedastik olması gerekmez. Homoskedastik olması varyansların homojen olması anlamına gelir.
- Hata terimlerinin normal dağılması gerekmez.
- Lojistik regresyon, bağımsız değişkenlerin aralık ölçekte olmasını istemez.

2.6.1 İkili (Binary) Lojistik Regresyon

Regresyon uygulamalarında, ilgili bağımlı değişken, sadece iki olası sonuca sahiptir ve bu nedenle, 0 ve 1 değerlerini alarak ikili bir gösterge değişkeniyle ifade edilebilir. (Bar-Hillel & Neter, 1996).

2.6.2 Basit Lojistik Regresyon

X ve Y ikili sonuçlar içeren bir veri kümesi olsun. X' deki her x_i değişkeni için sonuç y = 1 veya y = 0'dır.Bağımsız değişken y_i=1 olduğunda pozitif sınıfa, y_i=0 olduğundaysa negatif sınıfa ait olduğu söylenebilir. (Komarek & Moore, 2005)

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$Y_i = 0,1 \qquad i = 1,...,n \tag{2.24}$$

$$\mathrm{E}(\varepsilon) = 0$$

olduğunda;

$$\mathrm{E}\{Y_i\} = b_0 + b_1 X_i \tag{2.25}$$

Y_i olasılık dağılımı aşağıdaki gibi olan Bernoulli dağılımına sahiptir;

$$Y_i = 1 \quad P(Y_i = 1) = \pi_i \tag{2.26}$$

$$Y_i = 0 \quad P(Y_i = 0) = 1 - \pi_i \tag{2.27}$$

Bernoulli dağılımının beklenen değerinden ve varyansından faydalanıldığında;

$$\mathrm{E}\{Y_i\} = \beta_0 + \beta_1 X_i = \pi_i \tag{2.28}$$

$$\sigma^2\{Y_i\} = \mathrm{E}\{[Y_i - \mathrm{E}\{Y_i\}]^2\} = \mathrm{E}\{Y_i\}[1 - \mathrm{E}\{Y_i\}] \tag{2.29}$$

$$\varepsilon_i = Y_i - \pi_i \tag{2.30}$$

2.6.3 Lojistik Regresyon Modeli

Lojistik regresyon modeli şu şekilde formüle edilmiştir;(Harrell, 2015)

$$\pi(x_i) = P(y_i = 1 \vdots x_i) \tag{2.31}$$

$$= [1 + \exp(-X^T \beta)]^{-1} \tag{2.32}$$

$$X^T \beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_{p-1} x_{p-1} \tag{2.33}$$

$$\beta_{pX_1} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}, \quad X_{pX_1} = \begin{bmatrix} 1 \\ X_1 \\ \vdots \\ X_{p-1} \end{bmatrix}, \quad X_{ipX_1} = \begin{bmatrix} 1 \\ X_{i1} \\ \vdots \\ X_{i,p-1} \end{bmatrix} \tag{2.34}$$

$\pi(x_i)$, p bağımsız değişkenlere bağlı bir olayın olma olasılığını temsil eder.

$$\pi(x_i) = [1 + \exp(-X^T \beta)]^{-1} \tag{2.35}$$

$$= \frac{1}{1 + \exp(-X^T \beta)}$$

$$1 - \pi(x_i) = 1 - \frac{1}{1 + \exp(-X^T \beta)}$$

$$= \frac{[1 + \exp(-X^T \beta)] - 1}{1 + \exp(-X^T \beta)}$$

$$= \frac{\exp(-X^T \beta)}{1 + \exp(-X^T \beta)}$$

$$\ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \log it[\pi(x_i)]$$

$$\frac{\pi(x_i)}{1 - \pi(x_i)} = [\exp(-X^T \beta)]^{-1}$$

$$= X^T \beta$$

Lojistik regresyon dikotom bağımsız değişkenler ve tahminciler arasındaki ilişkiyi odds oranı yardımıyla belirler. Odds oranı, bir olayın oluşma olasılığı, olayın gerçekleşme ihtimaline bölünmesiyle elde edilir (Hayden, 2008).

$$odds = \frac{P(başarı)}{P(başarı)}$$

$$= \frac{\pi(x_i)}{1-\pi(x_i)}$$

$$= [\exp(-X^T\beta)]^{-1} \tag{2.36}$$

$$\text{Odds oranı} = \frac{Başarının\ odds\ değeri}{başarısızlığın\ odds\ değeri} \tag{2.37}$$

3
UYGULAMA

3.1 Uygulama Yöntemi ve Kullanılan Modeller

Bu tez kapsamında bir telekomünikasyon firmasına ait müşterilerin farklı verilerinden faydalanılarak müşteri kaybı tahminlenmesi hedeflenmiştir. Çalışma Matlab programı kullanılarak gerçekleştirilmiştir. Bu uygulamada kullanılan yöntemler yapay sinir ağları ve lojistik regresyon yöntemleridir. Lojistik regresyon kapsamında %90 eğitim %10 test, %80 eğitim %20 test, %70 eğitim %30 test ve %60 eğitim %40 test olmak üzere dört farklı veri bölümlemesi ile model çalıştırılmıştır. Yapay sinir ağları kapsamında ise 7 farklı yapay sinir ağı fonksiyonu kullanılmıştır. Her fonksiyon için hem %75 eğitim %15 test %15 geçerlilik ve %80 eğitim %10 test %10 geçerlilik veri bölümlemeleri çalıştırılmıştır. Ayrıca her veri bölümlemesi için 2, 3, 4, 5 olmak üzere farklı katmanlar kullanılmıştır. Bu kapsamda model yapay sinir ağları için toplamda 56 kez çalıştırılmıştır. Modeller doğruluk oranı, ortalama hata kareleri, Akaike ve Bayesgil kriterleri göz önünde bulundurularak değerlendirilmiştir. Lojistik regresyon için en iyi sonuç veren model, yapay sinir ağları için en iyi sonuç veren model ile karşılaştırılmıştır. Matlab programlama dili veri üzerinde farklı modeller yardımıyla yapay sinir ağları çalışması yapmayı kolaylaştıracak hazır algoritmalar yer almaktadır. Bu algoritmalar sayesinde derin istatistiksel bilgiye sahip olmadan farklı şekillerde modeli çalıştırmak mümkündür. Bu tez kapsamında yapay sinir ağları için kullanılan farklı algoritmaların açıklamaları aşağıdaki gibidir:

3.1.1 Trainscg - Ölçekli Eşlenik Gradyan Geri Yayılım

Trainscg, ağırlıkları ve sapmaları ölçekli eşlenik gradyan geri yayılım metoduna göre güncelleyen ağ eğitim yöntemidir.

trainscg, ağırlığı, ağ girişi ve transfer fonksiyonları türev fonksiyonlara sahip olduğu sürece herhangi bir ağı eğitebilir. Geri yayılım, ağırlık ve sapma değişkenleri X'e göre performansın türevlerini hesaplamak için kullanılır.

Ölçeklendirilmiş eşlenik gradyan algoritması, traincgp, traincgf ve traincgb'de olduğu gibi eşlenik yönlere dayanır, ancak bu algoritma her yinelemede bir satır araması yapmaz.

Aşağıdaki koşullardan herhangi biri gerçekleştiğinde eğitim durur:

Maksimum döngü sayısına (tekrar) ulaşıldığında.
Maksimum süre aşıldığında.
Performans amaç için minimize edildiğinde.
Performans gradyanı minimum gradyan değerinin altına düştüğünde.
Geçerlilik performansı, son düşüşünden bu yana maksimum hata süresinden daha fazla arttığında (geçerlilik (validation) kullanılırken).

3.1.2 Trainbfg - BFGS Yarı-Newton Geri Yayılımı

Newton'un yöntemi, hızlı optimizasyon için eşlenik gradyan yöntemlerine bir alternatiftir. Newton yönteminin basit adımı aşağıdaki gibidir:

Burada, ağırlıkların ve sapmaların mevcut değerlerindeki performans endeksinin Hessian matrisidir. Newton'un yöntemi genellikle eşlenik gradyan yöntemlerinden daha hızlı birleşir. Ne yazık ki, ileri beslemeli sinir ağları için Hessian matrisini hesaplamak karmaşık ve zaman alıcıdır.

trainbfg, ağırlığı, ağ girişi ve transfer fonksiyonları türev fonksiyonlara sahip olduğu sürece herhangi bir ağı eğitebilir.

Aşağıdaki koşullardan herhangi biri gerçekleştiğinde eğitim durur:

Maksimum dönem sayısına (tekrar) ulaşıldığında.
Maksimum süre aşıldığında.
Performans amaç için minimize edildiğinde.
Performans gradyanı minimum gradyan değerinin altına düşüyor.
Geçerlilik performansı, son düşüşünden bu yana maksimum hata süresinden daha fazla arttığında (geçerlilik kullanılırken).

3.1.3 Traingd - Gradyan Azalan Geri Yayılım

traingd, ağırlığı, ağ girişi ve transfer fonksiyonları türev fonksiyonlara sahip olduğu sürece herhangi bir ağı eğitebilir.

Geri yayılım, ağırlık ve sapma değişkenleri X'e göre performansın türevlerini hesaplamak için kullanılır. Her değişken, gradyan azalışına göre ayarlanır:

$$dx = lr * dperf \ / \ dX \tag{3.1}$$

Aşağıdaki koşullardan herhangi biri gerçekleştiğinde eğitim durur:

Maksimum döngü (epoch) sayısına (tekrar) ulaşıldığında.
Maksimum süre aşıldığında.
Performans amaç için minimize edildiğinde.
Performans gradyanı minimum gradyan değerinin altına düştüğünde.
Geçerlilik performansı, son düşüşünden bu yana maksimum hata süresinden daha fazla arttığında (geçerlilik kullanılırken).

3.1.4 Traingda - Uyarlanabilir öğrenme Oranı ile Geri Yayılımlı Gradyan Azalışı

Traingda yönteminde; standart en dik inişle, öğrenme oranı eğitim boyunca sabit tutulur. Algoritmanın performansı, öğrenme hızının uygun şekilde ayarlanmasına çok duyarlıdır. Öğrenme hızı çok yüksek ayarlanırsa, algoritma salınım yapabilir ve kararsız hale gelebilir. Öğrenme hızı çok küçükse, algoritmanın yakınsaması çok uzun sürer. Eğitimden önce öğrenme oranı için en uygun ayarı belirlemek pratik değildir ve aslında, algoritma performans yüzeyinde hareket ettikçe, eğitim süreci sırasında optimum öğrenme oranı değişir.

Eğitim sürecinde öğrenme hızının değişmesine izin verilirse en dik iniş algoritmasının performansı arttırılabilir. Uyarlanabilir öğrenme oranı, öğrenmeyi sabit tutarken öğrenme adımı boyutunu olabildiğince büyük tutmaya çalışır. Öğrenme hızı, yerel hata yüzeyinin karmaşıklığına duyarlı hale getirilir.

Uyarlanabilir öğrenme oranı, eğitim tarafından kullanılan eğitim prosedüründe bazı değişiklikler gerektirir. İlk olarak, ilk ağ çıktısı ve hata hesaplanır. Her deneme için mevcut öğrenme oranı kullanılarak yeni ağırlıklar ve sapmalar hesaplanır. Daha sonra yeni çıktılar ve hatalar hesaplanır.

traingda, ağırlığı, ağ girişi ve transfer fonksiyonları türev fonksiyonlara sahip olduğu sürece herhangi bir ağı eğitebilir.

Geri yayılım, ağırlık ve sapma değişkenleri X'e göre performans türevlerini (dperf) hesaplamak için kullanılır. Her değişken, gradyan azalışına göre ayarlanır:

$$dX = lr * dperf / dX \quad (3.2)$$

Aşağıdaki koşullardan herhangi biri gerçekleştiğinde eğitim durur:

Maksimum döngü sayısına (tekrar) ulaşıldığında.
Maksimum süre aşıldığında.
Performans amaç için minimize edildiğinde.
Performans gradyanı minimum gradyan değerinin altına düştüğünde.
Geçerlilik performansı, son düşüşünden bu yana maksimum hata süresinden daha fazla arttığında (geçerlilik kullanılırken).

3.1.5 Traingdx - Momentum ve Uyarlanabilir Öğrenme Oranı Geri Yayılımı ile Gradyan Azalışı

Traingdx işlevi, uyarlanabilir öğrenme oranını momentum eğitimi ile birleştirir. Ek bir eğitim parametresi olarak momentum katsayısına sahip olması dışında, traingda ile aynı şekilde düşünülebilir.

traingdx, ağırlığı, ağ girişi ve transfer fonksiyonları türev fonksiyonlara sahip olduğu sürece herhangi bir ağı eğitebilir.

Geri yayılım, ağırlık ve sapma değişkenleri X'e göre performansın türevlerini hesaplamak için kullanılır. Her değişken momentum ile gradyan azalışına göre ayarlanır,

$$dX = mc * dXprev + lr * mc * dperf / dX \quad (3.3)$$

burada dXprev (x'e göre performans türevi) ağırlık veya sapmada önceki değişikliktir.

Aşağıdaki koşullardan herhangi biri gerçekleştiğinde eğitim durur:

Maksimum dönem sayısına (tekrarlar) ulaşıldığında.
Maksimum süre aşıldığında.
Performans amaç için minimize edildiğinde.
Performans gradyanı minimum gradyan değerinin altına düştüğünde.
Geçerlilik performansı, son düşüşünden bu yana maximum hata süresinden daha fazla arttı (geçerlilik kullanılırken).

3.1.6 Trainr - Öğrenme İşlevleriyle Rastgele Sıralı Artımlı Eğitim

Her döngü (epoch) için, tüm eğitim vektörleri (veya dizileri), her bir sunumdan sonra ağ ve ağırlık ve sapma değerleri uygun şekilde güncellenen farklı bir rastgele sırada bir kez sunulur.

Bu koşullardan herhangi biri karşılandığında eğitim durur:

Maksimum döngü sayısına (tekrar) ulaşıldığında.
Performans amaç için minimize edildiğinde.
Maksimum süre aşıldığında.

3.1.7 Trainb - Ağırlık ve Sapma Öğrenme Kurallarıyla Toplu Eğitim

Her bir ağırlık ve sapma, her dönemden sonra öğrenme işlevine göre güncellenir (bir giriş vektörleri kümesinin tamamından geçer).

Bu koşullardan herhangi biri karşılandığında eğitim durur:

Maksimum döngü sayısına (tekrar) ulaşıldı.
Performans amaç için minimize edildiğinde.
Maksimum süre aşıldığında.
Geçerlilik performansı, son düşüşünden bu yana maximum hata süresinden daha fazla arttığında (geçerlilik kullanılırken).

3.2 Kullanılan Değişkenler

Analizde kullanılan değişkenlerin isimleri ve açıklamaları aşağıdaki gibidir:

Tablo 3.1 Kullanılan Değişkenler ve Açıklamaları

Değişken Adı	Açıklaması	Değişken Tipi
Hesap Uzunluğu	Müşterinin müşteri olma süresi	Tam sayı
Uluslararası Plan	Müşterinin uluslararası konuşma planı	Kategorik
Sesli Mesaj Planı	Müşterinin sesli mesaj planı	Kategorik
Sesli Mesaj Sayısı	Müşterinin kullandığı sesli mesaj sayısı	Tam sayı
Toplam Gün içi Dakika	Müşterinin toplamda yaptığı gün içi konuşma dakikası	Tam sayı
Toplam Gün için Arama	Müşterinin toplamda gün içinde yaptığı arama sayısı	Tam sayı
Toplam Gün içi Ücret	Müşterinin gün içi konuşmaları için ödediği ücret	Tam sayı
Toplam Akşam Dakika	Müşterinin toplamda akşam saatlerinde yaptığı konuşma dakikası	Tam sayı
Toplam Akşam Arama	Müşterinin toplamda akşam saatlerinde yaptığı arama sayısı	Tam sayı

Tablo 3.1 Kullanılan Değişkenler ve Açıklamaları (devamı)

Toplam Akşam Ücret	Müşterinin akşam saati konuşmaları için ödediği ücret	Tam sayı
Toplam Gece Dakika	Müşterinin toplamda yaptığı gece saati konuşma dakikası	Tam sayı
Toplam Gece Arama	Müşterinin toplamda gece saatlerinde yaptığı arama sayısı	Tam sayı
Toplam Gece Ücret	Müşterinin gece saatlerindeki konuşmaları için ödediği ücret	Tam sayı
Toplam Uluslararası Dakika	Müşterinin toplamda yaptığı uluslararası konuşma dakikası	Tam sayı
Toplam Uluslararası Arama Sayısı	Müşterinin toplamda yaptığı uluslararası arama sayısı	Tam sayı
Toplam Uluslararası Ücret	Müşterinin gün içi konuşmaları için ödediği ücret	Tam sayı
Müşteri Hizmetleri Çağrıları	Müşteri hizmetleri ile görüşme miktarı	Tam sayı
Müşteri Kaybı	Müşteri kaybı sonucu	Kategorik

3.3 T-Testi ile Değişken Seçimi

Analizde ilk olarak anlamsız olan değişkenleri modelden elemek adına tüm değişkenler için t-testi yapılmıştır. T testi sonuçları aşağıdaki gibidir:

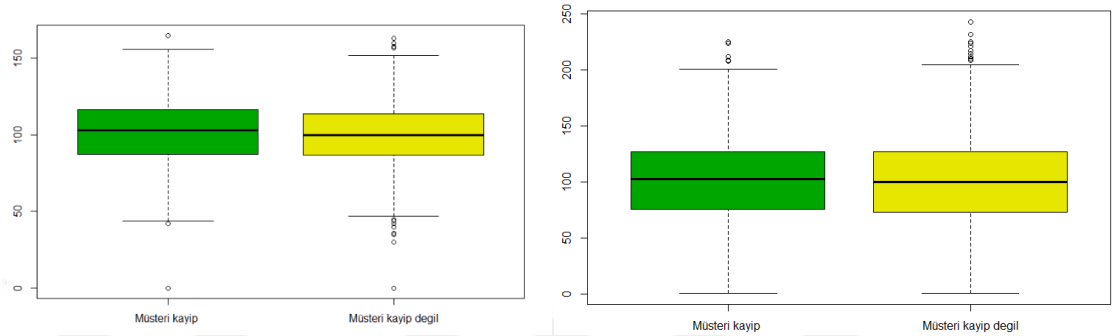
Tablo 3.2 T-testi Sonuçları

Değişkenler	**Kayıp**	**Sayı**	**Ortalama**	**Standart Sapma**	**T değeri**	**2 kuyruklu anlamlılık**
Hesap Uzunluğu	0 Hayır	2850	100,79	39,882	-0,955	**0,340**
	1 Evet	483	102,66	39,468		
Uluslararası Plan	0 Hayır	2850	0,07	0,247	-15,531	**0,000**
	1 Evet	483	0,28	0,451		
Sesli Mesaj Planı	0 Hayır	2850	0,30	0,456	5,926	**0,000**
	1 Evet	483	0,17	0,372		
Sesli Mesaj Sayısı	0 Hayır	2850	8,60	13,913	5,200	**0,000**
	1 Evet	483	5,12	11,860		
Toplam Gün içi Dakika	0 Hayır	2850	175,18	50,18	-12,098	**0,000**
	1 Evet	483	206,91	69,00		
Toplam Gün için Arama	0 Hayır	2850	100,28	19,801	-1,066	**0,287**

Tablo 3.2 T-testi Sonuçları (devamı)

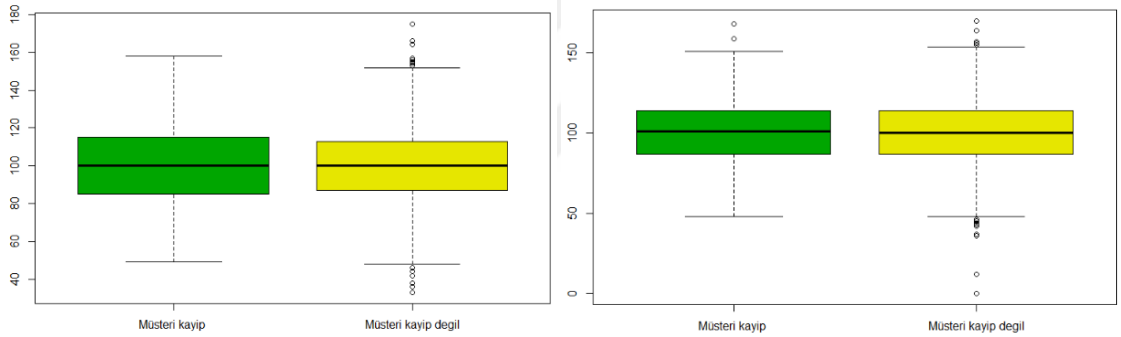
	1 Evet	483	101,34	21,582		
Toplam Gün içi Ücret	0 Hayır	2850	29,78	8,53	-12,098	**0,000**
	1 Evet	483	35,18	11,73		
Toplam Akşam Dakika	0 Hayır	2850	199,04	50,29	-5,379	**0,000**
	1 Evet	483	212,41	51,73		
Toplam Akşam Arama	0 Hayır	2850	100,04	19,958	-0,533	**0,594**
	1 Evet	483	100,56	19,725		
Toplam Akşam Ücret	0 Hayır	2850	16,92	4,27	-5,378	**0,000**
	1 Evet	483	18,05	4,40		
Toplam Gece Dakika	0 Hayır	2850	200,13	51,11	-2,050	**0,040**
	1 Evet	483	205,23	47,13		
Toplam Gece Arama	0 Hayır	2850	100,06	19,506	-0,354	**0,723**
	1 Evet	483	100,40	19,951		
Toplam Gece Ücret	0 Hayır	2850	9,01	2,30	-2,050	**0,040**
	1 Evet	483	9,24	2,12		
Toplam Uluslararası Dakika	0 Hayır	2850	10,16	2,78	-3,948	**0,000**
	1 Evet	483	10,70	2,79		
Toplam Uluslararası Arama Sayısı	0 Hayır	2850	4,53	2,442	3,054	**0,002**
	1 Evet	483	4,16	2,552		
Toplam Uluslararası Ücret	0 Hayır	2850	2,74	0,75	-3,949	**0,000**
	1 Evet	483	2,89	0,75		
Müşteri Hizmetleri Çağrıları	0 Hayır	2850	1,45	1,164	-12,319	**0,000**
	1 Evet	483	2,23	1,853		

Yapılan t testi sonucunda hesap uzunluğu, toplam gün içi arama, toplam akşam arama ve toplam gece arama değişkenlerinin anlamsız olduğu görülmüştür. Bu yüzden yapay sinir ağları ve lojistik regresyon modellerinde bu değişkenler hariç tutulmuştur. Aşağıda t-testi sonucuna göre anlamsız çıkan değişkenlerin kutu grafikleri gösterilmiştir:



Hesap Uzunluğu Gün İçi Konuşma

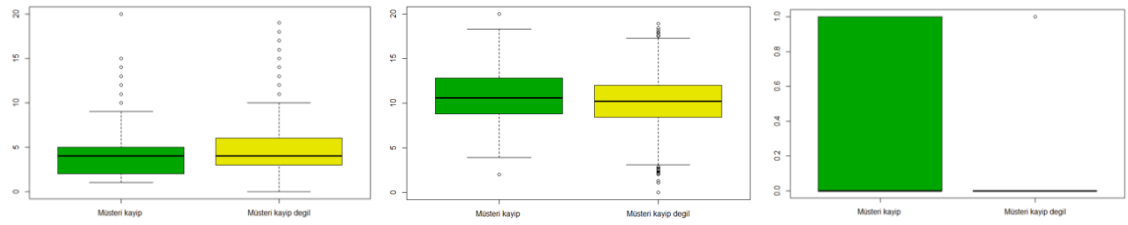
Şekil 3.1 Hesap Uzunluğu ve Gün İçi Konuşma Kutu Grafikleri



Akşam Konuşma Gece Konuşma

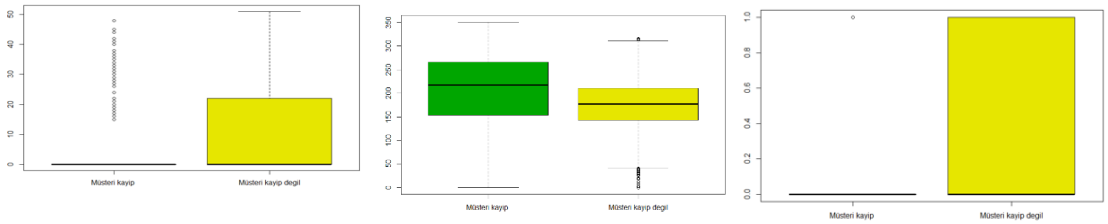
Şekil 3.2 Akşam Konuşma ve Gece Konuşma Kutu Grafikleri

Ayrıca diğer değişkenler için de kutu grafikleri aşağıda gösterilmiştir:



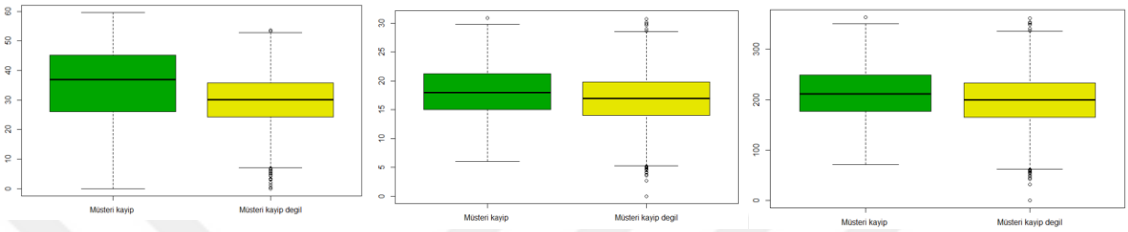
Uluslararası Plan Uluslararası Dakika Uluslararası Konuşma

Şekil 3.3 Uluslararası Plan, Dakika ve Konuşma Kutu Grafikleri



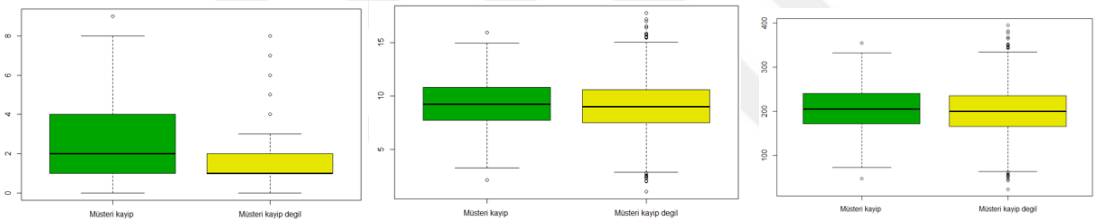
Sesli Mesaj Plan Sesli Mesaj Sayısı Gün Içi Dakika

Şekil 3.4 Sesli Mesaj Plan, Mesaj Sayısı ve Gün İçi Dakika Kutu Grafikleri



Gün İçi Konuşma Ücret Akşam Konuşma Akşam Ücret

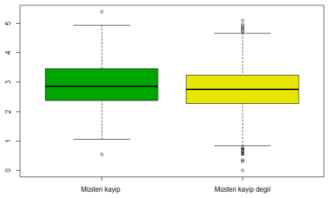
Şekil 3.5 Gün İçi Konuşma, Akşam Konuşma ve Akşam Ücret Kutu Grafikleri



Müşteri Hizmetleri

Gece Konuşma Dakika Gece Ücret Arama

Şekil 3.6 Gece Konuşma Dakika, Ücret ve Müşteri Hizmetleri Arama Kutu Grafikleri



Uluslararası Ücret

Şekil 3.7 Uluslararası Ücret Kutu Grafiği

Kutu grafiklerinden faydalanarak değişkenlerin en yüksek, en düşük noktalarını ya da çeyreklerine ait değerlerini görselleştirmek mümkündür. Yukarıdaki kutu grafiklerinden faydalanarak modelde anlamlı olup olmayacaklarını gözle görebiliriz. Örneğin T testi sonucunda anlamsız çıkan değişkenlerin kutu grafiklerinden de iki

grubun ortalamalarının neredeyse aynı olduğunu söylemek mümkündür. Uluslararası plan ve uluslararası dakikaya baktığımızda yine verilerin dağılımlarının yakın olduğunu söyleyebiliriz ama ortalamalar arasında ufak farklar olduğundan t-testinde anlamlı çıktıkları sonucuna ulaşabiliriz. Aynı şekilde uluslararası ücret, gece ücret, müşteri hizmetleri arama, akşam konuşma ve akşam ücret için de benzer bir durum gözlenmiştir. Uluslararası konuşma, sesli mesaj plan ve gün içi dakikaya baktığımızda ise iki grup arasında çok büyük fark olduğunu söylemek mümkündür. Sesli mesaj sayısı, gece konuşma dakika ve gün içi konuşma ücrette ise kaybedilen müşterilerin kutu gösterimlerinde verinin daha geniş bir alanda dağıldığı görülmektedir.

3.4 Yapay Sinir Ağları Modellerinin Uygulanması

Yukarıda bahsedilen metodoloji ile çalışılan yapay sinir ağları modellerinin çıktıları aşağıdaki tabloda yer almaktadır.

Tablo 3.3 Yapay Sinir Ağları Model Çıktıları

									MSE				Doğruluk oranı - Accuracy Rate			
				Katman	İterasyon	AIC	AICc	BIC	Tüm	Eğitim	Test	Geçerlilik	Tüm	Eğitim	Test	Geçerlilik
trainscg	0,7	0,15	0,15	2	71	-4572,64	-2236,54	-4376,97	0,07	0,07	0,06	0,07	92,5	93	92,5	91,8
	0,7	0,15	0,15	3	20	-3852,70	-1515,38	-3564,96	0,09	0,09	0,09	0,09	89,3	89,4	89,2	89,2
	0,7	0,15	0,15	4	28	-3611,75	-1272,73	-3231,93	0,10	0,10	0,10	0,10	85,8	85,4	85,8	86
	0,7	0,15	0,15	5	52	-4328,86	-1987,66	-3856,96	0,07	0,07	0,08	0,04	92,9	92,2	92,5	95,4
	0,8	0,1	0,1	2	44	-5458,10	-2788,14	-5257,88	0,06	0,06	0,06	0,06	92,7	93,4	92,6	92,8
	0,8	0,1	0,1	3	64	-5676,75	-3005,73	-5382,32	0,06	0,06	0,06	0,06	92,8	91,6	93	92,2
	0,8	0,1	0,1	4	55	-5409,16	-2736,66	-5020,51	0,06	0,06	0,06	0,06	92,9	92,8	92,9	93,1
	0,8	0,1	0,1	5	13	-3856,75	-1182,36	-3373,88	0,11	0,11	0,11	0,10	85,6	83,8	85,6	87,7
trainbfg	0,7	0,15	0,15	2	51	-4216,14	-1880,04	-4020,47	0,09	0,08	0,12	0,08	91,4	87,8	92	92,4
	0,7	0,15	0,15	3	71	-4572,64	-2236,54	-4376,97	0,07	0,07	0,06	0,07	92,5	93	92,5	91,8
	0,7	0,15	0,15	4	75	-4291,87	-1952,85	-3912,05	0,07	0,08	0,08	0,06	92,1	91	92,1	93,6
	0,7	0,15	0,15	5	67	-4356,22	-2015,02	-3884,32	0,08	0,07	0,08	0,08	91,8	91,2	92,2	90,4
	0,8	0,1	0,1	2	74	-5435,99	-2766,03	-5235,77	0,06	0,06	0,06	0,07	90,9	91	90,9	91,6
	0,8	0,1	0,1	3	82	-5462,06	-2791,04	-5167,63	0,06	0,06	0,08	0,06	92	91	92,3	90,7
	0,8	0,1	0,1	4	43	-5165,28	-2492,77	-4776,62	0,07	0,07	0,09	0,08	92,3	91,3	92,7	91,9
	0,8	0,1	0,1	5	47	-4974,00	-2299,60	-4491,13	0,07	0,07	0,10	0,07	91,5	88	91,8	92,5
traingd	0,7	0,15	0,15	2	1000	-3098,62	-762,52	-2902,95	0,13	0,13	0,12	0,12	85,5	86,8	85,2	85,8
	0,7	0,15	0,15	3	418	-3165,98	-828,66	-2878,24	0,12	0,12	0,14	0,11	85,3	83	85,4	87,4
	0,7	0,15	0,15	4	1000	-2987,96	-648,93	-2608,13	0,13	0,13	0,14	0,11	85,5	84,2	85,2	88,6
	0,7	0,15	0,15	5	1000	-3155,37	-814,17	-2683,47	0,12	0,12	0,13	0,11	85,5	84,2	85,5	87
	0,8	0,1	0,1	2	1000	-3596,19	-926,23	-3395,97	0,13	0,13	0,13	0,14	85,5	85,9	85,8	82,9
	0,8	0,1	0,1	3	1000	-3665,48	-994,45	-3371,05	0,12	0,12	0,12	0,12	85,5	85,3	85,6	85
	0,8	0,1	0,1	4	1000	-3562,74	-890,24	-3174,09	0,12	0,13	0,12	0,12	85,5	85,9	85,5	85,6
	0,8	0,1	0,1	5	1000	-3790,47	-1116,07	-3307,59	0,11	0,11	0,11	0,11	85,5	85,9	85,5	85,6
traingda	0,7	0,15	0,15	2	149	-3926,28	-1590,18	-3730,61	0,10	0,09	0,11	0,11	87,9	87,2	88,4	86,4
	0,7	0,15	0,15	3	152	-3761,58	-1424,26	-3473,84	0,09	0,10	0,08	0,09	86,8	89	86,2	87,4
	0,7	0,15	0,15	4	172	-3904,68	-1565,66	-3524,86	0,09	0,09	0,08	0,09	89,9	91	89,9	89
	0,7	0,15	0,15	5	168	-3770,27	-1429,06	-3298,36	0,09	0,09	0,07	0,09	88,3	91,2	87,8	87,6

Tablo 3.3 Yapay Sinir Ağları Model Çıktıları (devamı)

traingda	0,8	0,1	0,1	2	148	-4499,36	-1829,40	-4299,14	0,09	0,09	0,09	0,10	89,9	90,4	90	88,3
	0,8	0,1	0,1	3	158	-4186,09	-1515,07	-3891,66	0,10	0,10	0,11	0,08	87,2	85,3	87	91
	0,8	0,1	0,1	4	435	-5409,06	-2736,55	-5020,41	0,06	0,06	0,06	0,06	92,9	93,7	92,7	93,7
	0,8	0,1	0,1	5	156	-4561,42	-1887,02	-4078,54	0,09	0,09	0,09	0,11	90	89,2	90,6	86,8
traingdx	0,7	0,15	0,15	2	163	-3880,75	-1544,66	-3685,09	0,09	0,09	0,10	0,09	88,7	88,4	88,9	88,2
	0,7	0,15	0,15	3	187	-4324,96	-1987,63	-4037,21	0,08	0,08	0,08	0,09	91,6	91,0	92,1	90,2
	0,7	0,15	0,15	4	158	-3761,96	-1422,94	-3382,14	0,09	0,09	0,09	0,10	87,6	87,2	87,7	87,8
	0,7	0,15	0,15	5	73	-3247,31	-9061,10	-2775,41	0,12	0,12	0,15	0,13	85,5	81,8	86,6	84,2
	0,8	0,1	0,1	2	165	-4487,28	-1817,32	-4287,06	0,09	0,09	0,09	0,10	89,6	89,5	89,8	87,7
	0,8	0,1	0,1	3	185	-4921,30	-2250,27	-4626,86	0,08	0,08	0,08	0,09	91,9	92,2	92	90,4
	0,8	0,1	0,1	4	95	-3762,11	-1089,61	-3373,46	0,12	0,12	0,13	0,12	85,5	84,4	85,6	85,9
	0,8	0,1	0,1	5	196	-4944,01	-2269,61	-4461,13	0,07	0,07	0,07	0,07	92,1	91,9	92,2	91,9
trainr	0,7	0,15	0,15	2	36	-4783,04	-2398,97	-4586,68	0,06	0,07	0,06	0,06	92,1	92,6	92,1	92
	0,7	0,15	0,15	3	89	-4625,82	-2240,54	-4337,06	0,07	0,07	0,08	0,06	93	92,4	93,1	93,3
	0,7	0,15	0,15	4	99	-4999,80	-2612,86	-4618,63	0,06	0,06	0,07	0,06	92,9	90,1	93,4	93,3
	0,7	0,15	0,15	5	34	-4924,76	-2535,69	-2535,69	0,06	0,06	0,07	0,07	92,8	91,4	93,1	92,9
	0,8	0,1	0,1	2	83	-4728,08	-2058,12	-4527,86	0,08	0,08	0,10	0,07	91,1	89,2	91,2	92,2
	0,8	0,1	0,1	3	92	-5531,38	-2860,35	-5236,94	0,06	0,06	0,07	0,06	92	91	92,2	91,9
	0,8	0,1	0,1	4	55	-5488,92	-2816,41	-5100,27	0,06	0,06	0,07	0,07	92,8	92,2	93,1	90,7
	0,8	0,1	0,1	5	67	-5744,86	-3070,46	-5261,98	0,05	0,05	0,04	0,05	93,5	94,9	93,2	94,3
trainb	0,7	0,15	0,15	2	1000	-3305,91	-969,81	-3110,24	0,12	0,12	0,11	0,13	85,6	87,4	85,8	83
	0,7	0,15	0,15	3	1000	-3217,85	-880,53	-2930,11	0,12	0,12	0,11	0,12	85,5	86,6	85,3	85,4
	0,7	0,15	0,15	4	1000	-3026,04	-687,02	-2646,22	0,13	0,13	0,13	0,13	85,3	84,6	84,7	88,8
	0,7	0,15	0,15	5	1000	-3252,54	-911,34	-2780,64	0,12	0,12	0,13	0,12	85,5	83,4	85,9	86
	0,8	0,1	0,1	2	1000	-3646,93	-976,97	-3446,71	0,13	0,12	0,13	0,15	85,5	86,5	85,8	82
	0,8	0,1	0,1	3	1000	-3825,92	-1154,89	-3531,48	0,11	0,11	0,10	0,11	85,5	87,7	85,1	86,6
	0,8	0,1	0,1	4	1000	-3650,10	-977,59	-3261,44	0,12	0,12	0,10	0,16	85,5	88,3	85,8	80,5
	0,8	0,1	0,1	5	1000	-3747,36	-1072,96	-3264,48	0,11	0,12	0,10	0,12	85,6	87,4	85,5	84,4
Lojistik Regresyon	0,9	0,1				-3065,16	-665,70	-2675,52	0,14	0,13	0,15		86,375	86,53	84,98	
	0,8	0,2				-2874,33	-624,26	-2508,95	0,14	0,14	0,14		85,78	85,64	86,34	
	0,7	0,3				-3001,30	-651,83	-2619,78	0,14	0,14	0,14		86,232	86,25	86,19	
	0,6	0,4				-3068,96	-666,53	-2678,84	0,14	0,13	0,14		86,378	86,55	86,12	

Tüm çalıştırılan yapay sinir ağları modelleri arasında en başarılı modelin %80-%10-%10 veri bölümlemesine ve 5 katmana sahip "trainr" algoritmasının olduğu görülmüştür. Trainr algoritması kullanılarak çalıştırılan modelde %93,5 doğruluk oranı elde edilmiştir. Bu sonuç lojistik regresyon modeli ile karşılaştırılacaktır.

3.5 Lojistik Regresyon Modelinin Uygulanması

Yukarıda bahsedildiği şekilde çalıştırılan lojistik regresyon modelinin sonuçları aşağıdadır:

Tablo 3.4 Lojistik Regresyon Model Çıktıları

	Veri bölümleme					**MSE**			**Doğruluk oranı**		
	eğitim	**test**	**AIC**	**AICc**	**BIC**	**eğitim**	**test**	**tüm**	**eğitim**	**test**	**tüm**
Lojistik Regresyon	0,9	0,1	-3065,16	-665,70	-2675,52	0,13	0,15	0,14	86,5	85	86,4
	0,8	0,2	-2874,33	-624,26	-2508,95	0,14	0,14	0,14	85,6	86,3	85,8
	0,7	0,3	-3001,30	-651,83	-2619,78	0,14	0,14	0,14	86,3	86,2	86,2
	0,6	0,4	-3068,96	-666,53	-2678,84	0,13	0,14	0,14	86,6	86,1	86,4

Çalıştırılan lojistik regresyon modelleri arasında en başarılı modelin %60 eğitim %40 veri bölümlemesi yapılarak elde edilen model olduğu görülmüştür. Bu modelin doğruluk oranı %86,5'tir.

3.6 En İyi Modele Ait Detaylar

Yapay sinir ağları ve lojistik regresyona ait çıktılara bakıldığında %93,5 başarı oranı ile "trainr" fonksiyonu ile çalışılan yapay sinir ağları modelinin başarılı model olduğu görülmüştür. Bu model %80 %10 %10 veri bölümlemesi ve 5 katmana sahiptir. Sonuç 67 iterasyonda oluşmuştur. En iyi modele ait grafikler gösterilmiştir.

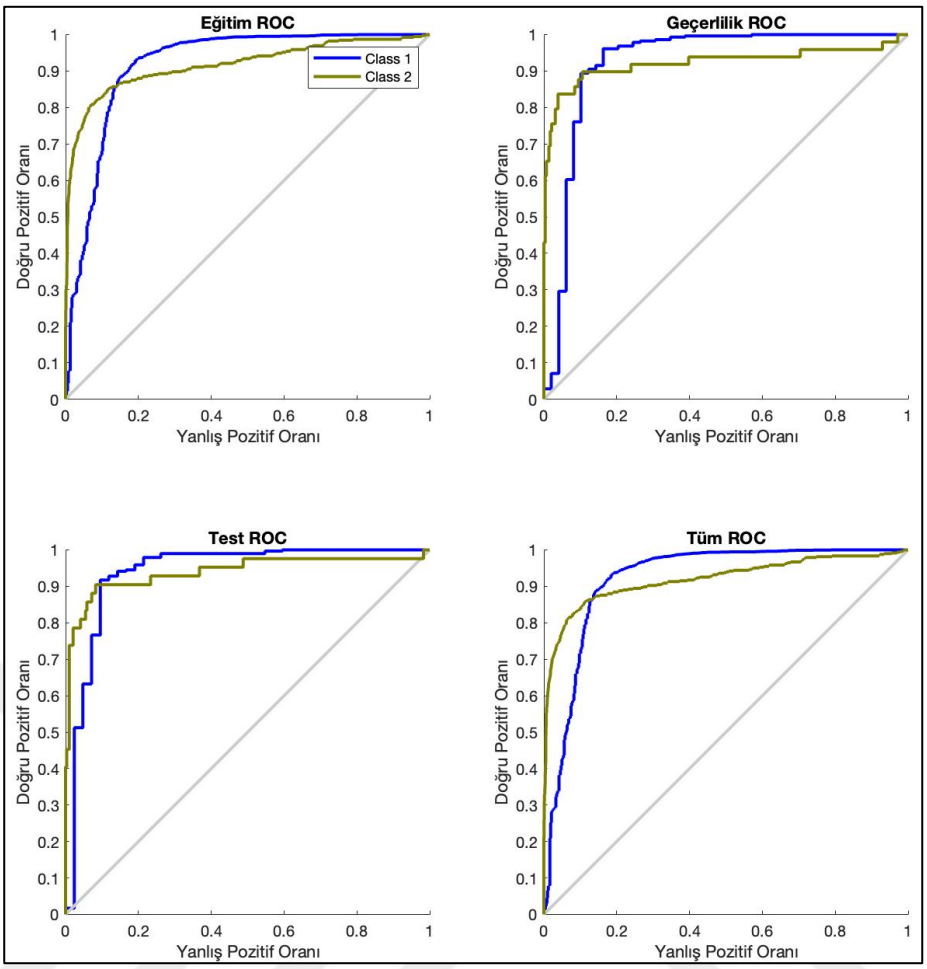
Hata matrisleri aşağıdaki gibidir:



Şekil 3.8 En İyi Modelin Hata Matrisleri

En iyi modele ait hata matrislerine baktığımızda doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif değerlerine ait dağılımları görebiliriz. Örneğin tüm hata matrisine baktığımızda 2803 müşteri için gerçek değer 1 iken tahmin değeri de 1 gerçekleşmiştir ve doğru tahmin edilmiştir. 47 müşteri için ise gerçek değeri 1 olduğu halde 2 tahmin edildiğini görüyoruz. Yani 47 müşteri aslında kayıp olmamasına rağmen bizim modelimiz o müşterileri kaybedilecek müşteri olarak tahminlemiştir. Tüm matrisleri incelediğimizde genel olarak doğru tahminleme yapıldığını görebiliriz.

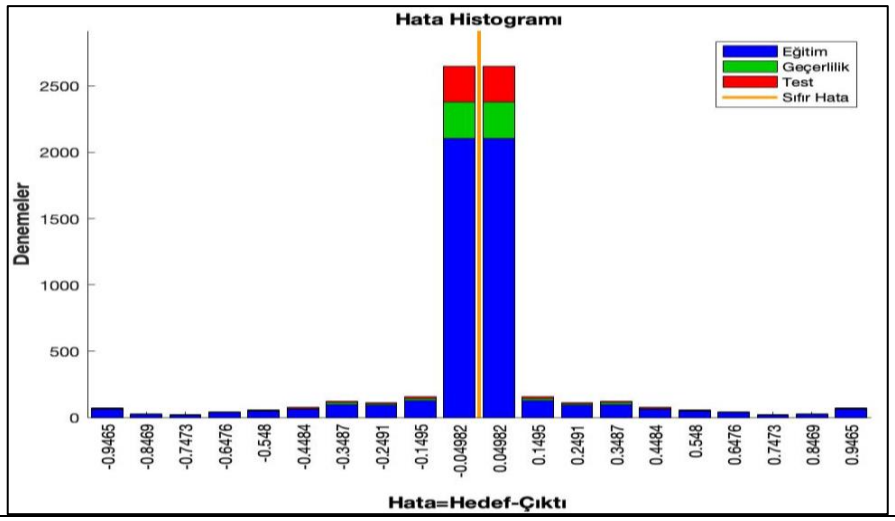
Modele ait roc eğrileri aşağıdaki gibidir:



Şekil 3.9 En İyi Modelin ROC Eğrileri

Tezin araştırma kısmında da bahsedildiği gibi ROC eğrisi tahminlemenin başarısını gösteren grafiklerdendir. Yukarıdaki grafiklerde lacivert eğrinin y grafiğinde 1'e yakın değerlerde ilerlemesi modelin başarısının yüksek olduğunu, doğru pozitif oranının yüksek olduğunu gösterir.

Hata histogramı aşağıda gösterilmiştir:



Şekil 3.10 En İyi Modelin Hata Histogramı

Hata histogramı sayesinde yapay sinir ağları modelleri çalıştırıldıktan sonra hedef değerler ile tahmin edilen değerlerin farkından oluşan hata değerlerini görselleştirebiliriz. Yukarıda en iyi modele ait hata histogramı gösterilmiştir. Burada hataların çok büyük kısmının -0,4982 ve +0,4982 arasında konumlandığını görüyoruz. Modelin başarısı yüksek olduğu için hata değerlerinin 0'a yakın değerlerde görülmesi beklenen bir durumdur.

4
SONUÇ VE ÖNERİLER

Yapılan tüm çalışmalar sonucunda elde ettiğimiz çıktılara baktığımızda yapay sinir ağlarının daha esnek yapısı nedeniyle lojistik regresyon modeline göre çok daha iyi başarı gösterdiği görülmüştür. Çalışmada ilk olarak veri seti üzerinden t-testi yaparak anlamsız değişkenler elenmiş, sonrasında ise sırasıyla yapay sinir ağları ve lojistik regresyon algoritmaları Matlab yardımı ile çalıştırılmıştır. Çalıştırılan tüm modeller sonucunda yapay sinir ağlarına ait Trainr algoritması kullanılarak elde edilen sonuçta %93,5 tahmin başarısı yakalanmıştır. Klasik yöntem olan lojistik regresyon %86,5 doğruluk oranı ile geride kalmıştır.

Müşteri kayıp analizinde literatürde farklı modeller kullanılmaktadır. Klasik yöntemler bu tarz çok değişkenlerin çözümünde başarı olarak yetersiz kalabilmektedir. Çalışmamızda da yapay sinir ağları algoritmaları, esnek yapısı nedeniyle lojistik regresyondan daha iyi sonuçlar vermiştir.

Yapılan çalışmada bir telekomünikasyon firmasının sadece yaptığı telefon görüşmeleri ve paketlerine ait verilerden faydalanılarak sonuçlara ulaşılmıştır ve müşterilerin telefon hareketlerine göre müşteriyi kaybedip kaybetmeyeceği tahminlenmeye çalışılmıştır. Günümüzde gelişen teknoloji ile telekomünikasyon şirketler yalnızca operatör olmanın çok daha dışında bir sürü farklı alanda hizmet vermektedirler. Dolayısıyla bu müşteri kaybı çalışmaları müşterinin kullandığı diğer hizmetlere ait verileri de eklenerek daha geniş çapta ilerletilebilir ve daha kesin sonuca ulaşabilirler. Ayrıca bu kapsamda ilerletecekleri müşteri kazanım kampanyalarını da verileri irdeleyerek belirleyebilirler ve uygun müşteriye uygun kampanyayı göstererek müşteriyi geri kazanabilirler. Örneğin Türkiye'de bugün telekomünikasyon şirketleri ayrıca çok büyük teknoloji şirketleri olarak da görev yapmaktadırlar ve bir sürü dijital ürüne sahiptirler. Bu dijital ürünlerde müşterilerin hareketlerini anlayarak kayıp analizi yapılabilir ve müşterilerin kullandıkları dijital ürünlere göre farklı kampanyalar önerilerek müşterinin başka firmaya geçmesini engelleyebilirler. Bu gibi çalışmalar ile sadece telekomünikasyon şirketleri değil , farklı sektörlerde hizmet veren büyük firmalar da aynı şekilde

kaybedecekleri müşterileri önceden anlayarak onları tutundurmak için çalışmalar yapabilir ve karlılıklarını bu kapsamda da artırmaya çalışabilirler.

KAYNAKÇA

Ağyar, Z. (2015). Yapay sinir ağlarının kullanım alanları ve bir uygulama. Mühendis ve Makine, 56(662), 22-23.

Ahmad, A.K., Jafar, A. & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. J Big Data 6, 28.

Arsan, T. ve Çimenli, S. (2016). Churn Analysis and Prediction. International Journal of Computer Science and Information Security 14(8), 550.

Baydemir M., (2014). Lojistik Regresyon Analizi Üzerine Bir İnceleme,

Bar-Hillel, M., & Neter, E. (1996). Why are people reluctant to exchange lottery tickets?. Journal of Personality and Social Psychology, 70(1), 17.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern recognition, 30(7), 1145-1159.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS inc, 9, 13.

Chen, K., Hu, Y.-H., & Hsieh, Y.-C. (2015). Predicting customer churn from valuable B2B customers in the logistics industry: A case study. Information Systems and e-Business Management, 13(3), 475-494.

Crone, S. F. (2004). Bussiness Forecasting with Artificial Neural Networks. Institute of Bussiness Forecasting.

Dahiya, K., & Bhatia, S. (2015). Customer churn analysis in the telecom industry. 4th International Conference on IEEE. India.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. 233-240.

Donner, A., & Klar, N. (1996). The statistical analysis of kappa statistics in multiple samples. Journal of Clinical Epidemiology.

Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R., & Basaran, M. A. (2009). A new approach based on artificial neural networks for high order multivariate fuzzy time series. Expert Systems with Applications, 36(7), 10589-10594.

Eltas, Ö . (2021). Biyoistatistik Çalışmalarında Kullanılan Küçük Örneklemlerde Mann-Whitney U Testi ve Bağımsız Örneklem T (Student's-t Independent Test) Testinin Güç Yönünden Karşılaştırılması . Atatürk Üniversitesi Veteriner Bilimleri Dergisi , 16 (1) , 88-94 .

Erilli, N. A., Eğrioğlu, E., Yolcu, U., Aladağ, Ç. H., & Uslu, V. R. (2010). Türkiye'de enflasyonun ileri ve geri beslemeli yapay sinir ağlarının melez yaklaşımı ile öngörüsü. Doğuş Üniversitesi Dergisi, 11(1), 42- 55.

Fletcher, D., & Gross, E. (1993). Forecasting with Neural Networks: An Application Using Bankrutcy Data. Information and Management, 24(3), 159-167.

Forhad, N., Hussain, M. S., & Rahman, R. M. (2014). Churn analysis: Predicting churners. 237-241.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge discovery in databases: An overview. AI magazine, 13(3), 57-57.
Gülpınar, V . (2015). Yapay Sinir Ağlari Ve Sosyal Ağ Analizi Yardimi İle Türk Telekomünikasyon Piyasasinda Müşteri Kaybi Analizi . Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi , 34 (1) , 331-350 .

Hamel, G. (2009.). Moon Shots for Management. 8.

Harrell Jr, F. E. (2015). Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer.

Hayden, R. W. (2008). A Review of:"Applied Regression Analysis and Other Multivariable Methods, by DG Kleinbaum, LL Kupper, A. Nizam, and KE Muller" , 906.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc Nat Acad Sci, 79, 2554–2558.

HUNG, Shin-Yuan, David C. Yen ve Hsiu-Yu Wang ,(2006). "Applying Data Mining to Telecom Churn Management", Expert Systems with Applications, Volume 31, Issue 3, 515-524.

Jadhav, R. J., & Pawar, U. T. (2011). Churn prediction in telecommunication using data mining technology. International Journal of Advanced Computer Science and Applications, 2(2).

Kabalcı, E. (2014). Yapay Sinir Ağları. Ders Notları

Kaur, I. & Kaur, J. (2020). "Customer Churn Analysis and Prediction in Banking Industry using Machine Learning," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), 434-437,

Kaynar, O., Taştan, S., & Demirkoparan, F. (2010). Ham petrol fiyatlarının yapay sinir ağları ile tahmini. Ege Akademik Bakış Dergisi, 10(2), 559-573.

Keskenler, M. F., & Keskenler, E. F. (2017). Bulanık mantığın tarihi gelişimi. Takvim-i Vekayi, 5(1), 1-10.

Komarek, P., & Moore, A. W. (2005). Making logistic regression a core data mining tool with tr-irls. In Fifth IEEE International Conference on Data Mining (ICDM'05)

Kubat, C. (2015). MATLAB yapay zeka ve mühendislik uygulamaları, 1. Baskı

Lewis, T. (2014). A Brief History of Artificial Intelligence. Live Science Retrieved.

Liu, P., Wu, N., Huang, C., Wei, B., Wang, L., & He, Z. (2006). Research on Applying Data Mining to Telecom CRM. International Forum of Information System Frontiers-2006 Xian International Symposium.

Lu, J. (2002). Predicting customer churn in the telecommunications industry––An application of survival analysis modeling using SAS, 114.

Ljung, L. (1999). System Identification: Theory for the User, Upper Saddle River, NJ, Prentice-Hall PTR,

McCulloch, W.S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5, 115–133.

Nath, S. V., & Behara, R. S. (2003). Customer churn analysis in the wireless industry: A data mining approach. 561, 505-510.

Öztemel, E. (2006). Yapay sinir ağları, 2. Baskı. İstanbul: Papatya Yayıncılık.

Reichheld, F. F., & Sasser, W. E. (1990). Zero defeofions: Quoliiy comes to services. Harvard business review, 68(5), 105-111.

Saraç, T. (2004). Yapay Sinir Ağları. Seminer Projesi. Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

Sund, R. (2002). Utilization of Administrative Registers using Statistical Knowledge Discovery. International Workshop on Mining Official Data, National Research and Development Centre for Welfare and Health, Finland.

Turanoglu-Bekar, E., Ulutagay, G., & Kantarcı-Savas, S. (2016). Classification of thyroid disease by using data mining models: A comparison of decision tree algorithms. Oxford Journal of Intelligent Decision and Data Sciences, 2, 13-28.

Van den Poel, D., & Lariviere, B. (2004). Customer attrition analysis for financial services using proportional hazard models. European journal of operational research, 157(1), 196-217.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79-82.

R. Zhang, W. Li, W. Tan ve T. Mo, (2017). "Deep and Shallow Model for Insurance Churn Prediction Service", 2017 IEEE International Conference on Services Computing (SCC), 346-353.

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirileri

1. Tüfekçi P.S, Büyüklü A.H, Parim C., (2021). Customer Churn Analysis With Artificial Neural Network Algorithms – An Example Of A Telecom Company. Middle East International Conference on Contemporary Scientific Studies-V Konferansı, Ankara, 496