

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**HÜCRESEL MOBİL İLETİŞİM SİSTEMLERİNDE DOLANDIRICILIK RİSK
YÖNETİMİ**

ONUR TÜFEKÇİOĞLU

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
PROF. DR. BANU DİRİ**

İSTANBUL, 2017

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**HÜCRESEL MOBİL İLETİŞİM SİSTEMLERİNDE DOLANDIRICILIK RİSK
YÖNETİMİ**

Onur TÜFEKÇİOĞLU tarafından hazırlanan tez çalışması 15.12.2017 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Prof. Dr. Banu DİRİ
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Prof. Dr. Banu DİRİ
Yıldız Teknik Üniversitesi

Yrd. Doç Dr. Oğuz ALTUN
Yıldız Teknik Üniversitesi

Yrd. Doç Dr. Tuğba YILDIZ
İstanbul Bilgi Üniversitesi

ÖNSÖZ

Son yıllarda gelişen yeni mobil ödeme yöntemleri, artan veri kullanımı ve yeni kurgulanan katma değerli servisler nedeniyle GSM operatörlerinin maruz kaldığı dolandırıcılık olayları ciddi manada artmıştır.

Bu çalışmada GSM operatörleri için risk oluşturan dolandırıcılık olaylarının tespiti için kullanılan risk yönetim sistemlerinin performansları incelenmiştir

Mevcut performansı arttırabilmek için 4 farklı Makine Öğrenmesi yöntemi (K-En Yakın Komşu, Naive Bayes, Rasgele Orman, Destek Vektör Makinesi) ile Derin Öğrenme (Deep Learning) yöntemleri ile X operatörüne ait sınıflandırılmış gerçek verilerdeki performans değerleri karşılaştırılmıştır.

Çalışma sırasında beni pozitif olarak motive eden ve bilgisi ile yardımlarını esirgemeyen tez danışmanım Prof. Dr. Banu Diri Hocama çok teşekkür ederim.

Derin Öğrenme metodları ile ilgili olarak verdiği dersine girmeme izin veren ve sorularıma sabırlı bir şekilde cevap veren Doç. Dr. Sırma Yavuz hocama teşekkür ederim.

Ayrıca Derin Öğrenme ilgili yaşadığım sıkıntılarda, günlerce beraber çalışarak aşmama yardımcı olan ve yol gösteren Araştırma Görevlisi Dr. Göksel Biricik'e nekadark teşekkür etsem azdır.

Son olarak her zaman bana destek olan ailem yanımda oldukları için kendimi çok şanslı hissediyorum ve hepsine sonsuz teşekkürlerimi sunuyorum.

Aralık, 2017

Onur TÜFEKÇİOĞLU

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	vi
KISALTMA LİSTESİ	vii
ŞEKİL LİSTESİ.....	viii
ÇİZELGE LİSTESİ	ix
ÖZET	x
ABSTRACT	xi
BÖLÜM1	
GİRİŞ	1
1.1 Literatür Özeti	2
1.2 Tezin Amacı	3
1.3 Hipotez	3
BÖLÜM2	
VERİ SETİ	4
2.1 Verinin Hazırlanması	4
2.2 Veri İşleme ve Özellik Çıkarılması	4
2.2.1 Çıkarılan Özellikler	5
2.3 Veri Seti	6
BÖLÜM 3	
İNCELENEN SİSTEMİN PERFORMANSI VE SİSTEMDEKİ ABONE KAYIPLARI	9
3.1 İncelenen Sistemin Performansı	9
3.2 Abone Kayıpları	10

BÖLÜM 4

GELİŞTİRİLEN SİSTEM	14
4.1 Kullanılan Yöntemler	15
4.1.1 K-En Yakın Komşuluk (k-NN) Yöntemi	15
4.1.2 Naive Bayes (NB) Yöntemi	18
4.1.3 Rastgele Orman (RO) Yöntemi.....	18
4.1.4 Destek Vektör Makinesi (DVM)	19
4.1.5 Weka Uygulaması	20
4.2 Makine Öğrenmesi Kullanılarak Alınan Sonuçlar	23
4.2.1 Aksiyon Tipine Göre.....	23
4.2.2 Kullanım Tipine Göre	25
4.3 Karşılaştırma	27

BÖLÜM 5

MEVCUT VERİ SETİNİN DERİN ÖĞRENME İLE İŞLENMESİ	30
5.1 Kullanılan Yöntemler	30
5.2 Alınan Sonuçlar	31
5.2.1 Aksiyon Tipine Göre Sonuçlar	32
5.2.2 Kullanım Tipine Göre Sonuçlar	33

BÖLÜM 6

PERFORMANS KARŞILAŞTIRMASI.....	34
6.1 Aksiyon Bazlı Performans Karşılaştırması	34
6.2 Kullanım Bazlı Performans Karşılaştırması.....	36

BÖLÜM 7

SONUÇ VE ÖNERİLER	39
KAYNAKLAR.....	40
EK-A	
ÖZELLİKLER.....	42
ÖZGEÇMİŞ.....	47

SİMGE LİSTESİ

$L2w$	L2Weightregularization - Ağırlık Düzenleme Katsayısı
$P(A)$	A olayının gerçekleşme olasılığı
$P(B)$	B olayının gerçekleşme olasılığı
$P(A\backslash B)$	B olayının gerçekleştiği durumda A olayının meydana gelme olasılığı
$P(B\backslash A)$	A olayının gerçekleştiği durumda B olayının meydana gelme olasılığı
Sr	SparsityProportion - Seyreklik Düzenleme Katsayısı
Sp	SparsityProportion - Seyreklik Oranı Katsayısı

KISALTMA LİSTESİ

Arff	Atribute relation file format
DVM	Destek Vektör Makinesi
GSM	Global System for Mobile Communication
k-NN	K-En Yakın Komşuluk
NB	Naive Bayes
PL/SQL	Procedural Language/Structured Query Language
RO	Rastgele Orman

ŞEKİL LİSTESİ

	Sayfa
Şekil 2. 1	Kullanım ve Kapama Bazlı Veri Seti..... 6
Şekil 2. 2	Kullanım tipine göre Sınıf bilgisi..... 7
Şekil 2. 3	Kapama tipine göre Sınıf bilgisi..... 6
Şekil 3. 1	Başarı Oranları..... 10
Şekil 3. 2	Kayıp yüzdesi..... 12
Şekil 3. 3	Kayıp Oranları..... 12
Şekil 3. 4	Tercih edilen Operatör bilgisi..... 13
Şekil 4. 1	2 boyutlu olarak sınıflandırılmış abone kümesi 16
Şekil 4. 2	Sınıfı bilinmeyen yeni örnek..... 16
Şekil 4. 3	En yakın 3 komşu 17
Şekil 4. 4	Sınıflandırılmış örnek 17
Şekil 4. 5	Rastgele Ağaç Orman Yapısı..... 19
Şekil 4. 6	Destek vektörlerince çizilen destek vektörleri, marjin ve hiperdüzlem 20
Şekil 4. 7	Weka Açılış Ekranı 20
Şekil 4. 8	Wekaya data yüklenmesi 21
Şekil 4. 9	Sınıflandırma işlemi ve Sonuç Bilgileri 22
Şekil 4. 10	Özellik seçme işlemi ve Sonuç Bilgileri..... 22
Şekil 4. 11	Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik) 24
Şekil 4. 12	Aksiyon Tipine Göre Başarı Oranları (%) (10 Özellik) 25
Şekil 4. 13	Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik) 26
Şekil 4. 14	Kullanım Tipine Göre Başarı Oranları (%) (10 Özellik) 27
Şekil 4. 15	76 Özellik için Performans Karşılaştırma 28
Şekil 4. 16	10 Özellik için Performans Karşılaştırma 29
Şekil 5. 1	AutoEncoder Mimari Modeli 31
Şekil 5. 2	Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik) 32
Şekil 5. 3	Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik) 33
Şekil 6. 1	Servis Kapama Aksiyonuna göre başarı oranları (%) (76 Özellik) 35
Şekil 6. 2	Tam Kapama Aksiyonuna göre başarı oranları (%) (76 Özellik) 35
Şekil 6. 3	Tüm Kapama Aksiyonları için başarı oranları (%) (76 Özellik) 36
Şekil 6. 4	Ses Kullanımı yapan aboneler için başarı oranları (%) (76 Özellik) 37
Şekil 6. 5	Mesaj Kullanımı yapan aboneler için başarı oranları (%) (76 Özellik) 37
Şekil 6. 6	Tüm Kullanımlar için başarı oranları (%) (76 Özellik) 38

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2. 1	Veri Seti 6
Çizelge 2. 2	Sınıflandırılmış Veri Seti 7
Çizelge 3. 1	Başarı Oranları (%)..... 10
Çizelge 3. 2	Abone Kayıpları 11
Çizelge 3. 3	Tercih edilen operatör bilgisi (%) 13
Çizelge 4. 1	Aksiyon Tipine Göre Başarı Oranları(76 Özellik) 23
Çizelge 4. 2	Aksiyon Tipine Göre Başarı Oranları (10 Özellik) 24
Çizelge 4. 3	Kullanım Tipine Göre Başarı Oranları(76 Özellik) 25
Çizelge 4. 4	Kullanım Tipine Göre Başarı Oranları(10 Özellik) 26
Çizelge 4. 5	Performans Karşılaştırma (%)..... 28
Çizelge 5. 1	Aksiyon Tipine Göre Başarı Oranları(76 Özellik) 32
Çizelge 5. 2	Kullanım Tipine Göre Başarı Oranları(76 Özellik) 33
Çizelge 6. 1	Aksiyon Tipine Göre Başarı Oranları (76 Özellik) 34
Çizelge 6. 2	Kullanım Tipine Göre Başarı Oranları(76 Özellik) 36

HÜCRESEL MOBİL İLETİŞİM SİSTEMLERİNDE DOLANDIRICILIK RİSK YÖNETİMİ

Onur TÜFEKÇİOĞLU

Bilgisayar Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Tez Danışmanı: Prof. Dr. Banu DİRİ

Bu çalışmada GSM operatörleri için risk oluşturan dolandırıcılık olaylarının tespiti için kullanılan risk yönetim sistemlerinin performansları incelenmiştir. Mevcut performansı arttırabilmek için 4 farklı makine öğrenmesi yöntemi (K-En Yakın Komşu, Naive Bayes, Rasgele Orman, Destek Vektör Makinesi) kullanılmıştır. X GSM operatörüne ait 2 aylık bir zaman diliminde dolandırıcılık şüphesiyle aksiyon alınan (Servis Kapama, Tam Kapama) 5.641 örnek için 76 farklı özellik çıkarılmış ve 4 farklı makine öğrenmesi yöntemi ile bu örnekler yeniden sınıflandırılmıştır. Aynı zamanda özellik azaltma yapılarak elde edilen 10 özellikli yeni veri seti üzerinde aynı makine öğrenmesi yöntemleri ile tekrar sınıflandırma yapılarak performans karşılaştırması yapılmıştır.

Son yıllarda özellikle ses, resim ve el yazısı tanımasında oldukça başarılı sonuçlar elde edilen Derin Öğrenmeye de çalışmada yer verilmiştir. Mevcut veri seti üzerinde Derin Öğrenme algoritmaları da test edilerek başarıları makine öğrenme yöntemleriyle karşılaştırılmıştır.

Anahtar Kelimeler: GSM dolandırıcılığı, Dolandırıcılık risk yönetimi, Makine öğrenmesi, K- En Yakın Komşu, Naive Bayes, Rasgele Orman, Destek Vektör Makinesi, Derin Öğrenme

**FRAUD RISK MANAGEMENT IN MOBILE CELLULAR COMMUNICATIONS
SYSTEMS**

Onur TÜFEKÇİOĞLU

Department of Computer Engineering

MSc. Thesis

Adviser: Prof. Dr. Banu DİRİ

Performance of the risk management system that is used for the detection of fraud on GSM Systems is investigated. In order to improve the current system performance, 4 different data mining methods is used. These methods are K-Nearest Neighbour, Naive Bayes, Random Forest, Support Vector Machine. Data Set is, 5641 GSM number belonging the X GSM operator. Suspected of fraud, X operator deactivated (service deactivated or full deactivated) these GSM numbers during 2 months. For 5641 GSM numbers, 76 feature is created and classified with 4 different machine learning method. Also using feature reduction, feature number reduces to 10 and classified with 4 different machine learning method. Performance of current system, developed system with 76 feature data set and reduced 10 feature data set is compared.

Speech, picture and hand writing recognition researches using deep learning methods are very popular and successful. Using the same data set Machine learning methods and deep learning methods performances are compared.

Keywords: Gsm fraud, Fraud risk management, Machine Learning, K-Nearest Neighbour, Naive Bayes, Random Forest, Support Vector Machine, Deep Learning

YILDIZ TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

GİRİŞ

GSM (Hücreli Mobil İletişim Sistemi) operatörlerinin temel amacı insanlar arasındaki iletişimi sağlamaktır. Gerek rekabet, gerekse gelişen teknoloji ve ihtiyaçlar, operatörlere müşterilerinin beklentilerini karşılamak ve yeni gelirler elde etmek için yeni hizmetler sunmayı zorunluluk haline getirmiştir. Numara taşınabilirliği ile operatörler arası rekabet çok ciddi manada artmıştır. Abone sayısındaki artış belli bir doyuma ulaştığı ve gelirlerde azalma olduğundan operatörler, hem gelirlerini hem de müşteri memnuniyetini arttırarak abone sayılarını koruyabilmek için katma değerli servislere önem vermeye başlamışlardır. Katma değerli servisleri ilk başta kendi bünyelerinde geliştiren operatörler çeşitliliği arttırmak ve riskleri azaltmak adına gelir paylaşım yöntemi ile farklı firmalarla işbirliğine gitmeye de başlamışlardır. Bu firmaların bir kısmı uluslararası hizmet veriyorken bir kısmı ise yerel firmalardan oluşmaktadır. Zamanla bu firmalar yerel hizmet veren şirketler olmaktan çıkıp, uluslararası hizmet veren firmalara dönüşmüşlerdir. Özellikle sundukları içerik açısından yasal olarak sorunlar yaşayan şirketler rahat hareket edebilecekleri ülkeler üzerinden hizmet vermeyi tercih edip, GSM operatörleri ile gelir paylaşımli anlaşmalar yapmışlardır. Zamanla artan kurgu sayısı, servis sayısındaki artış ve gelişen yeni ödeme kurguları (mobil ödeme, kredi kartı, temassız ödeme, cepte cüzdan, vs.) bu servislerin sunduğu hizmetlerin operatörler tarafından yeterince denetlenememesine ve ciddi dolandırıcılık vakalarına yol açmıştır [1-3].

Dünyada dolandırıcılık tespit etmek amacıyla çok sayıda ticari yazılım mevcuttur. Bu yazılımların performanslarını genel olarak kullanıcı tarafından yazılan kural setleri belirlemektedir. Mevcut incelenen X operatörünün sınıflandırma başarısı Bölüm 3’de

bildirildiği gibi %29,7'dir. Makine öğrenmesi yöntemleri kullanılmış ve sonuçları paylaşılmıştır [4][5-6].

İkinci bölümünde kullanılan veri seti, üçüncü bölümde de incelenen sistem hakkında bilgi verilmektedir. Bölüm dördte makine öğrenmesi yöntemleri ile geliştirilen sistem anlatılmış. Bölüm 5 te ise derin öğrenme yöntemleri ile mevcut veri setinin performansı incelenmiştir. Bölüm 6 da ise mevcut sistemin, makine öğrenmesi ve derin öğrenme yöntemleri performansları karşılaştırılmış ve sonuç bölümü ile devam etmiştir.

1.1 Literatür Özeti

Günümüzde GSM operatörleri dolandırıcılık vakalarını tespit edebilmek amacıyla çok çeşitli yazılımlar kullanmaktadır. Bu çalışmada incelenen X Operatörü, yurtdışı kökenli bir firmanın yazılımı ile kendi ihtiyaçlarına uygun geliştirdiği bir yazılımı aynı anda kullanarak ihtiyaçlarını karşılamaktadır. Finans ekibi altında çalışan bir ekibin belirlediği statik kural setlerine göre çalışan sistem, otomatik ya da manuel olarak aksiyon alma yeteneğine sahip. Aynı zamanda şüpheli hareketler incelenmek üzere ticket olarak ilgili ekiplerin kullandığı arayüzlerde oluşmaktadır. İhtiyaç duyulan kuralları yazan ve gerektiğinde güncelleyen ayrı bir ekip vardır. X operatörünün kendi bünyesinde yazdığı uygulamada ve yurtdışı firmaya yaptırılacak değişiklikleri yöneten teknik analiz ekibi ile yazılım ihtiyacını karşılamak amacıyla kod geliştiren bir ekip de mevcuttur. Uygulama yöneticileri, Database yöneticileri ve İşletim sistemi yöneticileri de tüm altyapının ve uygulamanın sağlıklı bir şekilde çalışmasından sorumludurlar.

Literatürde, Kaplan [10] Telekom sistemlerinde Fraud(Sahtekârlık) isimli çalışmasında muhtemel dolandırıcılık yöntemlerini genel olarak incelemiş ve her bir dolandırıcılık yöntemine ilişkin ayrı analiz yapılması gerektiğini vurgulamıştır. Kullanılacak yazılım ve raporlamalarla dolandırıcılık vakalarının önceden tespit edilebilmesinin mümkün olduğunu vurgulamıştır. Belçikada 1999 yılında yapılan bir sempozyumda "A hybrid system for fraud detection in mobile communication"[11] isimli çalışmada olası dolandırıcılık senaryoları ve dolandırıcılık tespit etme araçları incelenmiştir. Genel olarak kullanıcı profillemesi sistemi, öğrenmeli sinir ağı sistemi, kural bazlı tespit sistemi, öğrenmesiz sinir ağı sistemi, B number analiz sistemi incelenmiş ve bu sistemlerin

bütünleşik olarak kullanılması durumunda elde edilen kazançlar ortaya konmuştur. Fraud Detection in Telecommunications Industry: “Bridging the Gap with Random Rough Subspace Based Neural Network Ensemble Method” [12] isimli çalışmada ise dolandırıcılık tespitinde kullanılan yöntemler incelenmiştir. Mevcut veri setini alt parçacıklara ayırarak sinir ağlarının eğitilmesinde kullanılmasının performans etkileri incelenmiştir.

1.2 Tezin Amacı

GSM operatörleri son yıllarda dolandırıcılık olaylarının tespiti amacıyla çok ciddi yatırımlar yapmaktadır. Akıllı telefonların yaygınlaşması ve 4.5G hızlı internet bağlantısının kullanılmaya başlamasından sonra katma değerli servislerin sayıları ve kullanımları artmıştır. Tüm bu gelişmelere istinaden dolandırıcılık yöntemleri ve sayıları da hızla artmıştır. Operatörler bu dolandırıcılık olaylarını tespit etmek ve önlem almak için çok ciddi kaynaklar ayırmalarına rağmen doğru kitlenin tespiti noktasında ciddi eksiklikler vardır. Daha doğru bir kitleye aksiyon alabilmek amacıyla X operatörüne ait dolandırıcılık şüphesiyle aksiyon alınan 5.641 abone incelenmiş ve makine öğrenme yöntemleri (K-En Yakın Komşu, Naive Bayes, Rasgele Orman, Destek Vektör Makinesi) ve derin öğrenme yöntemleri ile sınıflandırılarak performans karşılaştırması yapılmıştır. Bu tezin amacı, X operatörüne ait sınıflandırılmış veri setinin başarı yüzdesini arttırabilmek için en uygun yöntemi tespit etmektir.[7-9]

1.3 Hipotez

Bu çalışmada X operatörüne ait dolandırıcılık şüphesiyle aksiyon alınan 5.641 abonenin makine öğrenme ve derin öğrenme yöntemleri ile sınıflandırılarak mevcut başarının üzerinde bir başarı elde edilmesi amaçlanmıştır.

İkinci bölümde mevcut veri seti anlatılmıştır, üçüncü bölümde mevcut sistemin performansı incelenmiştir. Dördüncü bölümde makine öğrenmesi yöntemleri ile geliştirilen sistem anlatılmıştır. Beşinci bölümde derin öğrenme yöntemi ile mevcut veri setinin işlenmesi anlatılmıştır. Altıncı bölümde makine öğrenmesi ve derin öğrenme yöntemi ile mevcut sistemin performansı karşılaştırılmıştır. Yedinci bölümde ise elde edilen sonuçlar paylaşılmıştır.

BÖLÜM 2

VERİ SETİ

X operatörünün 1 Ağustos 2014 - 1 Ekim 2014 tarihleri arasında dolandırıcılık yapma şüphesiyle ilgili hattın servise kapatıldığı (Ses veya mesaj kullanımı) ya da tam (tüm kullanımlar için: Ses, mesaj, internet) kapatılan 5.641 abonesi incelenmiştir.

2.1 Verinin Hazırlanması

Öncelikle verinin depolanması ve işlenmesi için Oracle 11G Express edition kurulmuştur. Oracle database'ine erişim ve PL/SQL dili kullanılarak özellik çıkarmak, verinin düzenlenmesi, verilerin analiz edilebilmesi, verinin dışa alınması için ticari bir yazılım olan Toad 12.0 uygulaması kurulmuştur. Hazırlanan veri setinin makine öğrenme yöntemleri ile işlenebilmesi için Weka 3.6 uygulaması, derin öğrenme yöntemleri ile işlenebilmesi için de Matlab R2017a uygulaması kurulmuştur.

2.2 Veri İşleme ve Özellik Çıkarılması

Öncelikle X Operatörüne ait 5.641 abonenin bilgi güvenliğine konu olan bilgileri silinmiştir. Daha sonra 15 farklı tabloda bulunan kullanım ve abonelik özelliklerinin yer aldığı veri kümesinden özellik çıkarma işlemi yapılmıştır. Veri Seti 800'den fazla sütun ve 5 milyona yakın satırdan oluşmaktadır. Bu veri setindeki hazır bilgilerden dolandırıcılık ile ilgili olabilecekler tespit edilmiş ve bunlara ilave olarak kullanım alışkanlıklarına dayalı yeni özellikler oluşturulmuştur. Daha sonra bu özellik kümesi ile ilgili varsa veri kalitesi problemi düzeltilmiştir. X operatörünün aldığı aksiyon bilgileri ile tespit edilen özelliklerin birleştirildiği ABONE isimli bir tablo oluşturularak tespit edilen

özellikler ilgili tablolardan alınmış ya da hesaplanarak ABONE tablosuna ilave edilmiştir. Bundan sonraki tüm işlemler ABONE tablosu üzerinden gerçekleştirilmiştir.

2.2.1 Çıkarılan Özellikler

Abonelere ait mevcut özellikler ve yeni oluşturulan özelliklerin bir kısmı açıklamaları ile aşağıda verilmiş olup, tümü EKA'de verilmektedir.

FRAUD_DURUMU: Abonenin gerçekte dolandırıcı olup olmadığını belirten ve sınıflandırmak amacıyla kullanılan özelliğidir. ADAY ve FRAUD olmak üzere 2 farklı sınıf mevcuttur.

ABONELIK_YASI: Abonenin hattını açtığı tarihten itibaren kaç gün geçtiğini gösterir.

KAPAMA_ONCESI_ODENEN_FAT_AD: Kapama öncesi abonenin kaç adet fatura ödediği bilgisidir.

KAPAMA_ONCESI_ODENEN_TUTAR: Kapama öncesi abonenin ödediği toplam fatura tutarı bilgisini gösterir.

GOVERNMENTFLAG: Hattın devlete ait olup olmadığını gösterir.

INDIVIDUALVIPSTATUS: Abonenin toplumda önemli bir etkisi olup olmadığını gösterir.

MVPNFLAG: Abonenin yüksek ödeme gücüne sahip olup olmadığını gösterir.

PISUBMARKETINGSEGMENT: Abonenin sahip olduğu alt pazarlama segmentini gösterir.

SIMCARDTYPE: Abonenin kullandığı simcard'ın tipini gösterir.

SUBSCRIBERTYPE: Faturalı ya da ön ödemeli abone gibi abonelik tipini gösterir.

TARIFF: Abonenin tarifesini gösterir.

GNCTRK: Abonenin genç tarifesinde olup olmadığını gösterir.

DUNNINGSEGMENT: Abonenin borç ödeme notunu gösterir.

DUNNINGFLAG: Abonenin ödenmemiş borcu olduğu için takipte olup olmadığı bilgisini gösterir.

CLUB1: X tarifesinde olup olmadığını bilgisini gösterir.

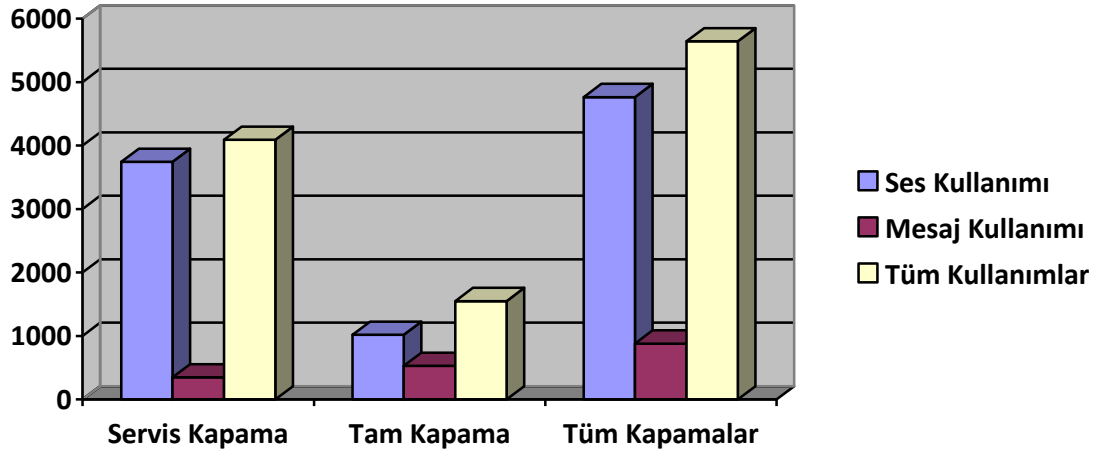
BILLINGCYCLE: Faturalama zamanını gösterir.

2.3 Veri Seti

X GSM operatörünün 1 Ağustos 2014 – 1 Ekim 2014 arasında dolandırıcılık yapma şüphesiyle ilgili hattın servise kapatıldığı (Ses veya mesaj kullanımı) ya da tam (tüm kullanımlar için: ses, mesaj, internet) kapatılan 5.641 örnek incelenmiştir. Aksiyon alınan abonelerin ses ve mesaj olmak üzere 2 farklı kullanım tipi mevcuttur. Veri setinin detayları Çizelge 2.1 ve Şekil 2.1 de gösterilmiştir.

Çizelge 2. 1 Veri Seti

	Servis Kapama	Tam Kapama	Tüm Kapamalar
Ses Kullanımı	3743	1019	4762
Mesaj Kullanımı	348	531	879
Tüm Kullanımlar	4091	1550	5641



Şekil 2. 1 Kullanım ve Kapama Bazlı Veri Seti

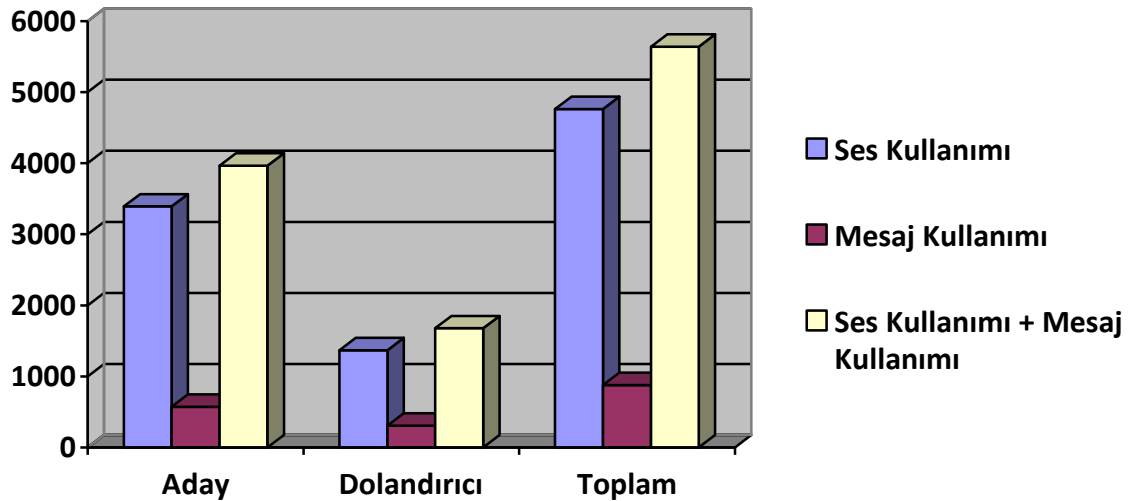
Eğer kapama aksiyonu alınan abone kendisinden istenen ara ödeme miktarını ödeyip hattını 1 ay içerisinde tekrar kullanıma açtırdı ise bu abone **aday** olarak sınıflandırılmıştır. Kendisinden istenen ödeme tutarını ödemeyen aboneler de **dolandırıcı** olarak sınıflandırılmıştır.

Kullanım ve aksiyon tiplerinin aday ve dolandırıcı olarak dağılımı Çizelge 2.2, Şekil 2.2 ve Şekil 2.3 te gösterilmiştir.

Çizelge 2. 2 Sınıflandırılmış Veri Seti

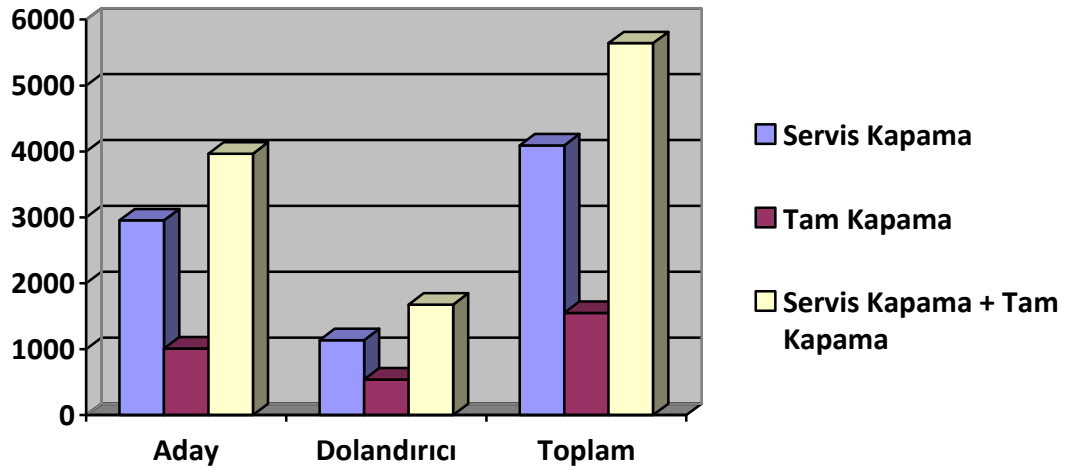
	Aday	Dolandırıcı	Aday+Dolandırıcı
Ses Kullanımı	3392	1370	4762
Mesaj Kullanımı	572	307	879
Ses Kullanımı + Mesaj Kullanımı	3964	1677	5641
Servis Kapama	2954	1137	4091
Tam Kapama	1010	540	1550
Servis Kapama + Tam Kapama	3964	1677	5641

Çizelge 2.2 Sınıflandırılmış veri setinin kullanım tipine ve aksiyon tipine göre dağılımını göstermektedir.



Şekil 2. 2 Kullanım tipine göre Sınıf bilgisi

Şekil 2.2 Sınıflandırılmış veri setinin kullanım tipine göre dağılımını göstermektedir.



Şekil 2. 3 Kapama tipine göre Sınıf bilgisi

Şekil 2.3 Sınıflandırılmış veri setinin aksiyon tipine göre dağılımını göstermektedir.

İNCELENEN SİSTEMİN PERFORMANSI VE SİSTEMDEKİ ABONE KAYIPLARI

Bu bölümde X operatörünün kullandığı dolandırıcılık sistemi ile mevcut veri seti üzerinde elde ettiği başarı incelenmiştir. Abone kayıpları ile hatalı kapatma arasında ilişki olup olmadığı ve bunun operatöre maliyeti incelenmiştir.

3.1 İncelenen Sistemin Performansı

X GSM operatörünün kullandığı sistemin başarısını ölçmek için denklem (3.1) kullanılmıştır.

Dolandırıcı Sayısı: sistem tarafından kapatılan ve 1 ay içerisinde açılmayan abonelerin sayısıdır.

Aday Sayısı: sistem tarafından kapatılan ve 1 ay içerisinde kendisinden istenen araödeme faturasını ödeyerek açılan abonelerin sayısıdır.

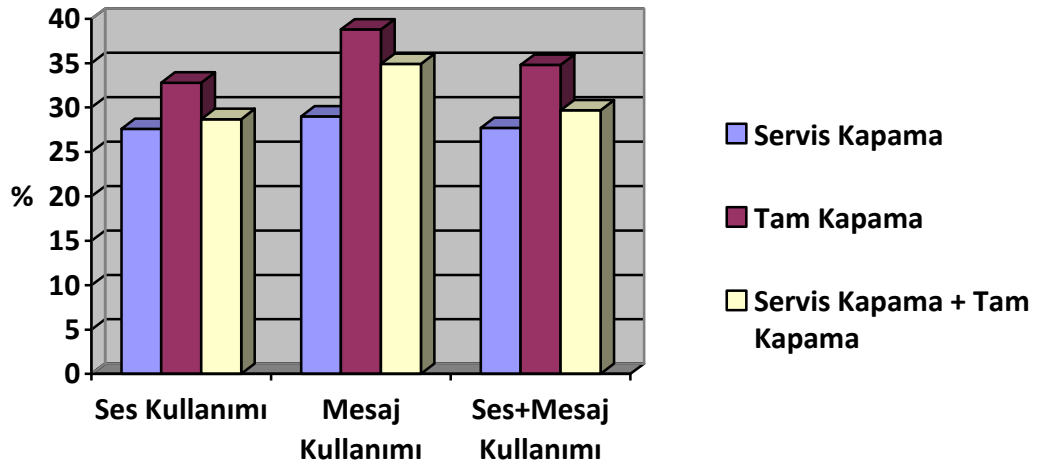
Başarı Oranı: Doğru bir şekilde tespit edilen dolandırıcıların yüzdesidir.

$$\text{Başarı Oranı} = (\text{Dolandırıcı Sayısı} / (\text{Dolandırıcı Sayısı} + \text{Aday Sayısı})) \times 100 \quad (3.1)$$

Ses Kullanımları, Mesaj Kullanımları ve her 2 kullanım türünün tamamı için, Servis Kapama, Tam Kapama ya da her 2 kapama türünün tamamı için hesaplanan başarı oranları Çizelge 3.1 de verilmiştir.

Çizelge 3. 3 Başarı Oranları (%)

	Servis Kapama	Tam Kapama	Tüm Kapamalar
Ses Kullanımı	27,6	32,8	28,7
Mesaj Kullanımı	29	38,8	34,9
Tüm Kullanımlar	27,7	34,8	29,7



Şekil 3. 1 Başarı Oranları

X operatörüne ait sistemin ortalama başarısı %29,7’dir. Mesaj kullanımında yapılan tam kapama başarısı %38,8 olmasına rağmen, Mesaj kullanımından dolayı tam kapama yapılan abonelerin sayısının düşük olması nedeniyle genel ortalamaya etkisi düşük olmuştur. Yaklaşık %30’luk bir başarı oranı oldukça düşük bir başarı oranıdır.

3.2 Abone Kayıpları

GSM operatörleri arasında numara taşınabilirliği çıktıktan sonra aboneler daha iyi hizmet alma, avantajlı kampanyalardan yararlanma, yüksek fatura ödememe, daha iyi çekim gücü ve operatörden duyulan memnuniyetsizlik gibi sebeplerden dolayı numaralarını diğer operatörlere taşımaya başlamıştır. Özellikle dolandırıcılık şüphesiyle

aksiyon alınan abonelerin, operatör memnuniyetlerinin düşmesi ve dolayısıyla hattını aksiyon alınmayan abonelerden daha fazla diğer operatörlere taşıması beklenen bir sonuçtur. Bu yüzden sistem tarafından yanlış sınıflandırma sonucu alınan her aksiyonun operatör genelinde yarattığı abone ve gelir kaybını ölçümleyebilmek GSM operatörleri için çok önemlidir.

Yanlış aksiyon alınması sebebiyle başka bir GSM operatörüne geçen abonelerin oranını karşılaştırabilmek amacıyla, daha önce hiçbir aksiyon alınmamış aboneler arasından rastgele seçilmiş 38.988 adet abone kontrol grubu olarak seçilmiş ve **normal** olarak etiketlenmiştir. Normal olarak etiketlenilmiş abone kitlesinde son 1 ayda başka bir GSM operatörüne geçenlerin sayısı *kayıp sayısı* olarak hesaplanmıştır. Aksiyon alınan abone kitlesinde ise aksiyon alındıktan 1 ay içerisinde abone başka bir operatöre geçtiyse bu abone kayıp olarak sınıflandırılmıştır. Kaybedilen abone yüzdesini hesaplamak için denklem (3.2) kullanılmıştır.

Kayıp Sayısı: Seçilen abone kitlesi içerisinde son 1 ay içerisinde başka bir GSM operatörüne geçenlerin sayısı.

Toplam Abone Sayısı: Seçilen abone kitlesindeki toplam sayıdır.

Kayıp Oranı: Toplam kitle içerisinde başka bir GSM operatörüne geçenlerin yüzdesi.

$$\text{Kayıp Oranı} = (\text{Kayıp Sayısı} / \text{Toplam Abone Sayısı}) \times 100 \quad (3.2)$$

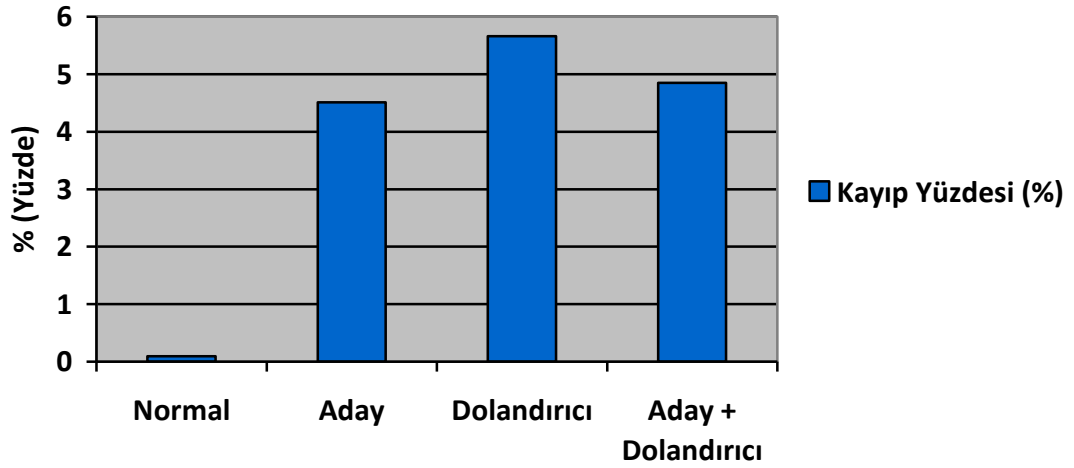
Çizelge 3. 2 Abone Kayıpları

	Toplam Abone	Kayıp Sayısı	Kayıp yüzdesi	Katı
Normal	38.988	36	% 0,0923	1
Aday	3.964	179	% 4,51	49
Dolandırıcı	1.677	95	% 5,66	61
Aday + Dolandırıcı	5641	274	% 4,85	52

Çizelge 3.2 de aksiyon alınan abonelerden aday olarak sınıflandırılanlarının operatör değiştirme oranının hiç aksiyon alınmayan normal abonelerin yaklaşık **49** katıdır. Bu

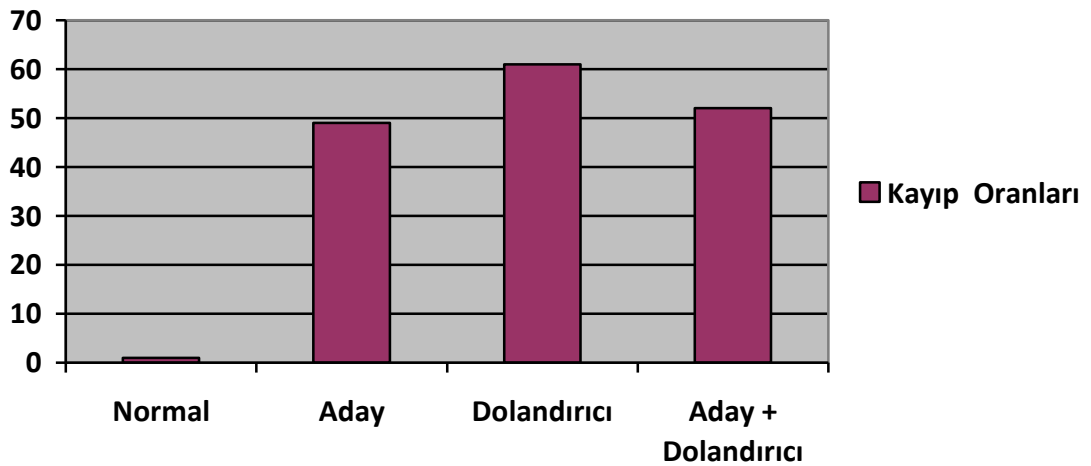
rakamlar hatalı aksiyon alınan abonelerin operatörler için çok ciddi bir problem olduğunu göstermektedir.

Şekil 3.2 de yüzdesel olarak kayıp oranları gösterilmiştir.



Şekil 3. 2 Kayıp yüzdesi

En düşük yüzde olan normal abonelerin kayıp yüzdesi 1 birim kabul edilerek diğer kayıp oranları bu yüzdeye göre oranlanmıştır. Şekil 3.3' te Kayıp oranları karşılaştırmalı olarak gösterilmiştir.

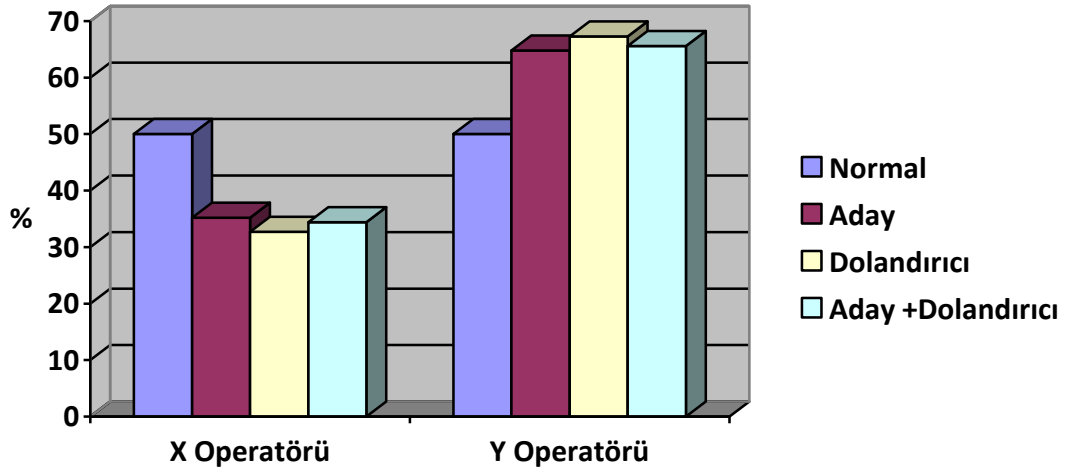


Şekil 3. 3 Kayıp Oranları

Mevcut tabloyu daha iyi analiz edebilmek adına diğer operatörlere geçen abonelerin tercih ettiği operatör bilgileri yüzdesel olarak Çizelge 3.3'te verilmiştir.

Çizelge 3. 3 Tercih edilen operatör bilgisi (%)

	Y Operatörü	X Operatörü
Normal	% 50	% 50
Aday	% 35	% 64,8
Dolandırıcı	% 32,7	% 67,3
Aday + Dolandırıcı	% 34,4	% 65,6



Şekil 3. 4 Tercih edilen Operatör bilgisi

Normal abonelerde değiştirilen operatör oranları eşit iken, aday ve dolandırıcı olarak sınıflandırılan abonelerin Z operatörünü Y operatörüne göre 2 kat daha fazla tercih ettiği görülmüştür. Bu oranlar Z operatörünün aday ve dolandırıcı olarak sınıflandırılan abonelerin beklentilerini, Y operatöründen daha iyi karşıladığını göstermektedir.

BÖLÜM 4

GELİŞTİRİLEN SİSTEM

X operatörünün kullandığı mevcut sistemin, normal kullanımlar dışında kullanım yapan aboneleri tespit etme ve doğru aksiyon alma başarısı %30 dur. Bu oran oldukça düşük olup, %70'lik bir abone kitlesine hatalı bir şekilde aksiyon alınmasına neden olmakta ve bu durum da çok ciddi müşteri memnuniyetsizliğine neden olmaktadır. Üçüncü bölümde incelendiği üzere aksiyon alınan abonelerin operatör değiştirme oranı, normal abonelere oranla 49 kat daha fazladır. Daha doğru bir ölçüm gerçekleştirebilmek adına aksiyon alınan tarihten itibaren 1 ay içerisinde başka bir operatöre geçip geçmediği incelenmiştir. Operatör değiştirme oranındaki 49 kat fazlalık net bir şekilde hatalı bir şekilde alınan her aksiyonun operatörlerlere ciddi bir müşteri kaybı ve memnuniyetsizliği olarak geri döndüğüdür. Sosyal medya'nın yaygınlaşması sonucu memnuniyetsizlik yaşayan abonelerin bu tecrübelerini başka kullanıcılarla da paylaştığı düşünülrse oluşan olumsuz etkinin gerçekte çok daha fazla olabileceği öngörülebilir.

Öncelikle verinin depolanması ve işlenmesi için Oracle 11G Express edition kurulmuştur. Oracle database'ine erişim ve PL/SQL dili kullanılarak özellik çıkarmak, verinin düzenlenmesi, analiz edilebilmesi, verinin dışa aktarılabilmesi için Toad 12.0 uygulaması kurulmuştur. Hazırlanan veri setinin makine öğrenme yöntemleri ile işlenebilmesi için Weka 3.6 uygulaması, derin öğrenme yöntemleri ile işleme yapılabilmesi için de Matlab R2017a uygulaması kurulmuştur.

4.1 Kullanılan Yöntemler

X operatörüne ait veri setindeki dolandırıcılık yapan aboneleri tespit etmek amacıyla 4 farklı makine öğrenmesi yöntemi kullanılmıştır

Bunlar k-En Yakın komşuluk (k-NN), Naive Bayes (NB), Rasgele Orman (RO) ve Destek Vektör Makinesi(DVM) dir.

Sistem modellenirken tespit edilen 76 farklı özellik kullanılmıştır. Sistemin başarı yüzdesini arttırmak ve daha hızlı çalışmasını sağlamak amacıyla özellik seçimine gidilerek özellik sayısı 10'a düşürülmüştür. Özellik seçimi ve abonelerin etiketlendirilmesinde açık kaynak kodlu Weka uygulaması kullanılmıştır. Özellik azaltma ve modellenen sistemin başarısını ölçümmlerken 10 katlı çapraz geçerleme (CV: Cross Validation k=10) kullanılmıştır. Özellik seçimi için de attribute evaluator: CfsSubsetEval, search method: BestFirst-D1-N5 kullanılmıştır.

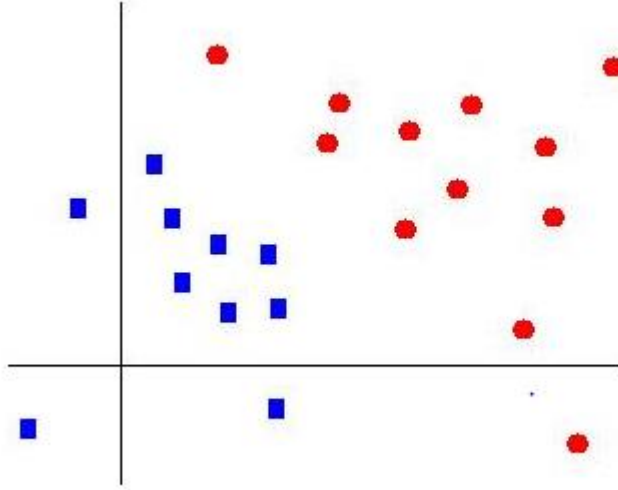
4.1.1 K-En Yakın Komşuluk (k-NN) Yöntemi

K-En Yakın Komşuluk algoritması N boyutlu yeni gelen bir örneğin sınıf etiketi bilinen örneklerle olan uzaklığını ölçerek, gelen örneğin hangi sınıfa ait olduğuna karar verir. Kendisine en yakın kaç adet örneğe bakarak karar vereceğini daha önceden belirlenmiş olan k parametresi belirler. Yakınlık ölçme yöntemi olarak genelde öklit distance kullanılmakla beraber farklı uzaklık ölçüm yöntemleri de kullanılabilir. K parametresi 1 olarak seçildiğinde gelen örnek kendisine en yakın olan sınıflandırılmış verinin sınıfı ile aynı sınıf olarak kabul edilir. K parametresi 3 olarak seçilirse, algoritma sınıflandırılacak örneğe en yakın 3 verinin sınıf etiketlerinden en fazla olan sınıfı seçer. K sayısı eşitlik olma durumu da düşünülerek genelde tek sayı seçilir. K parametresinin çok yüksek seçilmesi genelde başarı oranını düşürür. Çalışma mantığından dolayı sınıflandırılacak örnek için tüm veri setine olan uzaklık hesaplamasının yapılması gereklidir ve bu yüzden çalışma zamanı ve hesaplama süresi uzun olabilir.

N=2 ve K=3 olduğu bir durumu görsel olarak göstererek algoritmanın çalışma mantığını inceleyelim.

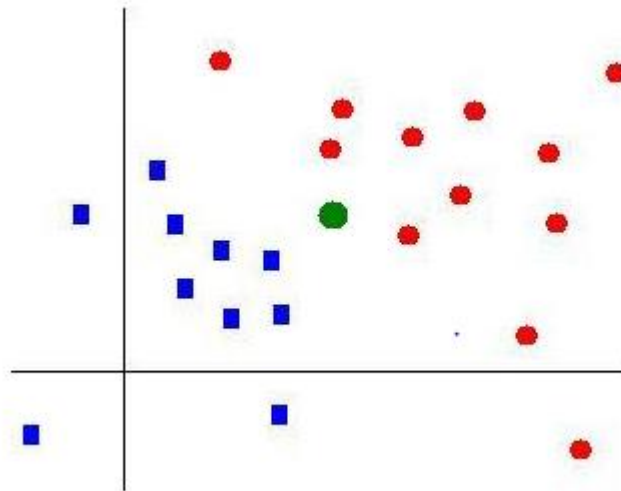
N=2 olması demek 2 boyutlu uzay için hesaplama yapılacak demektir. Başka bir ifadeyle sınıflandırılacak örnek sadece 2 tane özelliğe sahip demektir. Örneğin yaş ve

kilo gibi 2 özellik var demektir. $K=3$ olması gelen örneğe en yakın 3 noktayı tespit edip, bu 3 noktanın sınıfına bakarak en fazla olan sınıfı tespit ederek gelen örneği sınıflandıracak demektir.[13]

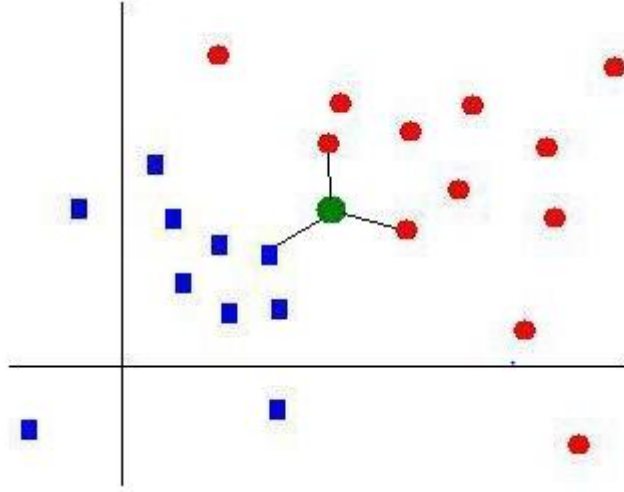


Şekil 4.1 2 Boyutlu olarak sınıflandırılmış Abone Kümesi

Şekil 4.1’de 2 boyutlu koordinat sistemine yerleştirilmiş örneklerin arasına sınıfını bilmediğimiz yeni bir örnek geldiğinde Yeşil ile gösterilen örneğin hangi sınıfa ait olduğunu tespit etmek için $k=3$ olacak şekilde en yakın 3 komşusu öklit mesafe bulma yöntemi ile tespit edilir.

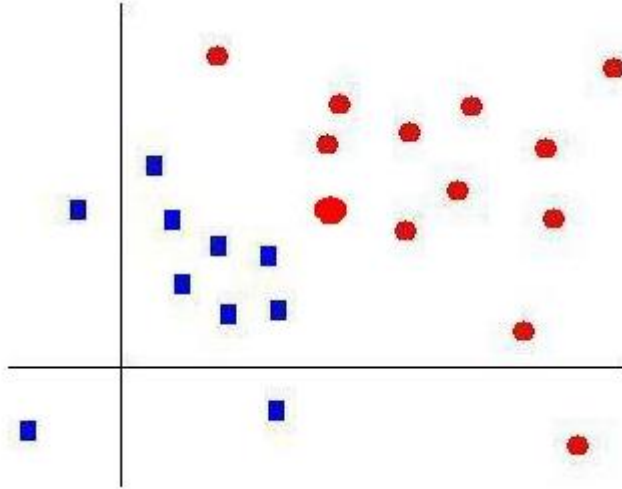


Şekil 4.2 Sınıfı Bilinmeyen Yeni Örnek



Şekil 4.3 En yakın 3 Komşu Nokta

Tespit edilen en yakın 3 noktanın, 2 tanesi kırmızı yuvarlak ile işaretlenmiş sınıfa aittir. Bir tanesi de mavi kare ile işaretlenmiş sınıfa aittir.



Şekil 4.4 Sınıflandırılmış Örnek

En yakın 3 noktanın 2 tanesi kırmızı yuvarlak olduğu için örnek kırmızı yuvarlak olarak sınıflandırılır.

4.1.2 Naive Bayes (NB) Yöntemi

Naive Bayes, genellikle metin sınıflandırma işlemlerinde kullanılan hızlı bir yöntemdir. Temelde olasılık hesabına dayanır. Her özellik eşit öneme sahiptir ve birbirleri hakkında bilgi barındırmazlar.

NB, Bayes teoremine dayanır. Bu teorem bir rassal değişken için koşullu olasılıklar ile önsel (marjinal) olasılıklar arasındaki ilişkiyi gösterir. Denklem (4.1)

$$P(A \setminus B) = \frac{P(B \setminus A)P(A)}{P(B)} \quad (4.1)$$

$P(A \setminus B)$: B olayının gerçekleştiği durumda A olayının meydana gelme olasılığıdır.

$P(B \setminus A)$: A olayının gerçekleştiği durumda B olayının meydana gelme olasılığıdır.

$P(A)$ ve $P(B)$: A ve B olaylarının gerçekleşme olasılığıdır.

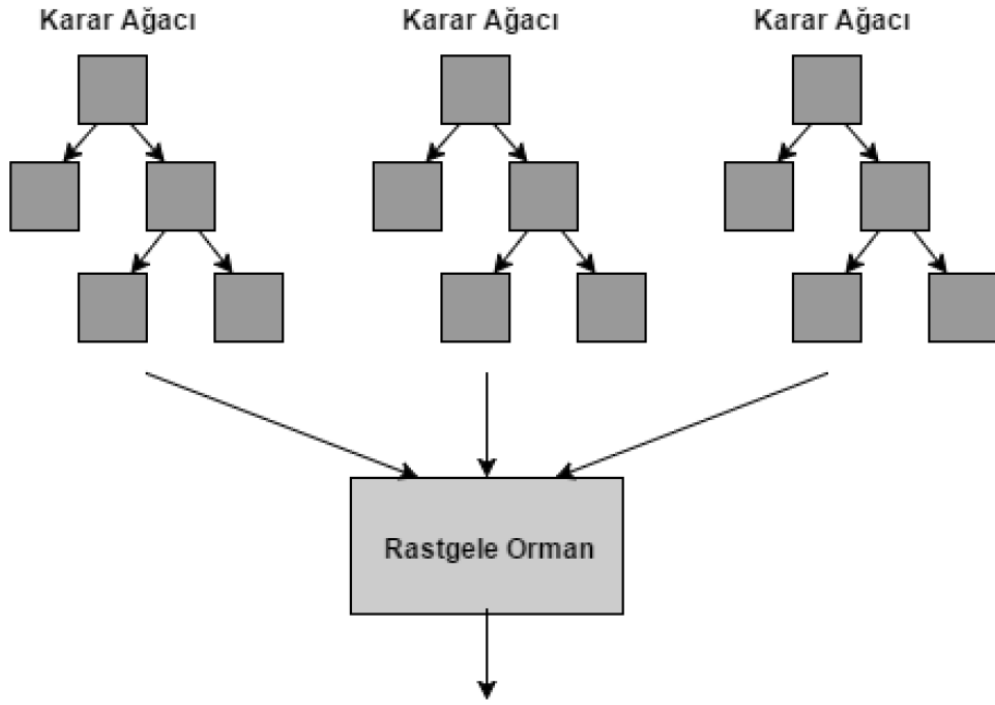
Bir sınıflandırma problemi birçok özellikten ve bir sonuç (hedef) değişkeninden oluşur. Naive Bayes Sınıflandırma Modeli Denklem (4.2) de gösterilmiştir.

$$p(C \setminus F_1, \dots, F_n) = \frac{P(C)p(F_1, \dots, F_n \setminus C)}{p(F_1, \dots, F_n)} \quad (4.2)$$

C, verilen hedefi F de özellikleri temsil eder. Naive Bayes sınıflandırıcı basitçe bütün koşullu olasılıkların çarpımı olarak düşünülebilir. Her 2 sınıf için aynı işlem yapılır ve sonuç olarak en yüksek olasılığa sahip olan sınıf seçilir. Olasılık hesaplaması yaparken bir özellik bir sınıfa ait değer üretmiyorsa olasılık hesabında 0 a bölme problemi çıkar. Bu durumun önüne geçebilmek için tüm özelliklere 1 eklenerek yumuşatma işlemi yapılır.[14]

4.1.3 Rastgele Orman (RO) Yöntemi

Rastgele Orman algoritması bir karar ağacı algoritmasıdır. Birden fazla karar ağacını kullanarak sonuc üretir. Breiman ve Cutler [15] tarafından ortaya konmuştur. Birden fazla sınıflandırıcıyla yapılan sınıflandırmalardan en popüler sonucu temel alır. Ağaçlarda düğümler dallara ayrılırken mevcut özelliklerin hepsi kullanılmaz. Rastgele seçilen bir alt küme içerisindeki en iyi sonucu verecek özellik kullanılır. (Şekil 4.5)



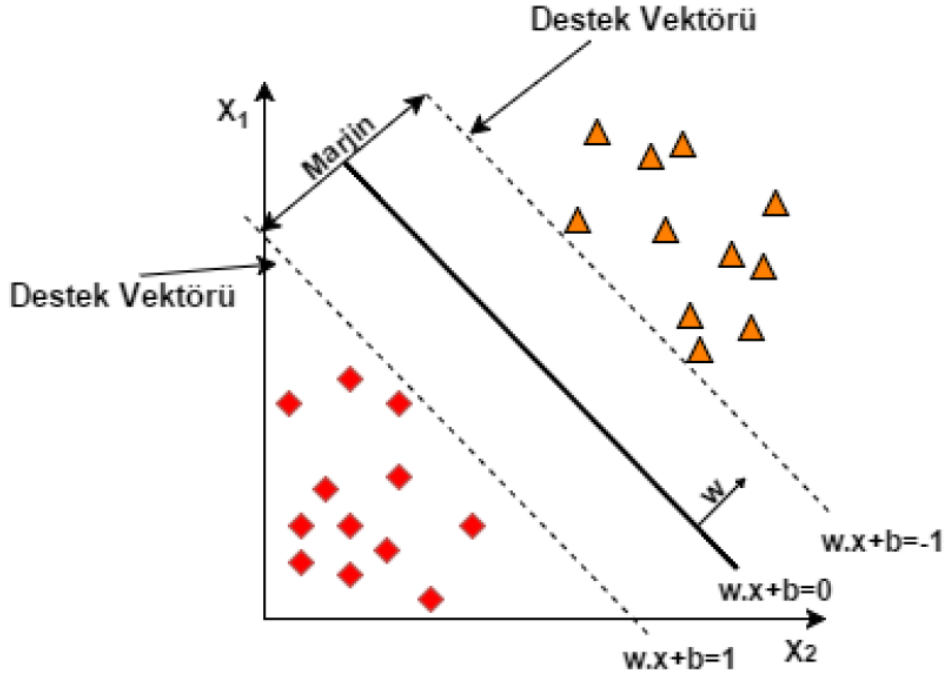
Şekil 4.5 Rastgele Ağaç Orman Yapısı

Bu çalışmada Rastgele Orman yönteminde kullanılan ağaç sayısı 10 dur.

4.1.4 Destek Vektör Makinesi (DVM)

Sınıflandırmada kullanılan basit ve etkili bir yöntemdir. Destek Vektör Makinesinde amaç bir düzlemde bulunan 2 grup arasında her iki gruba da mümkün olan en uzak doğruyu çizmektir.[16]

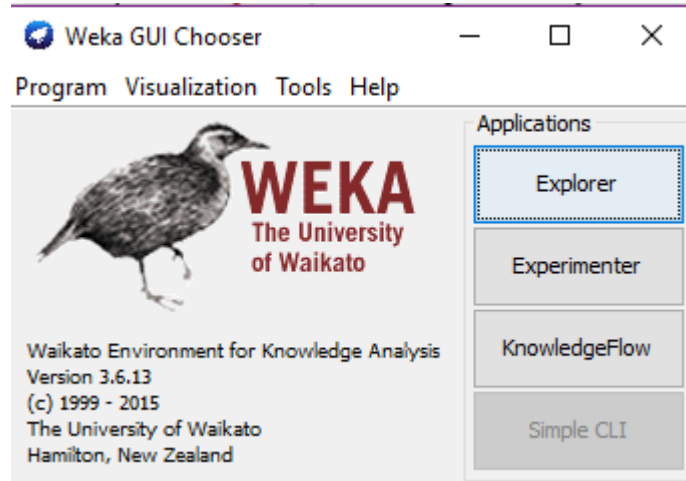
Destek Vektör Makinesinde sınıfları birbirinden başarıyla ayırabilen hiper düzlemlerden en yüksek marjine sahip olan hiper düzlemin seçilmesi amaçlanmıştır. En yüksek marjine sahip hiperdüzlemi bulmak için her 2 sınıfın birbirine en yakın noktalarından geçen ve birbirine paralel olan doğrulara destek vektörleri denir. (Şekil 4.6)



Şekil 4.6 Destek vektörlerince çizilen destek vektörleri, marjın ve hiperdüzlem

4.1.5 Weka Uygulaması

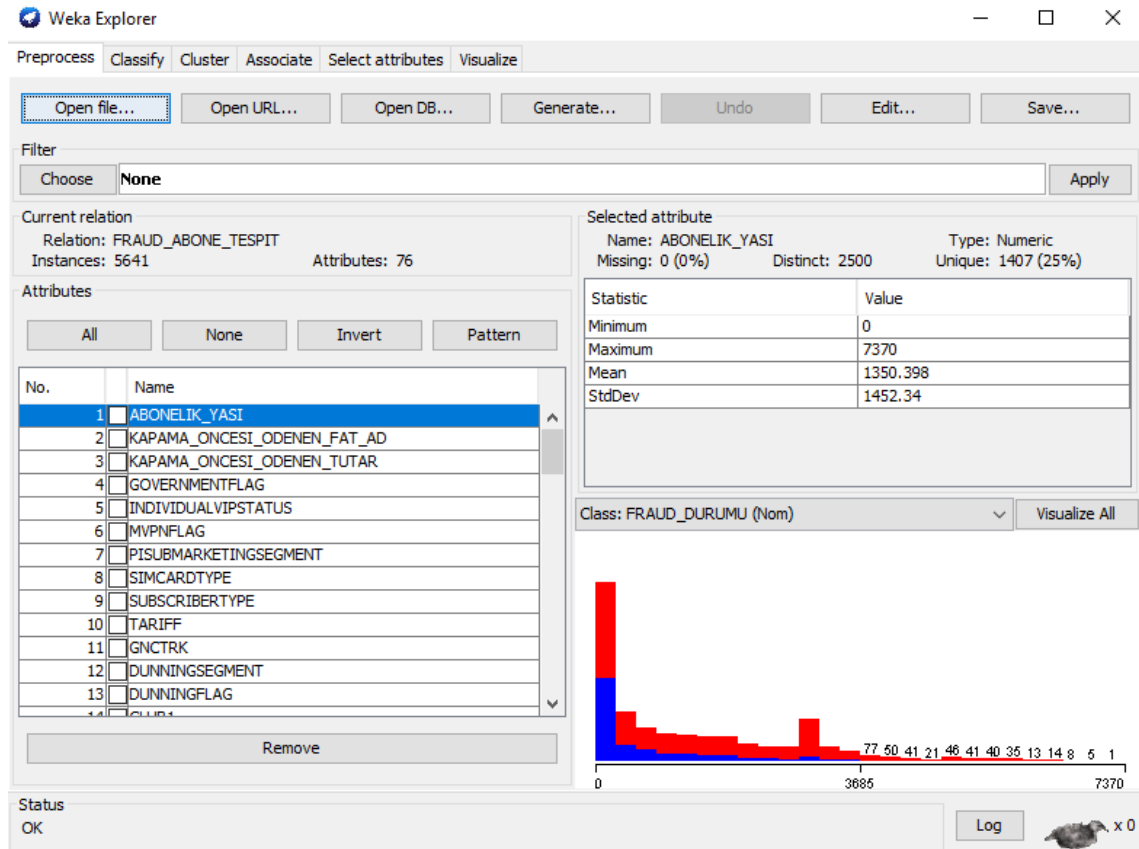
Yeni Zellanda'nın Waikato üniversitesi tarafından geliştirilmiş açık kaynak kodlu java tabanlı bir makine öğrenmesi uygulama platformudur. Kullanılabilecek çok sayıda makine öğrenmesi algoritmasına sahiptir. Görsel bir arayüzü vardır. (Şekil 4.7)



Şekil 4.7 Weka Açılış ekranı

Explorer, Experimenter, KnowledgeFlow ve Simple CLI olmak üzere 4 farklı moda sahiptir. Bu çalışmada Explorer modu kullanılmıştır.

Kendisine ait bir dosya formatı vardır. Arff (attribute relation file format) uzantısına sahip bir dosya uzantısı ile oluşturulan dosyaları işler. Dosya yapısı oldukça anlaşılırdır. Verinin isim ve özelliklerinin yer aldığı başlık kısmı ve verinin kendisi olmak üzere 2 bölümden oluşur. Ana ekranda Open File butonuna basılarak veri uygulamaya yüklenir.

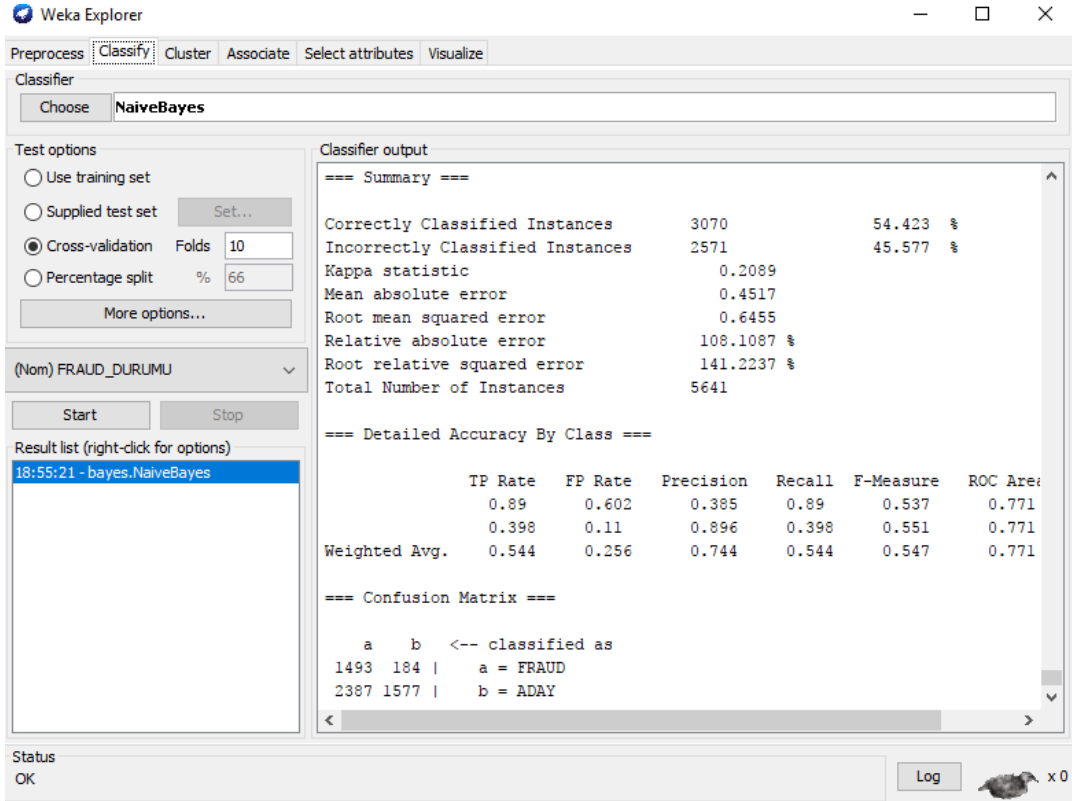


Şekil 4.8 Wekaya Data yüklemesi

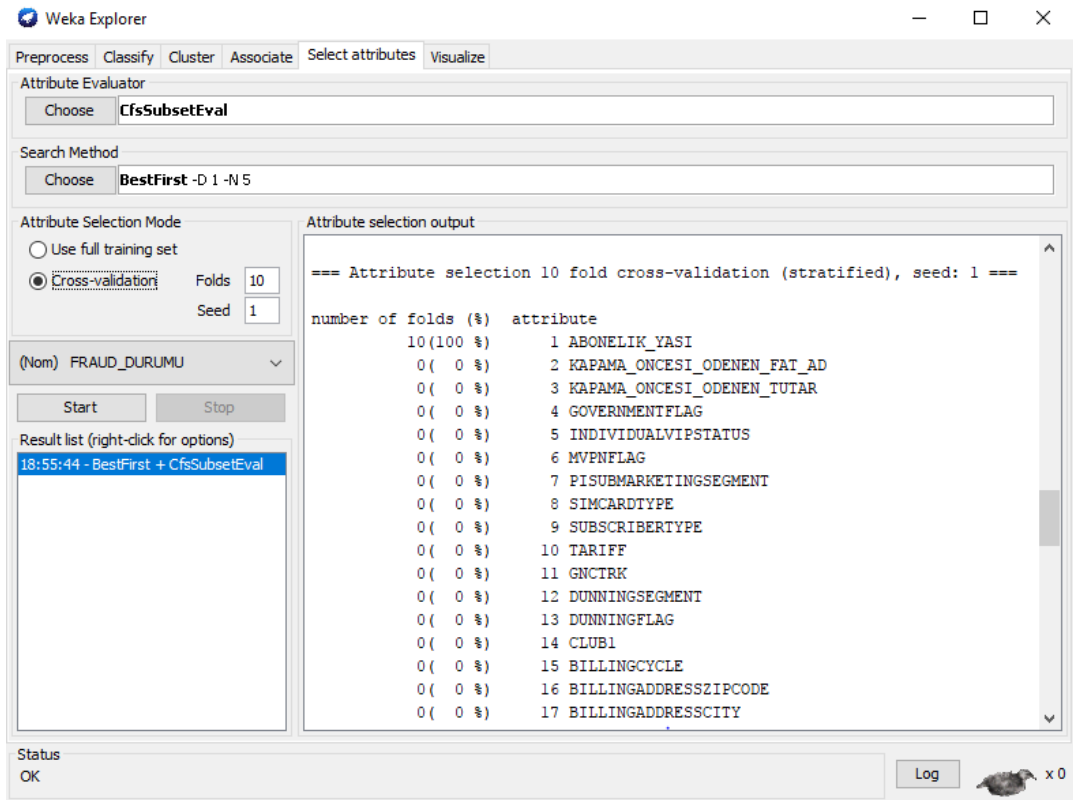
Şekil 4.8 de görsel olarak yüklenen verinin dağılımı görülebilmektedir. Eğer özellik numerik ise verinin minimum, maximum, ortalaması ve standart sapması gösterilir.

En üst menüde yer alan Classify Özelliği kullanılarak yüklenen veri, Weka üzerinde yüklü olan herhangi bir sınıflandırma algoritması kullanılarak sınıflandırılabilir. Sınıflandırma sırasında eğitim ve test verisinin ne olacağına dair çeşitli seçenekler seçilebilir.

Şekil 4.9 da Sınıflandırma işlemine ait örnek bir sonuç görülmektedir. Şekil 4.10 da ise Özellik Seçme işlemine ait bir örnek sonuç görülmektedir.



Şekil 4.9 Sınıflandırma İşlemi ve Sonuç Bilgileri



Şekil 4.10 Özellik Seçme İşlemi ve Sonuç Bilgileri.

4.2 Makine Öğrenmesi Kullanılarak Alınan Sonuçlar

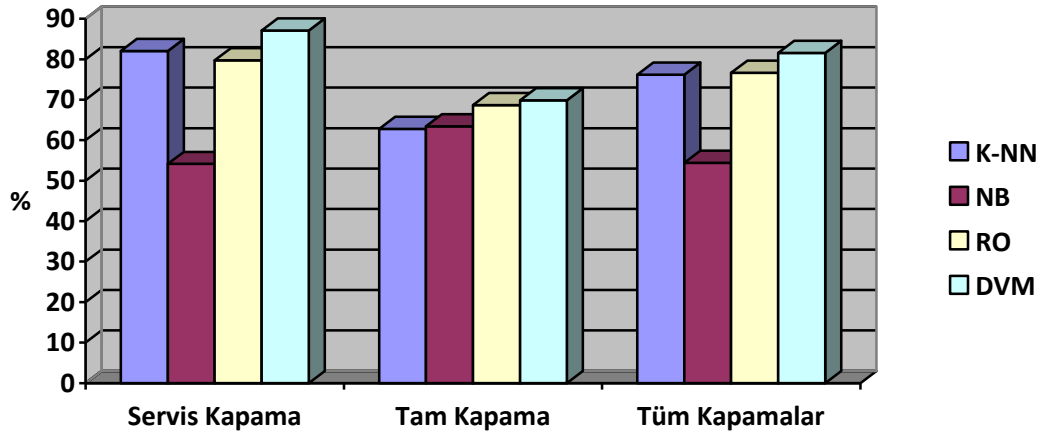
Weka da Aksiyon tipine ve Kullanım tipine göre hazırlanan veri setleri k-En Yakın komşuluk (k-NN), Naive Bayes (NB), Rasgele Orman (RO) ve Destek Vektör Makinesi(DVM) makine öğrenmesi yöntemleri ile sınıflandırılmıştır. Mevcut veri seti özellik indirgeme yöntemleri kullanılarak özellik sayısı azaltılmış ve aynı makine öğrenmesi yöntemleri ile tekrar sınıflandırılmıştır. Denemeler sonucunda alınan başarı değerleri aksiyon ve kullanım tipine göre ayrılarak verilmiştir.

4.2.1 Aksiyon Tipine Göre

İncelenen aksiyon tipleri abonenin sadece servis hizmetlerinin kapatılması ya da tüm kullanımlara kapatılması olup, alınan aksiyon tipine göre başarı oranları Çizelge 4.1 ve görsel olarak da Şekil 4.11’de gösterilmiştir. Aksiyon tipi ile sınıflandırma yaparken 76 özellik kullanılmıştır. Başarı Oranı doğru sınıflandırılan (Accuracy) abonelerin yüzdesidir.

Çizelge 4. 4 Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik)

	Servis Kapama	Tam Kapama	Tüm Kapamalar
K-NN	82,03	62,83	76,15
NB	54,14	63,41	54,42
RO	79,78	68,64	76,67
DVM	87,11	69,87	81,59

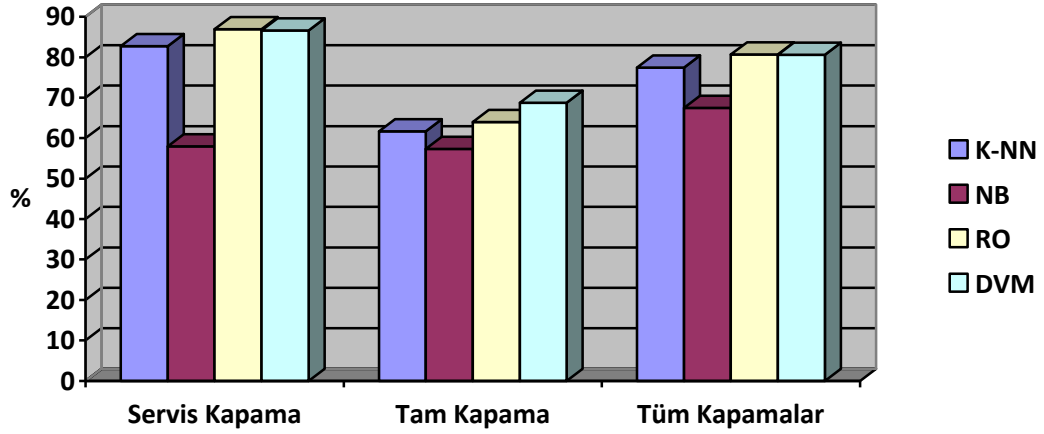


Şekil 4.11 Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik)

Çizelge 4.2 ve görsel olarak ta Şekil 4.12 de indirgenmiş 10 özellik için olan başarı oranları gösterilmiştir.

Çizelge 4. 2 Aksiyon Tipine Göre Başarı Oranları (%) (10 Özellik)

	Servis Kapama	Tam Kapama	Tüm Kapamalar
K-NN	82,74	61,67	77,45
NB	57,95	57,35	67,50
RO	86,94	64	80,69
DVM	86,67	68,77	80,64



Şekil 4.12 Aksiyon Tipine Göre Başarı Oranları (%) (10 Özellik)

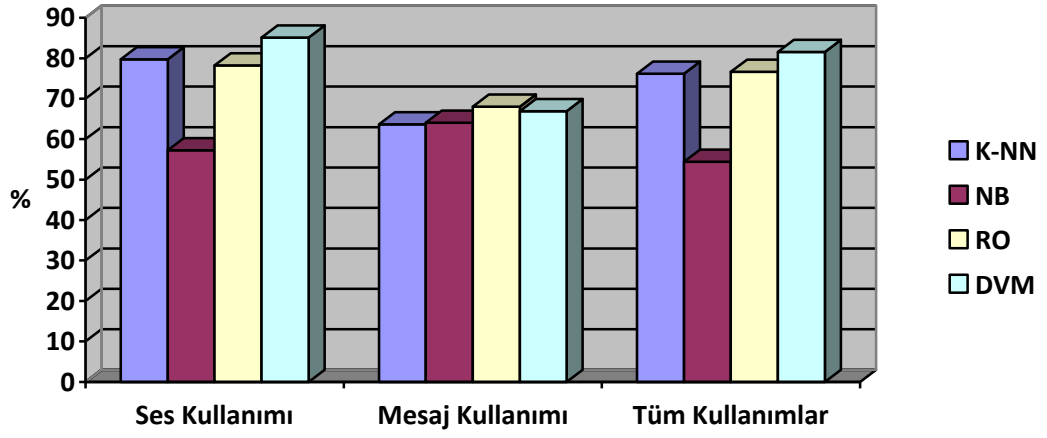
Özellik azaltma yöntemi ile mevcut 76 özellik, 10 özelliğe düşürülmüştür. En başarılı 2 yöntem olarak DVM ve RO yöntemleri ön plana çıkmıştır. RO hesaplama süresi olarak DVM'ye oranla daha hızlı sonuç alma imkanı sunduğu için en uygun hesaplama yöntemi olarak RO yöntemi tercih edilmelidir. Özellik sayısının az olması sistemin başarısını çok fazla etkilememiştir. DVM özellik sayısı fazla iken daha başarılı sonuçlar verirken, RO servis kapama ve servis+tam kapama da en başarılı sonuçları vermiştir.

4.2.2 Kullanım Tipine Göre

Abonelerin kullanım tipi seskullanımı, mesaj kullanımı, ses+mesaj kullanımı olarak 3 şekilde incelenmiştir. Tüm özellikler kullanılarak kullanım tipine göre alınan aksiyonların başarısı sırasıyla Çizelge 4.3 ve görsel olarak Şekil 4.13'te verilmiştir.

Çizelge 4. 3 Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik)

	Ses Kullanımı	Mesaj kullanımı	Tüm Kullanımlar
K-NN	79,73	63,70	76,15
NB	57,29	64,05	54,42
RO	78,24	68,03	76,67
DVM	85,15	66,89	81,59

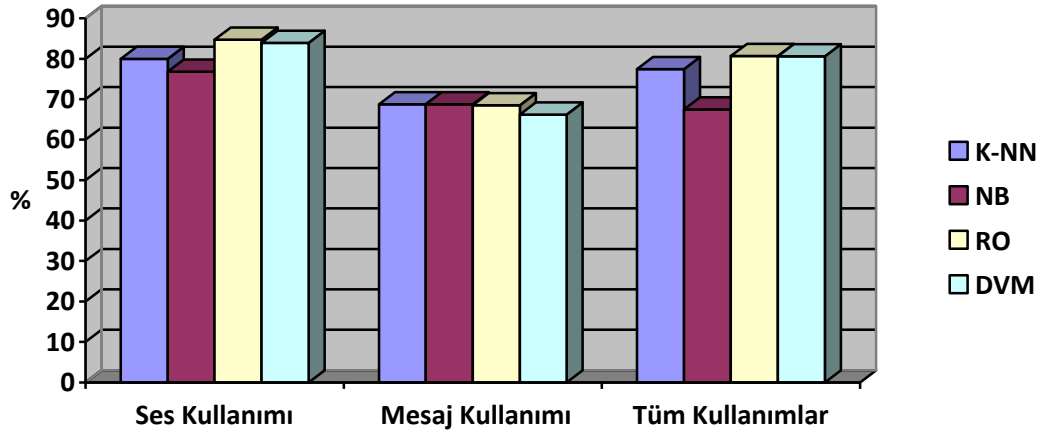


Şekil 4.13 Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik)

Özellik sayısını 76'dan 10'a indirdiğimizde kullanım tipine göre alınan yüzdesel başarı Çizelge 4.4'te, görsel olarak ta Şekil 4.14'te verilmiştir.

Çizelge 4. 4 Kullanım Tipine Göre Başarı Oranları (%) (10 Özellik)

	Ses Kullanımı	Mesaj Kullanımı	Tüm Kullanımlar
K-NN	80	68,71	77,45
NB	76,81	68,71	67,50
RO	84,73	68,48	80,69
DVM	83,93	66,21	80,64



Şekil 4.14 Kullanım Tipine Göre Başarı Oranları (%) (10 Özellik)

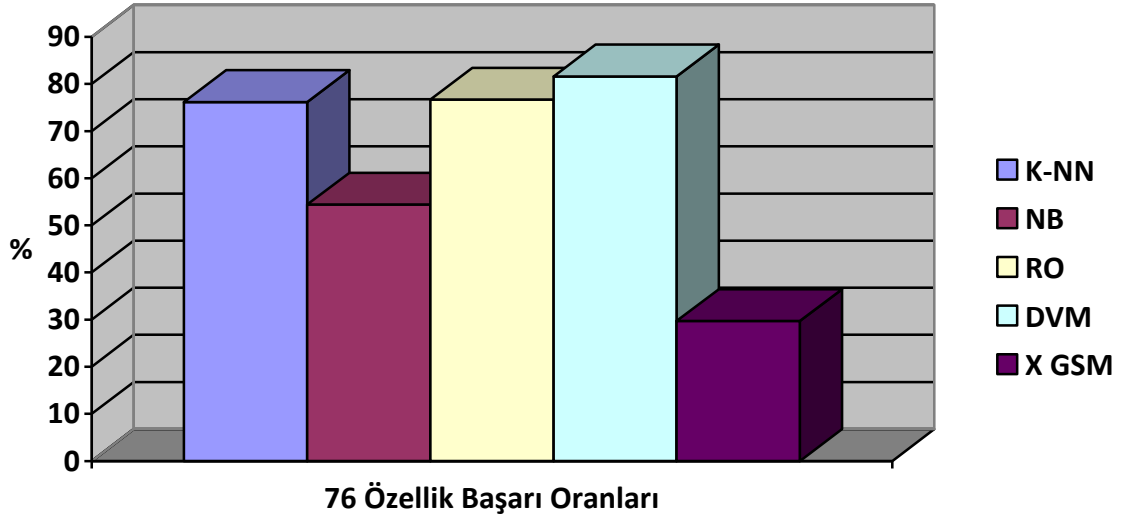
Özellik azaltma yöntemi ile mevcut özellik sayısı 76'dan 10'a düşürülmüş olmasına rağmen kullanım tipinde de başarı performansında büyük bir değişim olmamıştır. Özellik sayısı azaltıldığında DVM hariç tüm yöntemlerde başarı yükselmiştir. Ancak, en yüksek başarı 76 özellik ile tüm kullanım tiplerinde sırası ile %85,15, %66,89 ve %81,59 ile DVM'den alınmıştır. Özellik sayısı azaltıldığında yine en başarılı sınıflandırıcı DVM olsa da tüm özelliklerin kullanımında elde edilen %81,59'luk başarıya göre %2'lik düşme gözükmemektedir.

4.3 Karşılaştırma

X operatörüne ait mevcut sistemin doğru sınıflandırma yüzdesi ile dört farklı makine öğrenmesi yöntemi ile hem 76 hem de 10 özellik için alınan doğru sınıflandırma yüzdesi (Accuracy) sonuçları karşılaştırıldığında Çizelge 4.5'teki sonuçlar elde edilmiştir. Görsel olarak da 76 özellik için Şekil 4.15'te, 10 özellik için de Şekil 4.16'da verilmiştir.

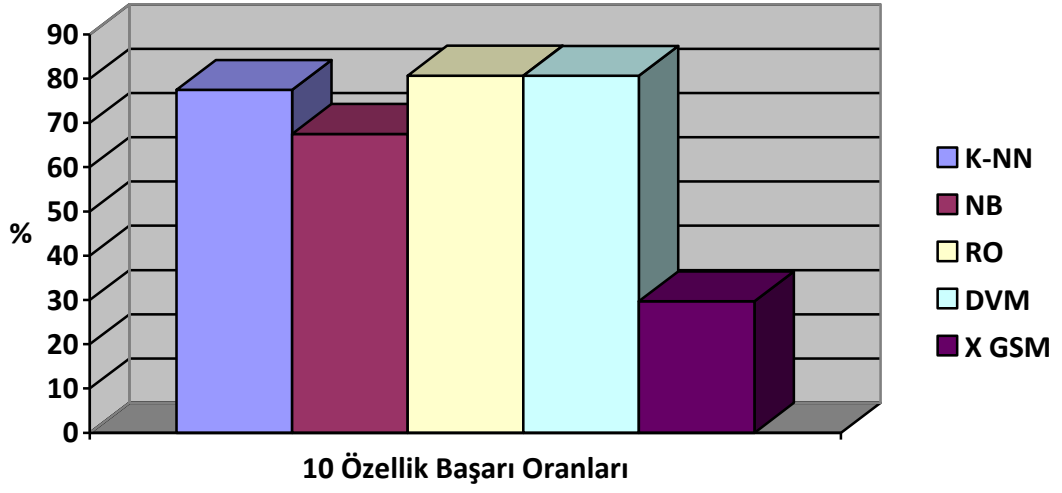
Çizelge 4. 5 Performans Karşılaştırma (%)

	76 Özellik	10 Özellik
K-NN	76,15	77,45
NB	54,42	67,50
RO	76,67	80,69
DVM	81,59	80,64
X GSM ¹	29,7	29,7



Şekil 4.15 76 Özellik için Performans Karşılaştırma

¹ X GSM Operatörüne ait Abonelerin doğru Sınıflandırma yüzdesi



Şekil 4.16 10 Özellik için Performans Karşılaştırma

Yapılan modellemede en başarısız sonucu veren NB yöntemi, X operatörüne ait mevcut sistemin başarısının yaklaşık 2 katından daha fazla başarılı olmuştur. RO ve DVM yöntemlerinde %80'in üzerinde bir başarı ile aboneler doğru tahmin edilmiştir. Böylece seçilen özellikler ve kurulan model ile yaklaşık %50'lik bir abone kitlesi için hatalı aksiyon alınmasının önüne geçilmiştir.

MEVCUT VERİ SETİNİN DERİN ÖĞRENME İLE İŞLENMESİ

Bu bölümde makine öğrenmesi yöntemleri ile elde edilen %80 lik bir başarıyı daha da yükseltmek amacıyla son yıllarda popülerliği gittikçe artan derin öğrenme metodları ile mevcut veri seti sınıflandırılmış ve elde edilen sonuçlar, makine öğrenmesi yöntemleri ve mevcut sistemin performansı ile karşılaştırılmıştır.

Derin Öğrenme, Yapay Sinir Ağlarının çok katmanlı olarak çalıştırılmasıdır.[17] Derin Öğrenme algoritmalarını kullanabilmek amacıyla Matlab 2017 üzerinde yer alan hazır derin öğrenme kütüphanesi kullanılmıştır.

Veri seti önce Autoencoder'dan geçirilmiş daha sonra, Softmax algoritması ve Destek Vektör Makinesi (DVM) ile 10 katlı Çapraz Geçerleme (CV: Cross Validation k:10) kullanılarak sınıflandırılmıştır.

Derin öğrenme de sayısal verilerle çalışabildiği için mevcut veri setindeki 76 özelliğten sayısal olmayan veriler numerik değerlere çevrilmiştir.

5.1 Kullanılan Yöntemler

Öncelikle her bir veri seti AutoEncoder'dan geçirildikten sonra 10 katlı Çapraz Geçerleme yöntemi ile Softmax ve Destek Vektör Makinesi algoritmaları kullanılarak sınıflandırılmıştır.

AutoEncoder ileri beslemeli Yapay Sinir Ağlarında kullanılır. Örneğin 20x20 pikseklik bir resmi verirken Yapay Sinir Ağının giriş boyutu 400 olacaktır. HiddenSize 50 seçilirse bu 400 özelliğten resmi ayırtedecek en başarılı 50 tanesi seçilerek ileri beslemeli Yapay Sinir Ağında sınıflandırma işlemi yapılacak demektir.[18]

AutoEncoder işlemi yapılırken kullanılan parametreler şu şekildedir.

HiddenSize=38 (Gizli Katman sayısı)

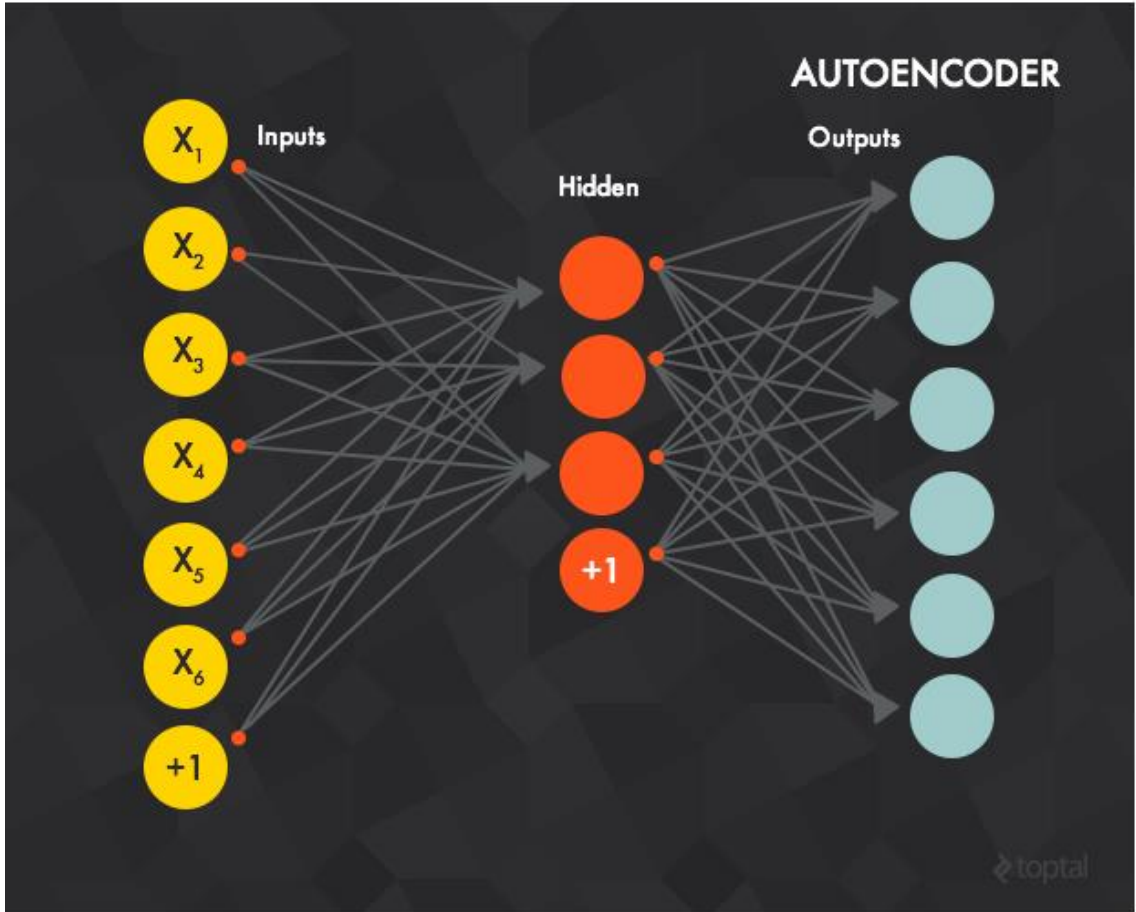
MaxEpoch=1000 (Maximum iterasyon Sayısı)

L2w=0.001 (L2WeightRegularization- Ağırlık Düzenleme Katsayısı)

Sr=4 (SparsityRegularization- Seyreklik Düzenleme Katsayısı)

Sp=0.005 (SparsityProportion – Seyreklik Oranı Katsayısı)

Şekil 5.1’de Autoencoder’in mimari modeli gösterilmiştir.



Şekil 5.1 AutoEncoder Mimari Modeli [18]

5.2 Alınan Sonuçlar

Makine öğrenmesi yöntemleri ile karşılaştırmak amacıyla mevcut veri seti hem aksiyon tipine göre hem de kullanım tipine göre sınıflandırılmıştır. Özellik indirgemesi

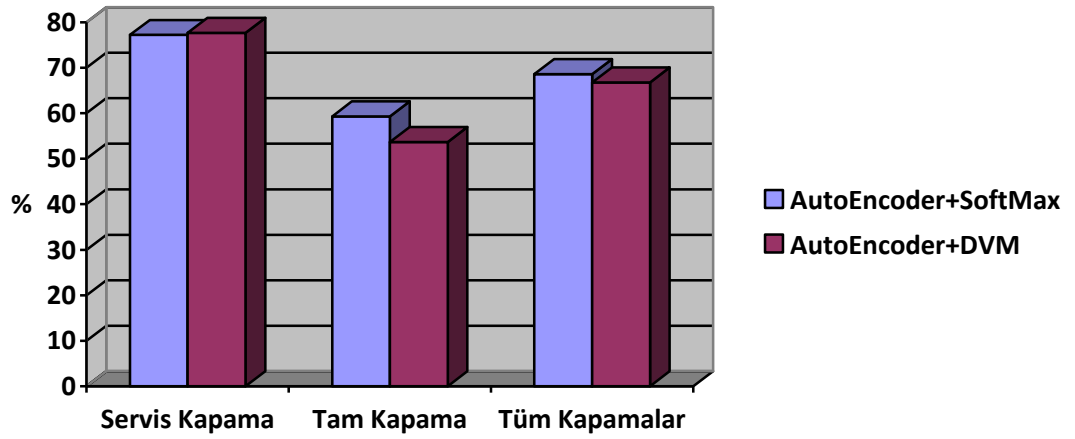
yapılmadan 76 özellik ile veri seti önce AutoEncoder'dan geçirilmiş. Daha sonra da 10 katlı Çapraz geçerleme yöntemiyle Softmax ve Destek Vektör Makinesi sınıflandırma algoritmalarıyla sınıflandırılarak başarıları ölçülmüştür.

5.2.1 Aksiyon Tipine Göre Sonuçlar

Alınan aksiyon tipine göre alınan yüzdesel başarı oranları Çizelge 5.1 de ve görsel gösterimi de Şekil 5.2'de paylaşılmıştır.

Çizelge 5. 1 Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik)

	Servis Kapama	Tam Kapama	Tüm Kapamalar
AutoEncoder +Softmax	77,22	59,31	68,58
AutoEncoder+DVM	77,68	53,68	66,76



Şekil 5.2 Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik)

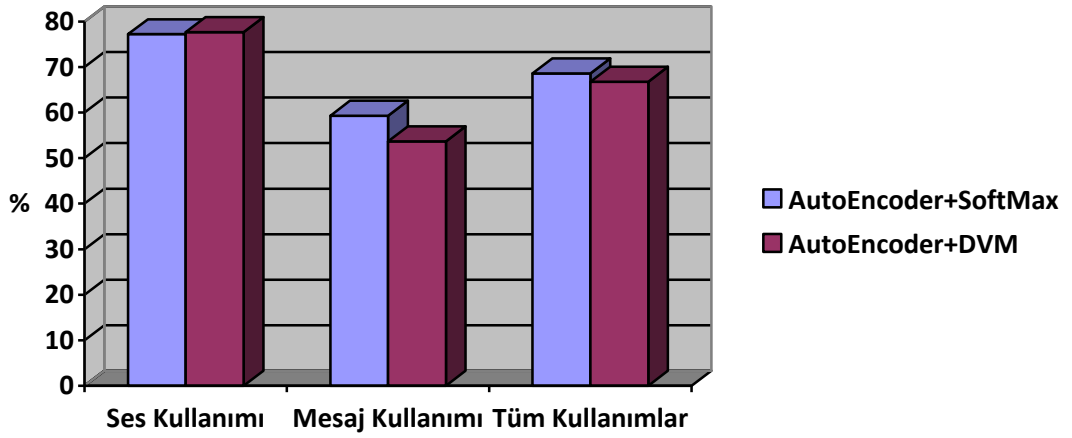
Kapama tipine göre yapılan sınıflandırmada Servis Kapama aksiyonunda elde edilen başarı oranları Makine Öğrenmesi ile elde edilen değerlere yakın olmasına rağmen, Tam Kapama aksiyonunda elde edilen düşük değerler, Tüm kapama Aksiyonları için elde edilen sonuçları da aşağıya çekmiştir. Tüm kapamalar için olan başarı oranı ortalama %67,7 de kalmıştır.

5.2.2 Kullanım Tipine Göre Sonuçlar

Kullanım tipine göre 76 özellik için alınan yüzdesel başarı oranları Çizelge 5.2’de paylaşılmıştır. Şekil 5.3’te de görsel olarak sonuçlar sunulmuştur.

Çizelge 5. 2 Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik)

	Ses Kullanımı	Mesaj Kullanımı	Tüm Kullanımlar
AutoEncoder +Softmax	73,74	54,48	65,58
AutoEncoder+DVM	72,36	45,21	66,76



Şekil 5.3 Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik)

Kullanım tipine göre yapılan sınıflandırma da Ses kullanımlarına ait elde edilen başarı oranları Makine Öğrenmesi ile elde edilen edilen değerlere yakın olmasına rağmen, Mesaj kullanımına ait elde edilen düşük değerler, Tüm kullanımlar için elde edilen sonuçları da aşağıya çekmiştir. Tüm kullanımlar için olan başarı oranı %67,7 de kalmıştır.

BÖLÜM 6

PERFORMANS KARŞILAŞTIRMASI

Bu bölümde mevcut sistemin, makine öğrenmesi yöntemleri ile geliştirilen sistemin ve derin öğrenme yöntemleri ile elde edilen sistemin performans karşılaştırması yapılmıştır. Başarı Oranı doğru sınıflandırma (Accuracy) yüzdesini gösterir.

6.1 Aksiyon Bazlı Performans Karşılaştırması

Çizelge 6.1’de makine öğrenmesi, derin öğrenme ve mevcut X operatörünün aksiyon bazlı aldıkları başarı 76 özellik için verilmiştir.

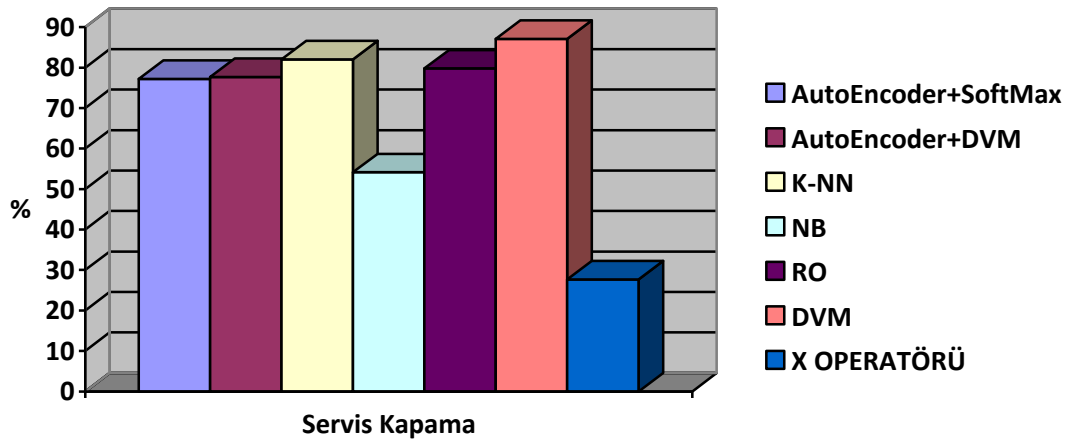
Çizelge 6. 1 Aksiyon Tipine Göre Başarı Oranları (%) (76 Özellik)

	Servis Kapama	Tam Kapama	Tüm Kapamalar
AutoEncoder +Softmax	77,22	59,31	68,58
AutoEncoder+DVM	77,68	53,68	66,76
K-NN	82,03	62,83	76,15
NB	54,14	63,41	54,42
RO	79,78	68,64	76,67
DVM	87,11	69,87	81,59
X Operatörü ¹	27,7	34,8	29,7

¹ X GSM Operatörüne ait Aksiyon alınan abonelerin aksiyon bazlı doğru sınıflandırma yüzdeleri

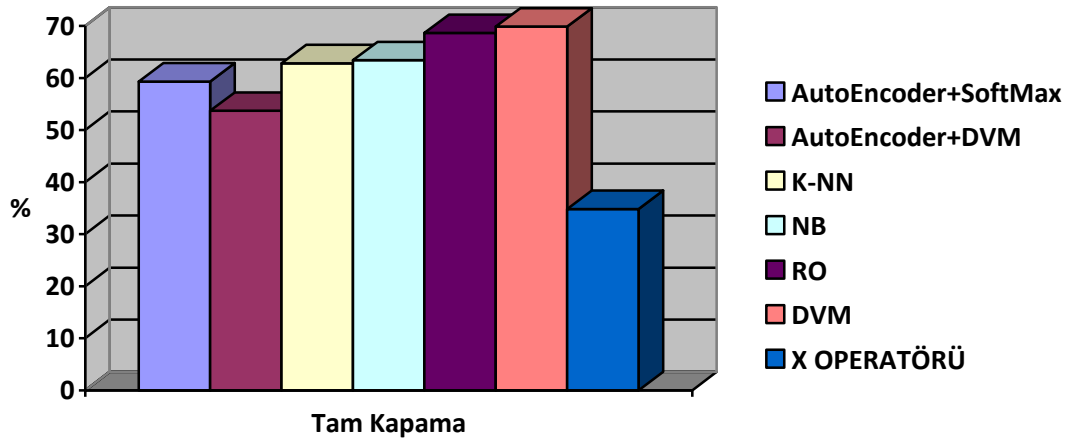
X GSM operatörünün aksiyon almada başarısızlığı oldukça yüksektir. Makine öğrenmesi ve derin öğrenme ile alınan başarılar yaklaşık 3 kat daha fazladır

Servis Kapama Aksiyonu alınan aboneler için başarı oranları Şekil 6.1’de görsel olarak gösterilmektedir.



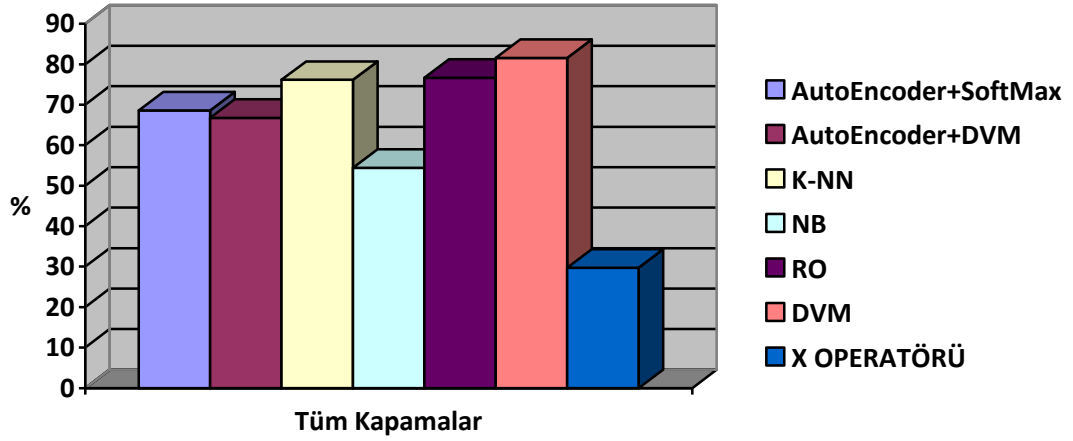
Şekil 6.1 Servis Kapama aksiyonuna göre başarı oranları (%) (76 Özellik)

Kapama Aksiyonu alınan aboneler için de başarı oranı Şekil 6.2’de gösterilmiştir.



Şekil 6.2 Tam Kapama aksiyonuna göre başarı oranları (%) (76 Özellik)

Tüm kapamalar için de başarı oranları Şekil 6.3’de görsel olarak verilmiştir.



Şekil 6.3 Tümü Kapama aksiyonları için başarı oranları (%) (76 Özellik)

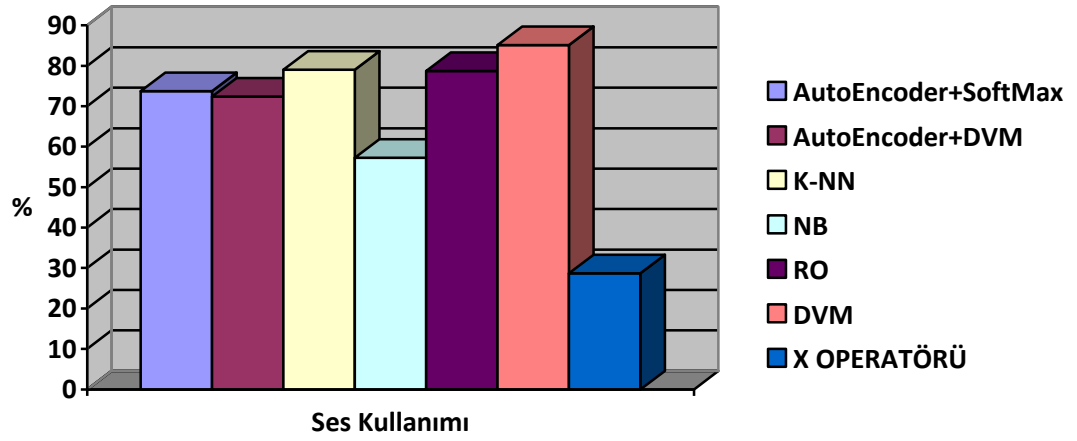
Aksiyon bazlı kapamada hem Servis Kapama, hem Tam Kapama hem de Tümü Kapamalar için DVM en yüksek başarı oranını göstermektedir.

6.2 Kullanım Bazlı Performans Karşılaştırması

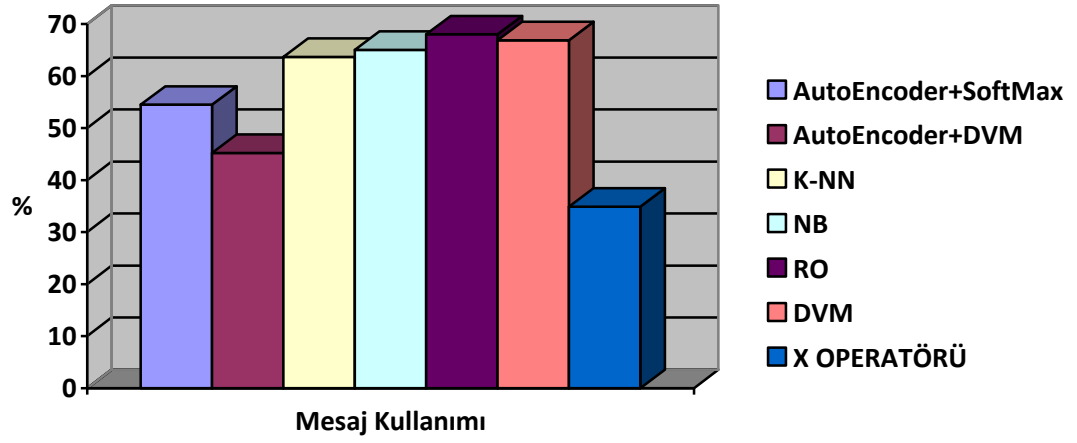
Çizelge 6.2’de makine öğrenmesi, derin öğrenme ve mevcut X operatörünün kullanım bazlı aldıkları yüzdesel başarı 76 özellik için verilmiştir. Şekil 6.4, Şekil 6.5, Şekil 6.6’da kullanım tipi olarak sırayla ses, mesaj ve tüm kullanımlar için başarılar görsel olarak verilmiştir.

Çizelge 6. 2 Kullanım Tipine Göre Başarı Oranları (%) (76 Özellik)

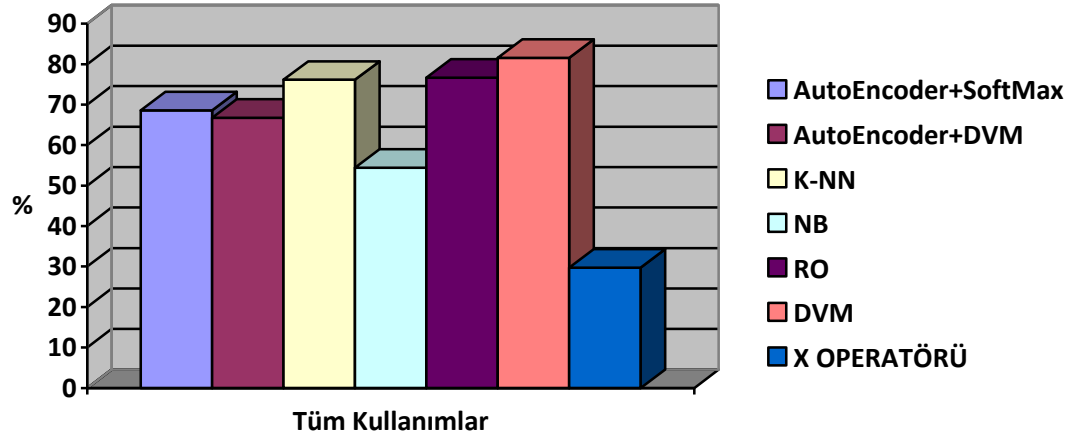
	Ses Kullanımı	Mesaj Kullanımı	Tümü Kullanımlar
AutoEncoder +Softmax	73,74	54,48	68,58
AutoEncoder+DVM	72,36	45,21	66,76
K-NN	79,03	63,70	76,15
NB	57,29	65,05	54,42
RO	78,74	68,03	76,67
DVM	85,15	66,89	81,59
X Operatörü	28,7	34,9	29,7



Şekil 6.4 Ses Kullanımı yapan aboneler için başarı oranları (%) (76 Özellik)



Şekil 6.5 Mesaj kullanımı yapan aboneler için başarı oranları (%) (76 Özellik)



Şekil 6.6 Tüm Kullanımlar için başarı oranları (%) (76 Özellik)

Ses kullanımında, mesaj kullanımında ve tüm kullanımlarda DVM en yüksek başarı oranını elde etmiştir.

SONUÇ VE ÖNERİLER

Yapılan bu çalışmada X operatörünün Fraud olarak sınıflandırdığı 5.641 abonesi alınan Aksiyon tipi ve Kullanım bazında olmak üzere incelenmiştir. Şu an da mevcut kullanılan sistemin başarısı, yaklaşık %30 doğruluk ile tespit edilebilmektedir. Yaklaşık %70’lik bir abone kitlesine alınan hatalı aksiyonların etkisi incelendiğinde, diğer operetörlere geçme oranlarının, hiç aksiyon alınmayan abonelere oranla 50 kat fazla olduğu görülmüştür. Günümüzde sosyal medyanın yaygın kullanılmasıyla başka operatörlere geçen abonelerin memnuniyetsizliklerini başka insanlarla paylaştığını düşünürsek tespit edilenin üzerinde bir sayının başka bir operatöre geçtiği ya da diğer operatörlerden X operatörüne geçmekten vazgeçtiği düşünülebilir.

DVM ile elde edilen % 81,6 ‘lık başarılı sınıflandırma ile yaklaşık %50’lik bir abone kitlesinin hatalı sınıflandırılması engellenmiş olacaktır.

Hatalı alınan aksiyonların engellenmesi sonucunda orta ve uzun vadede müşteri memnuniyetinin artması sağlanacaktır. 1500 aboneye hatalı aksiyon alınmasının önüne geçilecek, ayrıca aylık enaz 100 abonenin operatör değiştirmesi engellenecektir.

Aksiyon alınan her abone için ara ödeme çağrısı açılarak abonenin Müşteri Hizmetlerindeki çalışanlar aracılığıyla aranarak bilgi verilmesi gerekmektedir. 1500 abonenin aranmaması, her ay için operatöre 2 kişi daha az çalıştırma imkanı verecektir. Ayrıca, abonelerin müşteri hizmetlerini arayarak şikayet etmeleri de engellenmiş olacaktır.

KAYNAKLAR

- [1] Magnify, FraudFocus Advanced Fraud Detection, White Paper, Chicago, 2002.
- [2] Bolton, R., Hand, D.,(2002) "Statistical Fraud Detection: A Review (With Discussion)", Statistical Science, 17(3): 235-255.
- [3] Bolton, R. ve Hand, D. ,(2001) "Unsupervised Profiling Methods for FraudDetection", Credit Scoring and Credit Control VII, 2001.
- [4] Burge, P., ve Shawe-Taylor, J,(2001) "An Unsupervised Neural Network Approach to Profiling the Behavior of Mobile Phone Users for Use in Fraud Detection", Journal of Parallel and Distributed Computing, 61: 915-925.
- [5] Cahill, M., Chen, F., Lambert, D., Pinheiro, J., Sun, D.,(2002) "Detecting Fraud in the Real World. Handbook of Massive Datasets", 911-930.
- [6] Bhargava, B., Zhong, Y. ve Lu, Y.,(2003) "Fraud Formalization and Detection", Proc. of DaWaK2003, 330-339.
- [7] Wheeler, R., ve Aitken, S.,(2000) "Multiple Algorithms for Fraud Detection", Knowledge-Based Systems, 13(3): 93-99.
- [8] Maes S., Tuyls K., Vanschoenwinkel B. ve Manderick B.,(2002) "CreditCard Fraud Detection using Bayesian and Neural Networks", Proc. of the 1st International NAISO Congress on Neuro Fuzzy Technologies.
- [9] Foster, D., ve Stine, R.,(2004) "Variable Selection in Data Mining: Building aPredictive Model for Bankruptcy", Journal of American Statistical Association, 99: 303-313.
- [10] Kaplan, Y., "Telekom Sistemlerinde Fraud (Sahtekarlık)", <https://www.yasinkaplan.com/tr/docs/Fraud.pdf>, 11 Temmuz 2017.
- [11] Moreau, Y.,Lerouge, E., Verrelst, H., Vandewalle, J., Störmann, C.,Burge, P., "A hybrid system for fraud detection in mobile communications",ESANN'1999 proceedings - European Symposium on Artificial Neural NetworksBruges (Belgium), 21-23 April 1999, D-Facto public., 447-454.

- [12] Awoyelu, I., Adebomi, A., Amoo, A., Adebisi, R., Mabude, C., (2015) "Fraud Detection in Telecommunications Industry: Bridging the Gap with Random Rough Subspace Based Neural Network Ensemble Method", Computer Engineering and Intelligent Systems, 6:10.
- [13] Seker, S. , "KNN (K nearest neighborhood, en yakın k komşu", <http://bilgisayarkavramlari.sadievrenseker.com/2008/11/17/knn-k-nearest-neighborhood-en-yakin-k-komsu/>, 20 Ağustos 2017.
- [14] Seker, S., "Naive Bayes ile Metin Sınıflandırılması" , <http://bilgisayarkavramlari.sadievrenseker.com/2013/02/08/naive-bayes-ile-metin-siniflandirilmesi/>, 22 Ağustos 2017.
- [15] Breiman, L., ve Cutler, A., Random Forest, http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm, 22 Ağustos 2017.
- [16] Akarsu, C., (2016) Twitter Verileri ile Türk Televizyonları İzlenme Oranı Sıralamaları Tahmini, Yüksek Lisans Tezi, YÜ Fen Bilimleri Enstitüsü, İstanbul
- [17] Bakshi, A., "What is Deep Learning? Getting Started With Deep Learning", <https://www.edureka.co/blog/what-is-deep-learning>, 20 Eylül 2017.
- [18] Vasilev, I., "A Deep Learning Tutorial: From Perceptrons to Deep Networks", <https://www.toptal.com/machine-learning/an-introduction-to-deep-learning-from-perceptrons-to-deep-networks>, 21 Eylül 2017.

ÖZELLİKLER

Abonelere ait mevcut özellikler ve yeni oluşturulan özellikler ve açıklamaları aşağıdaki gibidir.

FRAUD_DURUMU: Abonenin gerçekte dolandırıcı olup olmadığını belirten ve sınıflandırmak amacıyla kullanılan özelliğidir. ADAY ve FRAUD olmak üzere 2 farklı sınıf mevcuttur.

ABONELIK_YASI: Abonenin hattını açtığı tarihten itibaren kaç gün geçtiğini gösterir.

KAPAMA_ONCESI_ODENEN_FAT_AD: Kapama öncesi abonenin kaç adet fatura ödediği bilgisidir.

KAPAMA_ONCESI_ODENEN_TUTAR: Kapama öncesi abonenin ödediği toplam fatura tutarı bilgisini gösterir.

GOVERNMENTFLAG: Hattın devlete ait olup olmadığını gösterir.

INDIVIDUALVIPSTATUS: Abonenin toplumda önemli bir etkisi olup olmadığını gösterir.

MVPNFLAG: Abonenin yüksek ödeme gücüne sahip olup olmadığını gösterir.

PISUBMARKETINGSEGMENT: Abonenin sahip olduğu alt pazarlama segmentini gösterir.

SIMCARDTYPE: Abonenin kullandığı simcard'ın tipini gösterir.

SUBSCRIBERTYPE: Faturalı ya da ön ödemeli abone gibi abonelik tipini gösterir.

TARIFF: Abonenin tarifesi gösterir.

GNCTRK: Abonenin genç tarifesinde olup olmadığını gösterir.

DUNNINGSEGMENT: Abonenin borç ödeme notunu gösterir.

DUNNINGFLAG: Abonenin ödenmemiş borcu olduğu için takipte olup olmadığını gösterir.

CLUB1: X tarifesinde olup olmadığını bilgisini gösterir.

BILLINGCYCLE: Faturalama zamanını gösterir.

BILLINGADDRESSZIPCODE: Faturalama adresine ait posta kodu bilgisini gösterir.

BILLINGADDRESSCITY: Faturalama adresine ait şehri gösterir.

CLUB2: Y tarifesinde olup olmadığını bilgisini gösterir.

CORPORATEMARKETINGSUBSEGMENT: Kurumsal alt pazarlama segment bilgisini gösterir.

CORPORATETYPE: Kurumsal segment bilgisini gösterir.

CREDITLIMITACCT: X yöntemiyle hazırlanmış müşteri limiti.

CUSTOMERTYPE: Müşteri tipini gösterir. (Şahıs ya da şirket gibi)

NUMPOSLINES: Sahip olduğu faturalı hat sayısını gösterir.

SEKTOR: Sektör bilgisini gösterir.

SASCREDITLIMITACCT: Y yöntemiyle hazırlanmış müşteri limiti.

ROAMFLAG: Abonenin yurtdışında olup olmadığını bilgisini gösterir.

SCOREBAND: Z yöntemiyle hazırlanmış müşteri limiti.

FINANCIALPRIVILEGEFLAG: Finansal olarak ayrıcalıklı bir abone olup olmadığını gösterir.

UCRETLI_VAS_KULLANIMI_TUTAR: Ücretli katma değerli servis kullanımı toplam tutarı.

UCRETLI_VAS_KULLANIM_ADET: Ücretli katma değerli servis kullanımı toplam adet.

PAYORDER: Abonenin otomatik ödeme talimatı olup olmadığını gösterir.

EINVOICE: Abonenin faturasını e- fatura olarak alıp almadığını gösterir.

INTERNATIONALSTATUS: Abonenin yurt dışındaki numaraları arayabilme durumunu gösterir.

ROAMINGSTATUS: Abonenin yurtdışında görüşme yapabilme durumunu gösterir.

ISPORTIN: Başka bir operatörden gelip gelmediği bilgisini gösterir.

VIDEOCALLSTAT: Abonenin video görüşme yapabilme durumunu gösterir.

IVR_ODEME_ADET: IVR sistemi üzerinden ödediği fatura adetini gösterir.

IVR_TOPLAM_ODEME: IVR sistemi üzerinden ödediği toplam fatura tutarını gösterir.

CSI_ODEME_ADET: Operatör üzerindeki kayıtlı kredi kartı ile yaptığı fatura ödeme sayısını gösterir.

CSI_TOPLAM_ODEME: Operatör üzerindeki kayıtlı kredi kartı ile yaptığı faturaların toplam tutarı

UCRETLI_SMS_CDR_ADET: Gönderdiği toplam ücretli SMS adetini gösterir.

TOPLAM_SMS_UCRET: Gönderdiği tüm ücretli SMS tutarının toplamını gösterir.

UCRETLI_SMS_CDR_ADET_5TL_U: Tanesi 5 lira ve üstü olan ücretli SMS gönderim adetini gösterir.

TOPLAM_SMS_UCRET_5TL_U: Tanesi 5 lira ve üstü olan ücretli SMS tutarının toplamını gösterir.

UCRETLI_SMS_CDR_ADET_10TL_U: Tanesi 10 lira ve üstü olan ücretli SMS gönderim adetini gösterir.

TOPLAM_SMS_UCRET_10TL_U: Tanesi 10 lira ve üstü olan ücretli SMS tutarının toplamını gösterir.

UCRETLI_SMS_CDR_ADET_20TL_U: Tanesi 20 lira ve üstü olan ücretli SMS gönderim adetini gösterir.

TOPLAM_SMS_UCRET_20TL_U: Tanesi 20 lira ve üstü olan ücretli SMS tutarının toplamını gösterir.

TOPLAM_ARAMA_SURESI: Abonenin yaptığı toplam arama süresini gösterir.

TOPLAM_ARAMA_SAYISI: Abonenin yaptığı toplam arama sayısını gösterir.

TOPLAM_ARAMA_SURESI_UCRETLI: Abonenin yaptığı toplam ücretli arama süresini gösterir.

TOPLAM_UCRETLI_ARAMA_SAYISI: Abonenin yaptığı toplam ücretli arama sayısını gösterir.

TOPLAM_ARAMA_SURESI_5TL_U: Abonenin yaptığı ücret tutarı 5 TL ve üzeri olan aramaların toplam süresini gösterir.

TOPLAM_ARAMA_SAYISI_5TL_U: Abonenin yaptığı ücret tutarı 5 TL ve üzeri olan aramaların toplam sayısını gösterir.

TOPLAM_ARAMA_SURESI_10TL_U: Abonenin yaptığı ücret tutarı 10 TL ve üzeri olan aramaların toplam süresini gösterir.

TOPLAM_ARAMA_SAYISI_10TL_U: Abonenin yaptığı ücret tutarı 10 TL ve üzeri olan aramaların toplam sayısını gösterir.

TOPLAM_ARAMA_SURESI_20TL_U: Abonenin yaptığı ücret tutarı 20 TL ve üzeri olan aramaların toplam süresini gösterir.

TOPLAM_ARAMA_SAYISI_20TL_U: Abonenin yaptığı ücret tutarı 20 TL ve üzeri olan aramaların toplam sayısını gösterir.

TOPLAM_ARAMA_SURESI_50TL_U: Abonenin yaptığı ücret tutarı 50 TL ve üzeri olan aramaların toplam süresini gösterir.

TOPLAM_ARAMA_SAYISI_50TL_U: Abonenin yaptığı ücret tutarı 50 TL ve üzeri olan aramaların toplam sayısını gösterir.

SURE_SANIYESI_5KURUS_U: Abonenin saniyesi 5 kuruş ve üzeri olan aramalarına ait toplam süreyi gösterir.

UCRET_SANIYESI_5KURUS_U: Abonenin saniyesi 5 kuruş ve üzeri olan aramalarına ait toplam tutarı gösterir.

SAYI_SANIYESI_5KURUS_U: Abonenin saniyesi 5 kuruş ve üzeri aramalarının toplam adetini gösterir.

SURE_SANIYESI_10KURUS_U: Abonenin saniyesi 10 kuruş ve üzeri olan aramalarına ait toplam süreyi gösterir.

UCRET_SANIYESI_10KURUS_U: Abonenin saniyesi 10 kuruş ve üzeri olan aramalarına ait toplam tutarı gösterir.

SAYI_SANIYESI_10KURUS_U: Abonenin saniyesi 10 kuruş ve üzeri aramalarının toplam adetini gösterir.

SURE_SANIYESI_20KURUS_U: Abonenin saniyesi 20 kuruş ve üzeri olan aramalarına ait toplam süreyi gösterir.

UCRET_SANIYESI_20KURUS_U: Abonenin saniyesi 20 kuruş ve üzeri aramalarının toplam tutarını gösterir.

SAYI_SANIYESI_20KURUS_U: Abonenin saniyesi 20 kuruş ve üzeri aramalarının toplam adetini gösterir.

SURE_SANIYESI_50KURUS_U: Abonenin saniyesi 50 kuruş ve üzeri olan aramalarına ait toplam süreyi gösterir.

UCRET_SANIYESI_50KURUS_U: Abonenin saniyesi 50 kuruş ve üzeri aramalarının toplam tutarını gösterir.

SAYI_SANIYESI_50KURUS_U: Abonenin saniyesi 50 kuruş ve üzeri aramalarının toplam adetini gösterir.

SURE_SANIYESI_1LIRA_U: Abonenin saniyesi 1 TL ve üzeri olan aramalarına ait toplam süreyi gösterir.

UCRET_SANIYESI_1LIRA_U: Abonenin saniyesi 1 TL ve üzeri aramalarının toplam tutarını gösterir.

SAYI_SANIYESI_1LIRA_U: Abonenin saniyesi 1 TL ve üzeri aramalarının toplam adetini gösterir.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Onur TÜFEKÇİOĞLU
Doğum Tarihi ve Yeri : 02.07.1979, Ankara
Yabancı Dili : İngilizce
E-posta : tufekciogluonur@yahoo.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Yüksek Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	2017
Lisans	İşletme	Anadolu Üniversitesi	2007
Lisans	Elektronik ve Haberleşme Mühendisliği	Yıldız Teknik Üniversitesi	2003
Lise	Fen	Isparta Anadolu Lisesi	1999

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2011-	Turkcell	Mühendis
2008-2011	Avea	Mühendis

YAYINLARI

Makale

1. Tüfekçioğlu, O., Biricik, G., Diri, B., (2016), "Fraud Risk Management For GSM Systems", "Sigma Journal of Engineering and Natural Sciences Dergisi", 8:169-174.

Bildiri

1. Tüfekçioğlu, O., Diri, B., (2016), "GSM Sistemlerinde Dolandırıcılık Risk Yönetimi", ASYU 2016 (Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu), 29 Eylül - 1 Ekim 2016, Düzce.