DOKUZ EYLÜL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SOSYAL MEDYADA SANAL ZORBALIK İÇERİKLERİNİN TESPİTİ

Mikayil SADİGZADE

Haziran, 2022
İZMİR

SOSYAL MEDYADA SANAL ZORBALIK İÇERİKLERİNİN TESPİTİ

Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü

Yüksek Lisans Tezi

Bilgisayar Bilimleri Anabilim Dalı

Mikayil SADİGZADE

Haziran, 2022

İZMİR

YÜKSEK LİSANS TEZİ SINAV SONUÇ FORMU

MİKAYİL SADİGZADE, tarafından **PROF. DR. EFENDİ NASİBOĞLU** yönetiminde hazırlanan **"SOSYAL MEDYADA SANAL ZORBALIK İÇERİKLERİNİN TESPİTİ"** başlıklı tez tarafımızdan okunmuş, kapsamı ve niteliği açısından bir Yüksek Lisans tezi olarak kabul edilmiştir.

Prof. Dr. Efendi NASİBOĞLU

Yönetici

Prof. Dr. Urfat NURİYEV

Jüri Üyesi

Dr. Öğr. Üyesi Can ATILGAN

Jüri Üyesi

Prof.Dr. Okan FISTIKOĞLU

Müdür

Fen Bilimleri Enstitüsü

TEŞEKKÜR

Tez çalışmam süresince değerli bilgi ve önerileriyle beni sürekli destekleyen, karşılaştığım problemlerde bana her daim yardımcı olan tez danışmanım Prof. Dr. Efendi Nasiboğlu'na,

Çalıştığım süre boyunca her çıkmaza girdiğimde yaratıcı fikirleriyle yolumu aydınlatan, yüksek lisansı bitirme sürecinde her türlü yardımı yapmaktan kaçınmayan Sayın Omar Bayramlı'ya,

Her zaman olduğu gibi yüksek lisans eğitimimde de benden desteklerini ve anlayışlarını esirgemeyen annem Shafiga Sadıgova'a, babam Sabir Sadıgov'a, kardeşim Maryam Sadıgov'a ve arkadaşlarıma,

Sonsuz minnet ve teşekkürlerimi sunarım.

Mikayil SADIGZADE

SOSYAL MEDYADA SANAL ZORBALIK İÇERİKLERİNİN TESPİTİ

ÖZ

Siber zorbalık, tüm dünyada olduğu gibi Türkiye'de de büyüyen bir sorundur. Şimdiye kadar elde edilen bulgulara göre, Türkiye'de sosyal medya kullananların siber zorbalığa maruz kalma olasılığı %20'i aşmıştır. Siber zorbalık tespiti İngilizcede çok olmasına rağmen Azerbaycan dili ve Türkçede çok az araştırma bulunmaktadır. Bu sorunu ortadan kaldırmak ve tespit etmek için genellikle makine öğrenimi kullanılmaktadır.

Bu çalışmamızda, Azerbaycan dili ve Türkçe metinler üzerinde yapılmış siber zorbalıkları tespit etmek için farklı makine öğrenmesi algoritmaları kullanılmıştır. Çalışmamız, toplam 4400 adet Azerbaycan dili ve Türkçe yazılmış ve sosyal medyadan toplanan cümlelerden oluşan bir veri seti üzerinde makine öğrenimi teknikleri kullanılarak yapılmıştır. Sınıflandırıcıların performansını değerlendirmek için kesinlik (precision), doğruluk (accuracy), duyarlılık (recall) ve F1-skor kullanılmıştır. Çalışmada, kullanılan iki farklı veri setini de ele aldığımızda Türkçe veri setine göre CountVectorizer için %85.98 doğruluk ve %96.94 F1-skor ile Linear SVM modeli en yüksek sonuçlar vermiştir. Yine aynı model ve veri seti ile Tf-IdfVectorizer için en yüksek %85.77 doğruluk ve %97.85 F1-skor sonuçlarına ulaşılmıştır.

Anahtar kelimeler: Siber zorbalık algılama; Makine Öğrenme; N-gram; Tf-Idf; Youtube'da video yorumları

DETECTION OF CYBERBULLYING CONTENT IN SOCIAL MEDIA

ABSTRACT

Cyberbullying is a growing problem in Turkey as well as all over the world. According to the findings obtained so far, the probability of being exposed to cyberbullying for those who use social media in Turkey has exceeded 20%. Although cyberbullying detection is abundant in English, there is very little research in Azerbaijani and Turkish. Machine learning is often used to eliminate and detect this problem.

In this study, different machine learning algorithms were used to detect cyberbullying on Azerbaijani and Turkish texts. Our study was carried out using machine learning techniques on a dataset consisting of 4400 sentences written in Azerbaijani and Turkish and collected from social media. Precision, accuracy, precision (recall) and F1-score were used to evaluate the performance of the classifiers. When we consider the two different datasets used in the study, the Linear SVM model gave the highest results with 85.98% accuracy and 96.94% F1-score for the CountVectorizer compared to the Turkish dataset. Again with the same model and dataset, the highest 85.77% accuracy and 97.85% F1-score results were achieved for Tf-IdfVectorizer.

Keywords: Cyberbullying detection; Machine learning; N-gram; Tf-Idf; Video comments on Youtube

İÇİNDEKİLER

Sayfa

YÜKSEK LİSANS TEZİ SINAV SONUÇ FORMU ii

TEŞEKKÜR iii

ÖZ iv

ABSTRACT v

ŞEKİLLER LİSTESİ viii

TABLOLAR LİSTESİ xi

KISALTMALAR xii

BÖLÜM 1 - GİRİŞ 1

BÖLÜM 2 – LİTERATÜR TARAMASI 5

BÖLÜM 3 – SİBER ZORBALIK 9

3.1 Siber Zorbalığın Türleri 9

3.2 Siber Zorbalık İstatistikaları 10

3.3 Siber Zorbalığın Etkileri 16

3.4 Siber Zorbalığa Tepkiler 19

BÖLÜM 4 – MAKİNE ÖĞRENİMİ YÖNTEMLERİ 21

4.1 Makine Öğrenimi Türleri 23

4.1.1 Denetimli Öğrenme 23

4.1.2 Denetimsiz Öğrenme 24

4.1.3 Pekiştirmeli Öğrenme 25

4.2 K-En Yakın Komşular (KNN) 26

4.3 Çok Terimli Naif Bayes (MNB) 29

4.4 Rastgele Orman (RF) 31

4.5 Karar Ağaçları (DT) 33

4.6 Lineer Destek Vektör Makinesi (SVM) 35

BÖLÜM 5 – KULLANILAN ARAÇLAR 38

5.1 Netlytic 38

5.2 Python Programlama Dili 47

BÖLÜM 6 – UYGULAMA 49

6.1 Değerlendirme Yöntemleri 49

6.2 Özellik Çıkarma 50

6.3 Çapraz Doğrulama 51

6.4 Veri Seti 52

6.4.1 Azerbaycan Veri Seti 54

6.4.2 Türkçe Veri Seti 57

6.5 Veri Ön İşleme 60

BÖLÜM 7 - HESAPLAMA SONUÇLARI VE DEĞERLENDİRMELER 61

BÖLÜM 8 - SONUÇ 64

KAYNAKLAR 65

ŞEKİLLER LİSTESİ

Sayfa

Şekil 1.1 Siber zorbalığa uğrayanların yüzdesi 2
Şekil 1.2 Siber zorbalığa maruz kalan çocukların yüzdesi 2
Şekil 2.1 Siber zorbalık alanında yapılan çalışmaların dağılımı 5
Şekil 3.1 Siber zorbalıkdaki en yaygın taciz türleri 11
Şekil 3.2 Çevrimiçi taciz biçimleri 12
Şekil 3.3 Siber zorbalığın nedenleri 13
Şekil 3.4 Kullanıcıların yaşadığı taciz türleri 14
Şekil 3.5 Sosyal medya uygulamalarında siber zorbalık 15
Şekil 3.6 Çevrimiçi zorbaların tercih ettiği oyun türleri 15
Şekil 3.7 Siber zorbalığın etkileri 17
Şekil 3.8 Siber zorbalığa maruz kalanlar üzerindeki psikolojik etki 17
Şekil 3.9 Çocukların siber zorbalık kurbanı olduğunu bildiren ebeveynler 18
Şekil 3.10 Google'un 2004'ten 2020'ye kadar "siber zorbalık" arama terimiyle ilgili veri göstergeleri 19
Şekil 3.11 Ebeveynlerin siber zorbalıkla ilgili çocuklarıyla konuşması 20
Şekil 4.1 Makine öğrenimi yöntemleri 22
Şekil 4.2 Sınıflandırma ve Regresyon 23
Şekil 4.3 Denetimli Makine öğrenmesi 24
Şekil 4.4 Denetimsiz Makine öğrenmesi 25
Şekil 4.5 Pekiştirmeli Makine öğrenmesi 26
Şekil 4.6 KNN'in sözde kodu 27
Şekil 4.7 KNN örnek çizim 28
Şekil 4.8 KNN adımları 29
Şekil 4.9 MNB'nin sözde kodu 30

Şekil 4.10 RF'in sözde kodu 31
Şekil 4.11 Meta Öğrenenler 32
Şekil 4.12 RF'nın akış şeması sembolleri 32
Şekil 4.13 DT öğrenmenin sözde kodu 34
Şekil 4.14 Olası hiper düzlemler 35
Şekil 4.15 SVM algoritmasının sözde kodu 37
Şekil 5.1 Oturum Açma/Kaydolma 39
Şekil 5.2 Twitter içe aktarma 39
Şekil 5.3 İçe aktarma başarılı 40
Şekil 5.4 İçe aktarma başarısız 40
Şekil 5.5 Youtube içe aktarma 40
Şekil 5.6 RSS içe aktarma 41
Şekil 5.7 Metin dosyası içe aktarma 42
Şekil 5.8 Metin Analizi için kategorilerin oluşturulması/düzenlenmesi 42
Şekil 5.9 Metin Analizi işleme seçenekleri 43
Şekil 5.10 Ağ görselleştirme, öne çıkan ilgi alanları 43
Şekil 5.11 Görünürlük özelliği 44
Şekil 5.12 Düzenleme özellikleri 45
Şekil 5.13 Renk özelliği 46
Şekil 5.14 Otomatik kümeler 47
Şekil 6.1 Önerilen işin akışı 52
Şekil 6.2 Veri açıklaması için geliştirilen web uygulamasının giriş sayfası 53
Şekil 6.3 Veri açıklaması için geliştirilen web uygulamasından bir anlık görüntü ... 53
Şekil 6.4 Kelimelerin görsel olarak yorumlarda ne kadar geçtiği 54
Şekil 6.5 Kelimelerin hangi cümlelerde kullanıldığı 55
Şekil 6.6 Azerbaycan dilinden oluşan veri setindeki kelimeler 55

Şekil 6.7 Yorumların yazılma tarihleri 56
Şekil 6.8 Yorumlarda ortak kelimler kulannanlar arasındaki ilişki 56
Şekil 6.9 Türkçe kelimelerin görsel olarak yorumlarda ne kadar geçtiği 57
Şekil 6.10 Türkçe veri setinde kelimelerin hangi cümlelerde kullanıldığı 58
Şekil 6.11 Türkçe veri setindeki kelimeler 58
Şekil 6.12 Türkçe yorumların yazılma tarihleri 59
Şekil 6.13 Yorumlarda bir birlerine yanıt veren kullanıcılar 60
Şekil 6.14 Metin ön işleme yönteminin sözde kodu 60
Şekil 7.1 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Azerbaycan dili ile hazırlanmış veri seti) 62
Şekil 7.2 Tf-Idfectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Azerbaycan dili ile hazırlanmış veri seti) 62
Şekil 7.3 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Türkçe ile hazırlanmış veri seti) ... 63
Şekil 7.4 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Türkçe ile hazırlanmış veri seti) ... 63

TABLOLAR LİSTESİ

Sayfa

Tablo 2.1 İlgili çalışmaların özeti 8

Tablo 6.1 Karmaşıklık Matrisi 49

Tablo 6.2 Azerbaycan dilinde toplanmış veri kümesindeki bazı cümleler 54

Tablo 6.3 Türkçe yazılmış yorumlardan toplanmış veri kümesindeki bazı cümleler.57

Tablo 7.1 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları(Azerbaycan dili ile hazırlanmış veri seti) 61

Tablo 7.2 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenim Modellerinin Değerlendirme Sonuçları(Azerbaycan dili ile hazırlanmış veri seti) 61

Tablo 7.3 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları(Türkçe ile hazırlanmış veri seti) 61

Tablo 7.4 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları(Türkçe ile hazırlanmış veri seti) 61

KISALTMALAR

ABD	:Amerika Birleşik Devletleri
LGBTQ	:Lesbian, Gay, Bisexual, Transgender (Lezbiyen, Gey, Biseksüel, Transgender)
SVM	:Support Vector Machine (Destek Vektör Makinesi)
LSA	:Latent Semantic Analysis(Gizli Anlamsal Analiz)
MNB	:Multinomial Naive Bayes(Çok terimli Naif Bayes)
KNN	:k-Nearest Neighbors (k-En Yakın Komşular)
LSTM	:Long Short-Term Memory (Uzun Kısa Süreli Bellek)
D-CNN	:Deep Convolutional Neural Network (Derin Evrişimli Sinir Ağı)
RF	:Random Forest (Rastgele Orman)
YSA	:Yapay Sinir Ağı
ANN	:Artificial Neural Network (Yapay sinir ağı)
CNN	:Convolutional neural network (Evrişimli sinir ağı)
RNN	:Recurrent Neural network (Tekrarlayan sinir ağı)
LR	:Logistic Regression (Lojistik Regresyon)
NB	:Naive Bayes (Naif Bayes)
SGD	:Stochastic Gradient Descent (Stokastik Gradyan İniş)
GHSOM	:Growing Hierarchical Self-Organizing Map(Büyüyen Hiyerarşik Kendi Kendini Düzenleyen Harita)
RL	:Reinforcement Learning (Pekiştirmeli Öğrenme)
AUROC	:Area Under the Curve — Receiver Operating Characteristics (Eğri Altındaki Alan — Alıcı Çalışma Karakteristikleri)
BERT	:Bidirectional Encoder Representations from Transformers (Transformatörlerden Çift Yönlü Kodlayıcı Temsilleri)
BOW	:Bag Of Words (Kelime torbası)
RO	:Reverse Osmosis (Ters osmoz)
ROC	:Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)

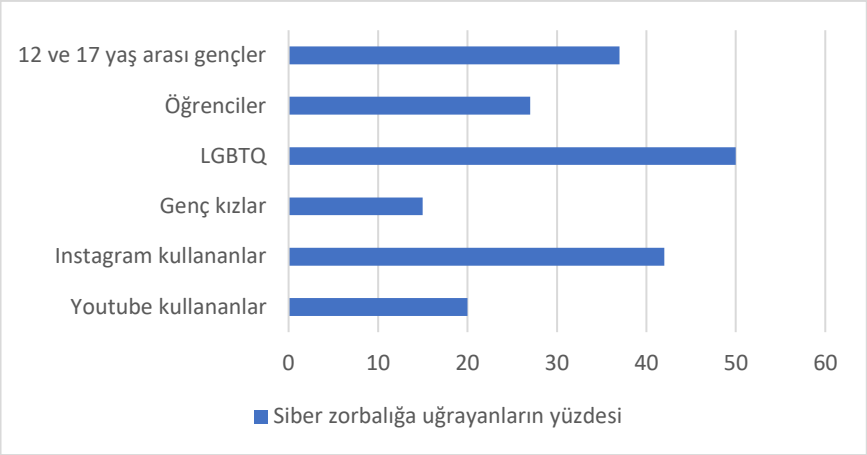
CRT	:Cardiac Resynchronization Therapy (Kardiyak Resenkronizasyon Tedavisi)
DRL	:Deep Reinforcement Learning (Derin Pekiştirmeli Öğrenme)
JRİP	:Repeated Incremental Pruning to Produce Error Reduction (RIPPER) (Hata Azaltma (RIPPER) Üretmek için Tekrarlanan Artımlı Budama)
NVEEE	:National Voices For Equality, Education and Enlightenment (Eşitlik, Eğitim ve Aydınlanma İçin Ulusal Sesler)
İBK	:İnstance-Bases Learning With Parameter k) (k Parametresi ile Örnek Tabanlı Öğrenme)
SMO	:Sequential Minimal Optimization (Sıralı Minimum Optimizasyon)
DT	:Decision Trees (Karar Ağaçları)
DAG	:Directed Acyclic Charts (Yönlendirilmiş Döngüsel Olmayan Grafiklerin)
CPA	:Cutting Plane Algorithm (Kesme Düzlemi Algoritması)
RBF	:Radial Based Function (Radyal Tabanlı Fonksiyon)

BÖLÜM 1
GİRİŞ

Sosyal ağın gelişmesiyle birlikte dünyada siber zorbalık sayısı artmaya başladı. Siber zorbalık tespiti, özellikle Makine Öğrenimi kullanımı ile birlikte de artan bir ilgi görmektedir. Siber zorbalığın artmasının nedeni, zorbaların internetteki zorbalığın tespit edilememesi veya tespit edilse dahi yasal yaptırım uygulanmayacağını düşünmeleridir. Bu tür siber zorbalık suçları, insanlar üzerinde psikolojik baskı oluşturarak gelecekte hayatlarında ruhsal izler bırakmaktadır. Siber zorbalığı zamanında tespit etmek ve karşı koymak çok zordur. Dünya çapında çok sayıda çocuk cinsel ayrımcılığa ve cinsiyete dayalı şiddete, fiziksel cezaya, savaşa ve diğer şiddet biçimlerine maruz kalmaktadır. Ayrıca birçoğu okul bahçesinde akranları tarafından çete şiddeti, silahlı saldırı, tecavüz, taciz, cinsel ve toplumsal cinsiyete dayalı şiddete maruz kalıyor. Şiddetin yeni bir türü olarak özellikle cep telefonları, bilgisayarlar, web siteleri ve sosyal paylaşım siteleri aracılığıyla gerçekleşen siber zorbalık gibi olaylar çocukların yaşamlarını olumsuz etkilemektedir (Altay ve Betül, 2017). Bilgi teknolojilerinin gelişmesi ve iletişim araçlarının hızlı bir şekilde kullanıcıların hayatına girmesi, sosyal medya platformlarının, web sitelerinin ve paylaşım ağlarının gelişiminin ve çeşitliliğinin önünü açmıştır (Altay ve Bilal, 2018).

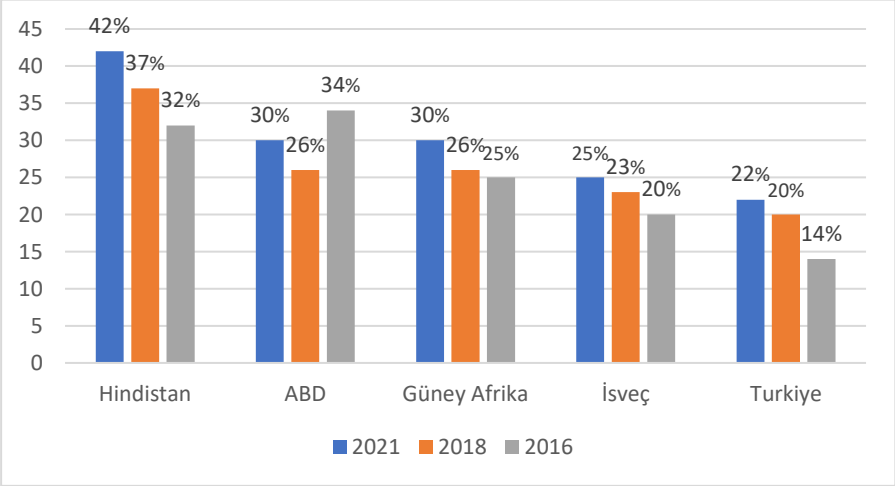
İnternetin dünyada ve toplumumuzda da hızla gelişmesiyle birlikte siber zorbalık kavramı hayatımızı daha fazla etkilemeye başlamıştır. Özellikle sosyal medyayı kullanan herkes siber zorbalığa maruz kalmış veya etkisinde olmuştur. Siber zorbalığa maruz kalan kişilerin sıklıkla psikolojik sorunlar yaşadığı görülmektedir. Bu nedenle siber zorbalık göz ardı edilmemelidir. Siber zorbalıkla mücadele için birçok ülkede yeni açılan kurumlar vardır. Bu kurumlara örnek olarak ABD'de Ulusal Suç Önleme Konseyini ve Türkiye'de Bilgi Teknolojileri ve İletişim Kurumu'nu gösterebiliriz (Shukan ve ark., 2019).

Şekil 1.1'de siber zorbalıkla ilgili araştırmalar sonucunda bazı istatistiksel sonuçları verilmiştir. Bu sonuçlara göre, LGBTQ bireylerin sosyal medyada siber zorbalığa uğrama olasılığı daha yüksektir (Ravichandran, 2017).



Şekil 1.1 Siber zorbalığa uğrayanların yüzdesi (Ravichandran, 2017)

Çocuklar ve ergenler için tehlikeli olabilen siber zorbalık vakalarının dünya çapında giderek büyüyen bir sosyal sorun haline gelmesi, eğitimcileri, akademisyenleri ve yasa koruyucularını siber zorbalığın risklerini, yaygınlığını ve sonuçlarını incelemeye ve bu konuda çözüm üretmeye yöneltmiştir (Şekil 1.2). Özellikle son yıllarda giderek daha fazla ülkenin bu konuda eğitim politikaları oluşturduğu, araştırma ve projeleri desteklediği ve bu konuda çalışan sivil toplum kuruluşları, akademisyenler, eğitimciler ve ailelerle ortaklıklar kurduğu gözlemlenmektedir (Akca ve İdil, 2017).



Şekil 1.2 Siber zorbalığa maruz kalan çocukların yüzdesi

Bilgi ve iletişim teknolojilerinin getirdiği bu kolaylıkların yanı sıra aşırı ve kontrolsüz kullanımın çeşitli sorunlara yol açtığı açıktır. Bunlara pornografi, bilgi ve

iletişim teknolojilerinin aşırı kullanımı, öğrencilerin akademik performanslarına müdahale, sanal ortamdaki herkesle kişisel bilgileri paylaşma, zararlı ilaçları satın almak için sanal ortamı kullanma gibi sorunlar dahildir. Örneğin, birçok kişinin üye olduğu bir sosyal ağ olan Facebook'ta, bağımlıların buluşup sosyalleştiği, sentetik uyuşturucu olarak adlandırılan ve Türkiye'de sıklıkla gündemde olan bonzai sattıkları birçok sanal grup bulunmaktadır (Eroğlu, 2014).

Çeşitli ülkelerdeki akademik araştırmalar, siber zorbalığa uğrayan çocukların algıları ve farkındalıkları, siber zorbalığın onlar üzerindeki etkileri ve sorunla nasıl başa çıkmaya çalıştıkları üzerine odaklanmıştır. Siber zorbalığın doğası, siber zorbalığın davranışları ve bu olgunun sonuçları ülkeden ülkeye farklılık göstermekle birlikte, yeni medyanın yarattığı kültüre bağlı olarak birçok ortak noktanın olduğu açıktır (Akca ve İdil, 2017).

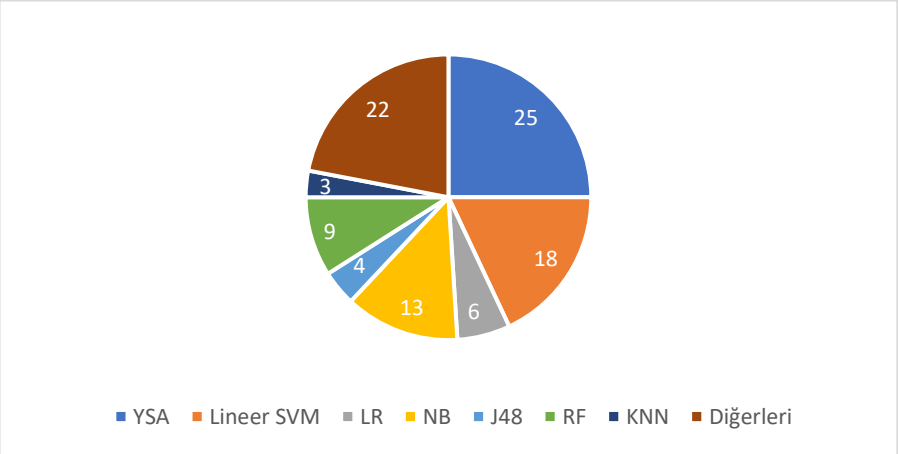
Siber zorbalık üzerine yapılan birçok araştırma sonucunda, siber zorbalık ve mağduriyetin insanların sosyal, akademik ve duygusal yaşamları üzerinde son derece kötü bir etkisinin olduğu gösterilmiştir (Eroğlu ve Neşe, 2015). İlişkisel tarama modeli kullanılarak yürütülen (Polat ve Seda 2016) çalışmasında, ergenlerin siber zorbalık ve mağduriyetini etkilediğine inanılan sosyo-demografik değişkenlerin yanı sıra internet kullanımı, öfke ve öfkenin ifade biçimleri, stres yönetimi ile ilgili değişkenlere ilişkin veriler kullanılmıştır. Davranış değişkenleri toplanmış ve sistematik ve istatistiksel olarak değerlendirilmiş, analiz edilmiş ve elde edilen veriler betimsel bir yöntemle açıklanmıştır. (Ünver ve Zihni, 2017) çalışması, ortaokul öğrencilerin yaptığı siber zorbalığın, problemli internet kullanımı ve riskli çevrimiçi davranışlarla ilgilisinin ne kadar olduğunu belirlemek için yapılmıştır. Ayrıca buna yaş, cinsiyet, günlük internet kullanımı ve tercih edilen internet siteleri gibi farklı faktörlerinde etkisi vardır.

Siber zorbalık tüm dünyada olduğu gibi Türkiye'de de büyüyen bir sorundur. Yapılan araştırmalara göre son zamanlarda sosyal medya kullanan kitlenin artmasıyla birlikte Türkiye'de sosyal medya kullananların siber zorbalığa maruz kalma olasılığı %20'i aşmıştır. Bu bağlamda literatürlerde siber zorbalık tespiti İngilizce metinler üzerinde çok olmakla birlikte Azerbaycan dili ve Türkçe metinler üzerinde az sayıda çalışma bulunmaktadır. Ayrıca Türkçe çalışmalarda sadece sınırlı sayıda algoritma ve yöntem kullanılmaktadır.

Bu çalışmamızda Netlytic sitesinin özelliği kullanılarak Youtube'da Azerbaycan dili ve Türkçe yazılmış yorumlardan iki farklı veri seti oluşturularak kullanılmıştır. İlk olarak, ilgili veri kümesine çeşitli ön işleme adımları uygulandı. Ön işleme adımlarından sonra veri setindeki metinler üzerinde öznitelik olarak için stop word, n-gram ve Tf-Idf yaklaşımları uygulandı. Daha sonra geliştirilen farklı makine öğrenme algoritmaları, siber zorbalık içeren bu mesajları tespit etmek için test edildi. Son olarak geliştirilen modellerin sonuçları tahmin başarısı açısından karşılaştırıldı. Sonuç olarak çalışmada, kullanılan iki farklı veri setini de ele aldığımızda Türkçe veri setine göre CountVectorizer için %85.98 doğruluk ve %96.94 F1-skor ile Linear SVM modeli en yüksek sonuçlar vermiştir. Yine aynı model ve veri seti ile Tf-IdfVectorizer için en yüksek %85.77 doğruluk ve %97.85 F1-skor sonuçlarına ulaşılmıştır.

BÖLÜM 2
LİTERATÜR TARAMASI

Literatürde siber zorbalığın tespiti ile ilgili farklı yaklaşımlar bulunmaktadır. Bu alanda yapılan çalışmaların dağılımı Şekil 2.1'de görülmektedir. Bu alandaki bir başka (Reynolds ve ark., 2011) çalışmada, zorbalığın çok daha yüksek düzeyde gerçekleştiği bir soru-cevap sitesi olan Formspring.me'nin içeriğinden anketden veri seti oluşturulmuştur. Bu çalışmada elde edilen büyük miktardaki verinin etiketlenmesi işlemi Amazon Mekanik Türk web servisi kullanılarak gerçekleştirilmiştir. Oluşturulan veri kümesine C4.5 karar ağacı yönteminin uygulanması, siber zorbalık içeren metinlerin %78,5 doğrulukla tespit edildiğini ortaya çıkarmıştır.



Şekil 2.1 Siber zorbalık alanında yapılan çalışmaların dağılımı

(Arroyo-Fernández ve ark., 2018), saldırganlık düzeylerinin her birini ayrı ayrı tahmin etmek için birkaç iyi bilinen eğitim makinesi (sınıflandırıcılar) ile birlikte bir dizi vektör uzayı modelleme tekniğini test etmiştir. Vektör uzayı modelleme teknikleri, Tf-Idf vektörlerinin gizli anlamsal analizini farklı boyutlarda LSA içermektedir. Hem Tf-Idf, hem de LSA'ya göre karakter veya kelimeler n-gram özellikleri kullanılarak hesaplanmıştır.

(Sugandhi ve ark., 2016) çalışmasında, hangi sınıflandırıcının en iyi doğruluğu verdiğini görmek için veri seti üzerinde farklı sınıflandırıcılar test edilmiştir. Bunu yapmak için, verinin %80'ini eğitim için ve %20'sini test için kullanarak etiketlenmiş verileri iki kümeye ayrılmıştır. Çalışmada, 393 zorbalık olan ve 2886 zorbalık olmayan

mesajlardan oluşan etiketli verileri, Lineer SVM, MNB ve KNN sınıflandırıcıları kullanılarak diğer çalışmalarla karşılaştırılmıştır.

(Novalita ve ark., 2019), hem siber zorbalık hem de siber zorbalık olmayan tweet'leri ele almışlar. Rastgele orman sınıflandırma testinin sonuçlarına bakıldığında, sistemin siber zorbalık tweet'lerini en iyi F1-skoru 90% olan başarıyla bulduğu tespit edilmiştir. (Isa ve Ashianti, 2017), siber zorbalık sınıflandırması için en uygun SVM çekirdeğini araştırmışlar. Kullanılan veriler doğrusal olarak ayrılabilir olmadığından, en iyi modelin %97,11 doğrulukla çok çekirdekli bir model olduğunu bulmuşlardır.

Türkçe ile ilgilenen (Pervan ve Keleş, 2019) makalesinde, LSTM, D-CNN, RF gibi YSA modelleri kullanılmıştır. N11.com web sitesinden toplanan veriler üzerinde çalışmalar yapılmıştır. (Agrawal ve Awekar 2018) çalışmasında, ANN, CNN, RNN, LSTM gibi yapay sinir ağı modelleri kullanılmış ve bunların arasında LSTM en yüksek doğruluğu göstermiştir. Bir sosyal medya aracı olan Twitter'dan toplanan veriler üzerinde çalışmalar yapılmıştır. (Yuvaraj ve ark., 2021) çalışmasında, İngilizce metinleri inceleyen bu makalede sadece CNN algoritması kullanılmış ve Formspring.me'den toplanan veriler üzerinde çalışmalar yapılmıştır. Çalışmada, en yüksek puan F1-skorda alınmışdır.

(Zhao ve ark., 2020) ve (Rezvani ve ark., 2020) makalelerinde Twitterden toplanan İngilizce farklı veriler ele alınarak LSTM algoritması üzerinde çalışmalar yapılmıştır. Çalışmaların herikisinde, LSTM en yüksek puanını F1-skorda göstermiştir.

(Rachid ve ark., 2020) makalesinde İngilizce metinlerden oluşturulan veriler üzerinde CNN tekniği kullanılmıştır. Twitter'dan toplanan veriler üzerinde araştırmalar yapılmış ve çalışmada, CNN için en yüksek F1-skoru bulunmuştur. Bir diğer (Pawar ve Raje, 2019) çalışmasında veriler İngilizce metinlerden oluşturulmuş olup NB, LR, SGD makine öğrenme teknikleri kullanılmıştır. Formspring.me'den toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, LR en yüksek doğruluğu ve F1-skor sonuçlarını göstermiştir. İngilizce metin üzerinde çalışılan (Di Capua ve ark., 2016) makalesinde GHSOM tekniği kullanılmıştır. Burada Formspring.me, Youtube, Twitter'dan toplanan veriler üzerinde çalışmalar yapılmıştır.

(Aind ve ark., 2020) makalesinde makine öğreniminin bir alanı olan RL tekniği kullanılmıştır. Yorumlardan toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, pekiştirmeli öğrenme en yüksek doğruluk sonucunu göstermiştir. (Soni ve

Singh, 2018) çalışmasında, LR tekniği kullanılmıştır. Gönderilerden toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, LR en yüksek sonucunu AUROC 'da göstermiştir.

(Haidar ve ark., 2019) makalesin'de makine öğrenimi algoritmalarının kararlılığını ve doğruluğunu geliştirmek için kullanılan bir meta-algoritma tekniği olan torbalama kullanılmıştır. Twitter'dan toplanan veriler üzerinde araştırmalar yapılmış ve çalışmada, SVM en yüksek puanı F1-skorda göstermiştir. (Paul ve Saha, 2020) makalesinde İngilizce olarak BERT makine öğrenimi tekniği kullanılmıştır. Wikipedia, Formspring.me ve Twitter'dan toplanan veriler üzerinde çalışmalar yapılmıştır. Çalışmada, BERT en yüksek sonucunu F1-skorda göstermiştir.

Türkçe metinler üzerinde çalışılan (Karcioğlu ve Aydin, 2019) makalesinde BOW, Lineer SVM, LR gibi algoritmalar kullanılmıştır. Twitter'dan toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, SVM en yüksek sonucunu doğrulukda göstermiştir. (Altay ve Alatas, 2018) çalışmasına baktığımızda nispeten daha az kullanılan RO, Çok katmanlı algılayıcı, J48, SVM gibi teknikler kullanılmıştır. Formspring.me web sitesinden toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, en yüksek sonuçları F1-skor ve ROC'da göstermiştir.

(Kumari ve ark., 2021) makalesinde RF kullanılmış ve görüntülü konuşmalarda yazılan yorumlardan toplanan veriler üzerinde çalışmalar yapılmıştır. Araştırmada, RF en yüksek puanını F1-skorda göstermiştir. (Hani ve ark., 2019) makalesinde SVM, Sinir ağı teknikleri kullanılmıştır. E-posta ve buna ek olarak anlık mesajlaşmadan toplanan veriler üzerinde çalışmalar yapılmış ve ayrıca çalışmada, SVM en yüksek sonucunu F1-skorda göstermiştir. (Song ve Song, 2021) makalesinde CRT algoritması kullanılmıştır ve buzzes'tan (telefon görüşmeleri) toplanan veriler üzerinde çalışmalar yapılmış ve çalışmada, CRT en yüksek sonucu doğrulukda göstermiştir.

Tablo 2.1'de siber zorbalık üzerinde yapılmış literatür taramasının sonucunda elde edilen bilgiler verilmiştir.

Tablo 2.1 İlgili çalışmaların özeti

Çalışmalar	Kullanılan yöntemler	Veriler	Sonuçlar	Çalışılan Dil
Pervan ve Keleş, 2019	LSTM, D-CNN, RF	N11 internet sitesi	LSTM Doğruluk=%94,21	Türkçe
Agrawal ve Awekar, 2018	CNN	Formspring.me	CNN F1-skor = %93	İngilizce
Yuvaraj ve ark., 2021	ANN, DRL	Twitter	ANN Doğruluk=%80,7	İngilizce
Zhao ve ark., 2020	Dikkat ağı ile LSTM	Twitter	LSTM F1-skor =%86,3	İngilizce
Van ve ark., 2020	CNN	Gönderiler	CNN F1-skor =%88,5	İngilizce
Rezvani ve ark., 2020	LSTM	Instagram, Twitter	LSTM F1-skor =%92,8	İngilizce
Rachid ve ark., 2020	RNN, CNN	Yorumlar	RNN F1-skor = %84	İngilizce
Pawar ve Raje, 2019	NB, LR, SGD	Formspring.me	LR Doğruluk=%90,24, F1-skor = %90,56	İngilizce
Di Capua ve ark., 2016	GHSOM ağ algoritması	Formspring.me, Youtube, Twitter	GHSOM Doğruluk = %73, F1-skor = %74	İngilizce
Aind ve ark., 2020	RL	Yorumlar	RL Doğruluk=%89	İngilizce
Soni ve Singh, 2018	LR	Gönderiler	LR AUROC = %83,4	İngilizce
Haidar ve ark., 2019	SVM	Twitter	SVM F1-skor =%90,5	İngilizce
Paul ve Saha, 2020	BERT	Twitter Wikipedia Formspring.me	BERT F1-skor = %94	İngilizce
Karcioğlu ve Aydin, 2019	Lineer SVM, LR	Twitter	Lineer SVM Doğruluk =%65,62, LR Anma ortalaması =%60,99	İngilizce
Altay ve Alatas, 2018	RO, J48, SVM	Formspring.me	RO F1-Skor=%83,2; ROC=%92	İngilizce
Kumari ve ark., 2021	RF	Yorum içeren resimler	RF F1-Skor = %74	İngilizce
Hani ve ark., 2019	SVM, Sinir Ağı	E-posta ve anlık mesajlaşma	SVM F1-Skor =%89,8	İngilizce
Song ve Song, 2021	CRT	Telefon konuşmaları	CRT Doğruluk=%74,5	İngilizce
Reynolds ve ark., 2011	J48, JRIP, IBK ve SMO	Formspring.me	J48 Doğruluk=%81,7	İngilizce
Bozyiğit ve ark., 2019	YSA modelleri	Twitter	YSA2 F1-skor = %91	İngilizce
Bozyiğit ve ark., 2018	SVM, MNB, RF, KNN, NB, Torbalama, C4.5	Twitter	SVM F1-skor = %91	Türkçe
Nasiboglu ve Gencer, 2022	Derin Öğrenme Modelleri	İngilizce haber makaleleri	Spacy NER F1-skor = %97	İngilizce
Sadigzade ve Nasibov, 2022	RF, MNB, DT, Lineer SVM	Twitter	Lineer SVM F1-skor = %99	İngilizce

BÖLÜM 3
SİBER ZORBALIK

3.1 Siber Zorbalığın Türleri

Genel olarak siber zorbalığın insanlar üzerindeki etkisine göre farklılıklar göstermesi nedeniyle literatürde birçok siber zorbalık türü tartışılmaktadır (Weru ve ark., 2017). En yaygın siber zorbalık türleri şunlardır:

- **Taciz**

Taciz, siber zorbalığa uğrayan kişinin karşı tarafdan hakaret ve küfür içeren mesajların iletilmesidir (Akca ve Sayımer, 2017; Öztürk, 2017; Yazğili ve Baykara 2021).

- **İfşa**

İfşa (Outing), siber zorbalığa uğrayan kişinin kişisel bilgilerinin, kimsenin görmemesi gereken bilgilerin veya fotoğraflarının herkesin görebileceği şekilde telefon, sosyal medya gibi ortamlarda paylaşılmasıdır (Akca ve Sayımer, 2017; Öztürk, 2017; Yazğili ve Baykara, 2021).

- **Dışlama**

Siber zorbalığa maruz kalan kişinin whatsapp ve instagram gibi gruplardan kasıtlı olarak çıkarılması veya kısıtlanmasıdır (Öztürk, 2017; Yazğili ve Baykara, 2021).

- **Taklit etme**

Kimliğe bürünme veya kimlik avı, siber zorbalığa uğrayan kişi veya kişileri incitmek, toplumda ve arkadaşlarının huzurunda kötü görünmelerini sağlamak amacıyla siber zorbalığa uğrayan kişinin veya başkasının kılığına girerek çeşitli mesajlar göndermek veya paylaşmaktır (Yazğili ve Baykara, 2021).

- **Saldırı**

Saldırı, siber zorbalığa uğrayan kişi hakkında çeşitli yalan, kötü ve dedikodu ibareleri veya bilgilerin diğer üçüncü ve dördüncü kişilere gönderilmesidir (Öztürk, 2017; Yazğili ve Baykara, 2021).

- **Öfke**

Alevli kışkırtma olarak da bilinen bu eylem, sosyal medyayı kullanan kişiler arasındaki ilişkilerde genellikle düşük gösterici hareketler ve düşmanca bir etki yaratmaktadır (Akca ve Sayımer, 2017; Öztürk, 2017; Yazğili ve Baykara, 2021).

- **Sahte profil oluşturma**

Sahte profil oluşturma (Create fake profile), web siteleri ve çevrimiçi gruplar gibi sahte platformlar üzerinden siber zorbalığa uğrayan kişilere utanç verici eylemlerde bulunmak ve kötü niyetli haberlerle güvenliklerini azaltmak için yapılan bir eylemdir (Öztürk, 2017; Yazğili ve Baykara, 2021).

- **Akran siber takip**

Akran siber tacizi, siber zorbalığın başka bir biçimi olan tacizin daha şiddetli bir çeşididir. Bu türün tacizden farkı, siber zorbalığa uğrayan kişinin farklı iletişim teknolojileri aracılığıyla sürekli takip edilmesi, tehdit edilmesi, sindirilmesi ve kendini sürekli tehlikede hissetmesidir (Öztürk, 2017; Yazğili ve Baykara, 2021).

3.2 Siber Zorbalık İstatistikaları

Siber zorbalık, son zamanlarda ne yazık ki tüm dünyada hızlı şekilde büyüyen bir sorun haline gelmeye başladı . İnsanların alıştığı normal zorbalıklardan pek farkı olmayan bu türün tek farkı, çevrimiçi gerçekleşmesidir. Bununla ilgili olarak aşağıda, gitgide bu artan sorunun kapsamını ve neden etkili bir çözüme ihtiyacı olduğunu gösteren bir dizi siber zorbalık istatistiği bulunmaktadır.

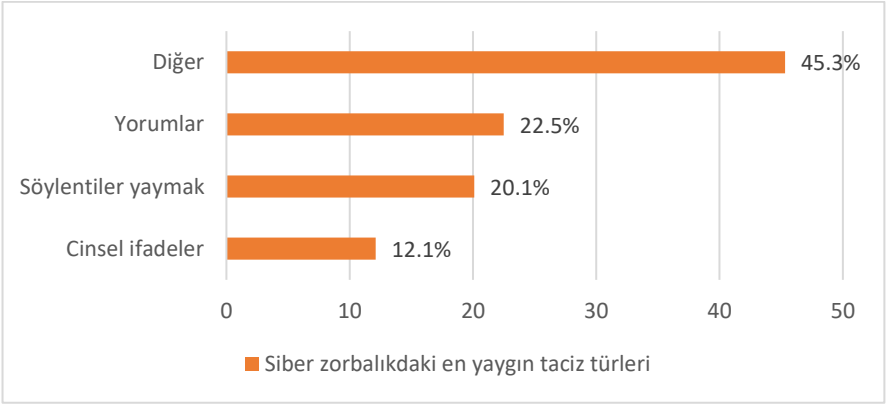
Yapılmış araştırmalara göre siber zorbalık hakkında genel istatistikler (Justin ve Sameer, 2022):

- Halihazırda yaygın olarak kullanılan sosyal medyalar platformlarında yorum yazmak artık herkes için sıradan hal almıştır ve siber zorbalıkların %22.5 yazılan yorumlar içermektedir.
- Herhangi bir kişinin paylaştığı kişisel durumunu veya fotoğrafını onun izni olmadan ekran görüntüsünü alarak paylaşanların oranı %35 dir.

- Özellikle sosyal medyada daha fazla zaman geçiren gençlerin %61'i siber zorbalığa uğrama nedeninin onların fiziksel ve dış görünüşleri yüzünden olduğunu söylüyorlar.
- Yapılan araştırmalar sonucunda siber zorbalığın bir türü olan tacize maruz kalanların %56'sı Facebook kullanıcılarının olduğu gözlemlenmiştir.
- Çocuk yaşda teknoloji ve sosyal medyayla buluşan gençlerin %70'nin daha 18 yaşına gelmeden siber zorbalığa maruz kalmaktadır.

Siber zorbalık yapan kişiler karşı tarafı taciz edici, küçük düşürücü, tehdit edici ve ayrıca utandırıcı eylemlerde bulunurlar. Yapılan bu zorbalıklar genel olarak görünüm, zeka, ırk veya cinsellik üzerinde olup yorumlarla yapılmaktadır. Yazılan yorumlara ve mesajlara göre gözlemlenmiş istatistikler aşağıda bilgilerde ayrıntılı şekilde belirtilmiştir.

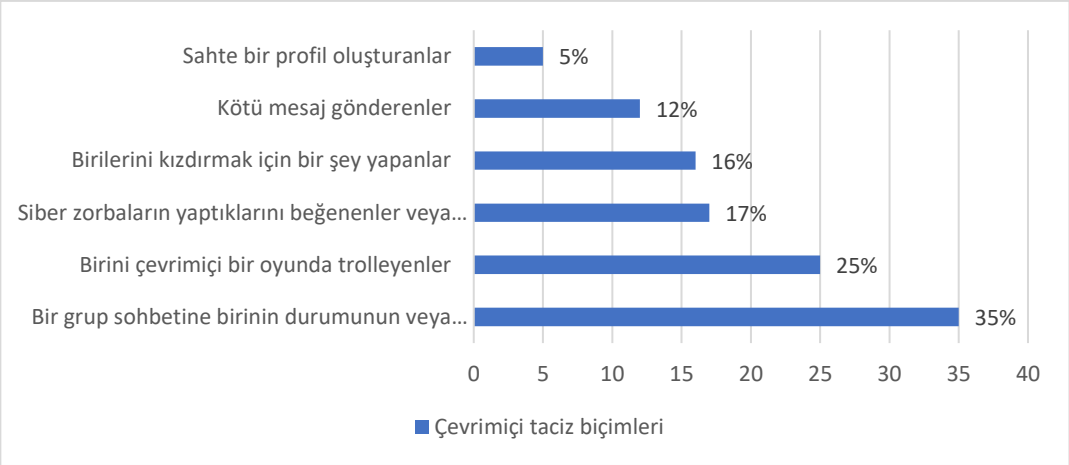
Amerikada yapılmış araştırmalara göre siber zorbalığa uğrayan öğrenciler daha çok zorbalığın taciz türü olan yorumlarla karşı karşıya kalmaktadır. Bu araştırmalar sonucunda ise genel olarak bu taciz türleri yüzde oranlarıda belirtilerek 4 çeşite ayrılmıştır. Bu oranlar Şekil 3.1'de gösterilmiştir.



Şekil 3.1 Siber zorbalıkdaki en yaygın taciz türleri

Siber zorbalığa maruz kalan mağdurlardan alınan bilgilere göre onlara bir anlık öfkeyle taciz türlü mesaj gönderen kişilerin %64'nü normal hayatda da yüz yüze tanıdıklarını söylüyorlar. Siber zorbalar zorbalık yaptığı kişileri birebir tanımaklarına rağmen onların kişisel fotoğrafıyla alay ederek karşı tarafı küçük düşürecek, üzecek ve utandıracak eylemlere başvuruyorlar. Bu tarz eylemler çevrimiçi oyunlarda da sık

sık görülmektedir. Yapılmış bazı araştırmalara göre çevrimiçi taciz biçimlerine göre 6 çeşite ayrılmış ve onların istatistikleri aşağıda Şekil 3.2'de belirtilmiştir (Amy, 2022):

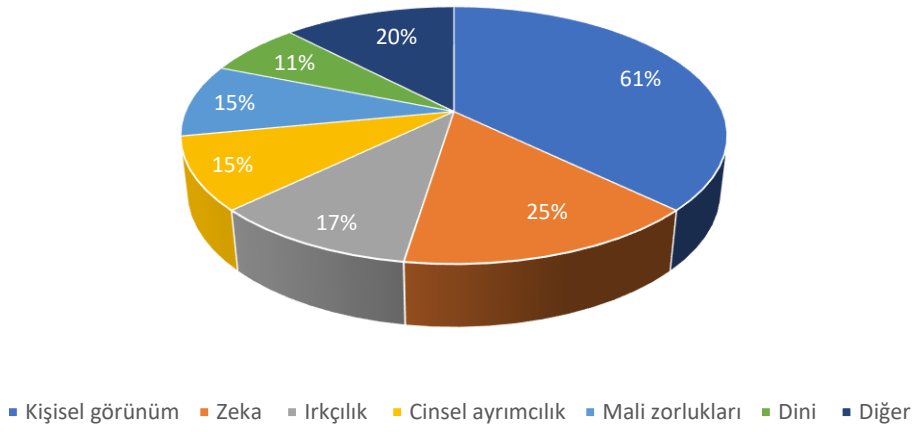


Şekil 3.2 Çevrimiçi taciz biçimleri

(Amanda, 2022)'nın yaptığı araştırmalara göre sosyal medya kullanan 15-17 yaşlarındaki gençlerin %18'i ve yaşları daha az olan çocukların %11'i zorbaların onların şahsi mesajlarını herkesin göre biliceği şekilde paylaşmasını yada başkalarına göndermelerini deneyimlemiştirler. (Florida Atlantic University, 2022)'da kaynağında belirtilen bilgilerde Florida Atlantik Üniversitesinin on yılda yaptığı araştırmalardan sonra toplam 20000 kişiden oluşan bir anketde okul öğrencilerinin %70'nin birisi tarafından dedikodusunun yaydığını söylediler.

(Reportlinker, 2022)'nın araştırmalarında siber zorbalığa uğrayan çocukların normal çocuklara göre kimlik sahtekarlığı kurban olma oranın 9 kat daha fazla olduğunu açıklamışdır.

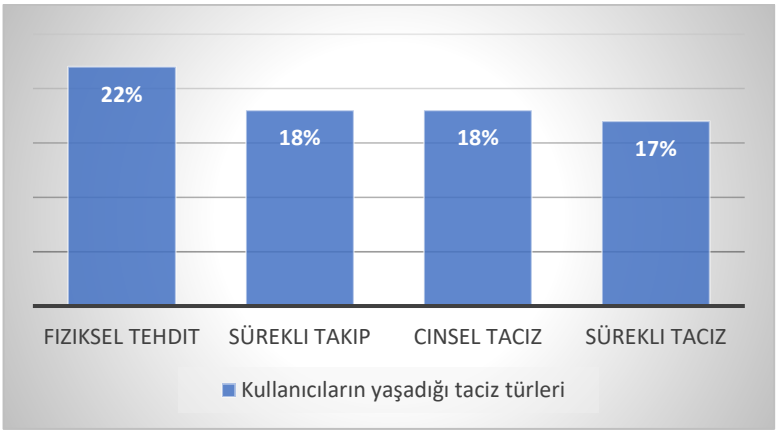
(Ditch the Label, 2022)'ın kaynağının yaptığı anketlerden sonra siber zorbaların hedefleri arasında malesef engelli ve akıl hastalarıda yer almaktadır. Bunun nedeni ise engelliler ve akıl hastaları bu tarz siber zorbalıklara genellikle karşı koyamazlar ve bu yüzdende zorbalar için onları taciz etmek daha kolay gelir. Bir başka (NVEEE, 2022) kaynağın topladığı verilere göre siber zorbalık yapanların sık kullandığı araçlar Şekil 3.3'de detaylı bir şekilde istatistikleri ile birlikte verilmiştir.



Şekil 3.3 Siber zorbalığın nedenleri

(Joseph, 2022)'da yapılan araştırmalara göre 2007-2016 seneleri aralığında uzun bir süre gençlerin siber zorbalığa maruz kalma oranı %32'ler civarında pek değişiklik göstermemiştir. En son 2019 yapılan araştırmaların sonucunda çoğunluğunun çocuklar, öğrenciler, kızlar ve LGBT topluluğu üyelerinin teşkil etdiyi kişilerin %43'nün siber zorbalığa maruz kaldığı gözlemlenmiştir. (https://netsanity.net/)'de siber zorbalıkla ilgili yaptığı çalışmaların sonucunda LGBT üyesi gençlerin %12'den fazlası sürekli olarak siber zorbalıkla karşılaşdığı ve bu mağdurların %35'i çevrimiçi tehditlere, %58'si ise zorbaların nefret söylemleriyle uğraşmak zorunda kaldıkları tespit edilmiştir.

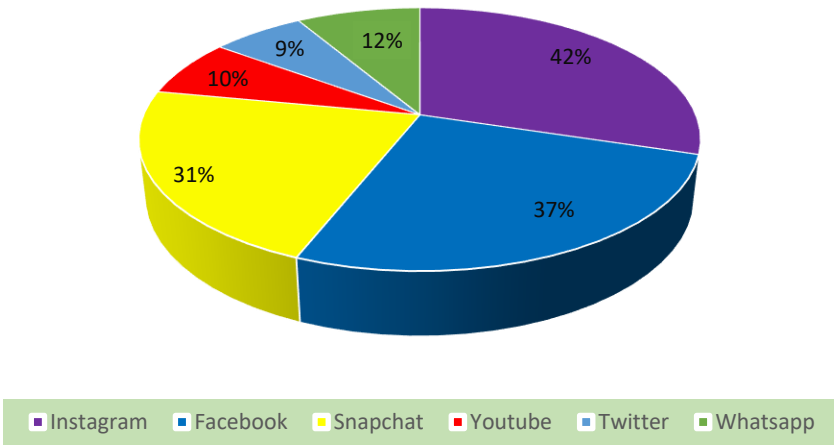
Kızlar erkeklere göre daha çok zorbalığa maruz kalmaktdırlar. Kayıtlarda (Do Something, 2022)'nın yaptığı araştırmalardan sonra 12-17 yaş aralığında erkeklerin %6'sı, kızların ise %15'i siber zorbalığa uğradığı bulunmaktadır. Daha büyük kızlarda ise bu oranın %41'e kadar yükseldiği tespit edilmiştir. (Sam, 2022)'nın çalışmasında gençlerin %60'ı ve 12-17 aralığındaki çocukların ise %37'si zorbalığa tanık kalmasına rağmen bunu durdurmamışdırlar. Bunun nedeni olarak onların mağdur olmak ve kötü durumda kalmamak için müdahele etmedikleri bilinmektedir. Yetişkinler arasında siber zorbalığa maruz kalma oranına bakıldığında bu oranın %53'den fazla olduğu farkedilmiş ve onlara karşı yapılan zorbalıkların çocuk ve gençlere göre farklılık gösterdiği tespit edilmiştir. Tespit edilen bu zorbalıklar fiziksel tehdit, cinsel taciz, sürekli takip ve sürekli taciz gibi ciddi çevrimiçi taciz türleridir ve onların oranları Şekil 3.4'de gösterilmiştir.



Şekil 3.4 Kullanıcıların yaşadığı taciz türleri

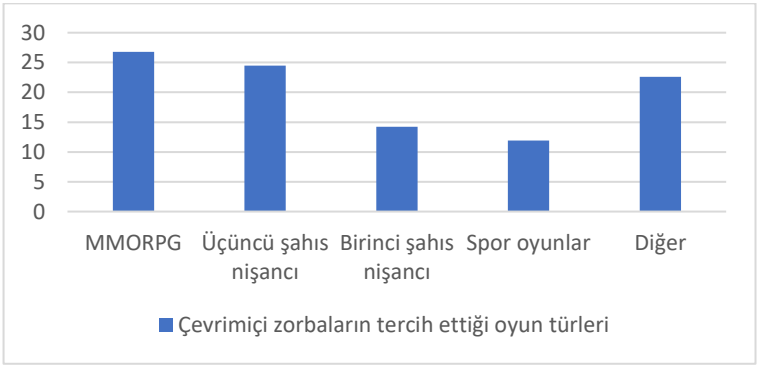
(Do Something, 2022)'nın yaptığı araştırmalara göre özellikle gençlerin %95'i çevrimiçidirler ve onlarında çoğu bunu mobil cihazlardan yapmaktadırlar. Buradan da gelinen sonuç mobil cihazların daha yaygın kullanılması siber zorbalığın artmasına neden olmasıdır. İnsanların siber zorbalığa en fazla nerede maruz kaldığına baktığımızda, özellikle tüm dünyada yaygın bir şekilde kullanılan platformalar karşımıza çıkacaktır. Bu platformalar siber zorbalıkların karşısını almak için çeşitli çalışmalar yapsalar bile halada tamamen arındırılmış değillerdir. (Enough Is Enough, 2022)'yaptığı araştırmalar sonucunda İnstagram kullanıcılarının diğer platformalara göre daha çok zorbalığa uğraması gözlemlenmiştir. İnstagram'dan sonra Facebook ve Snapchat'in bu konuda üst sıralarda yer almaktadırlar. Şuan dünyada ciddi sayda kullanıcısı olan Youtube'da bu rakamlar diğerlerine göre nispeten daha düşük olup, yapılan anketlerin sonucunda her on kişiden sadece birinin siber zorbalığa maruz kaldığı elde edilmiştir. Şekil 3.5'de sosyal medya platformaların siber zorbalık oranları verilmiştir.

(Joseph, 2022) kaynağından alınan verilere göre internet trolleri daha çok sosyal medyada aktif halde gözükmektedirler. Buradaki istatistikalara göre zorbaların yaptığı trollerin %38'i sosyal medyadaki insanlar üzerinde yazılmıştır. Ayrıca %23'ü Youtube ve diğer benzer platformalarda paylaşılmış videolar altında yazılmış yorumlardan ibaretdir.



Şekil 3.5 Sosyal medya uygulamalarında siber zorbalık

(Telenor, 2022)'nın ebeveynlerle yaptığı anketlerin sonucunda onlar çevrimiçi oyun oynayan çocuklarının %79'nun fiziksel tehditler aldığı belirlenirken, %41'nin de cinsiyetçi veya ırkçı sözler aldığını söylediler. Bir başka kaynak (https://cyberbullying.org/)'nın yaptığı ankete göre MMORPG'leri oynayan oyuncuların siber zorbalık yapma oranı daha yüksektir. MMORPG ve diğer oyunların siber zorbalık oranı Şekil 3.6'da verilmiştir.



Şekil 3.6 Çevrimiçi zorbaların tercih ettiği oyun türleri

İnternetde blog yazarların yaklaşık olarak 500 milyondan fazla olması nedeniyle onlarında siber zorbalıkla tanışmaları şaşırtıcı değildir. Yazılan blogların yorum kısmının insanların sorular veya konu üzerinde tartışma yapmaları için olmasına rağmen insanlar kin ve nefret dolu yorumlar yapmaktadır. Burada blok zorbalığı adlandırılan bu zorbalığa maruz kalanlar sadece blog yazarları değil, aynı zamanda yorumculardır. Blog zorbalığın az görülen başka bir yanıda blog yazarların yazdığı

içeriklerin başkalarını üzmesi, utandırması veya aşağılaması gibi taraflarının olmasıdır. Bu yetişkinler, okul çocukları ve öğrencilerde gözükmektedir.

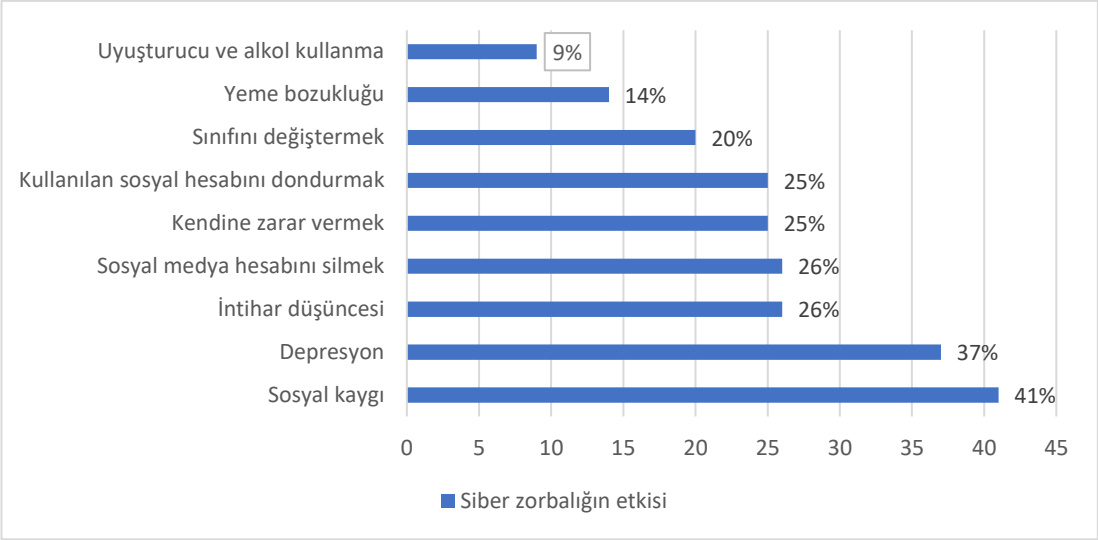
3.3 Siber Zorbalığın Etkileri

Genel olarak zorbalığın mağdurların genel yaşamına ve zihinsel sağlığına da etkileri çok yüksektir. Yapılan araştırmalara göre siber zorbaların çevrimiçi ortamda daha agresif olması nedeniyle onların kurbanlarının üzerinde daha fazla etkisinin olduğu tespit edilmiştir. Siber zorbalığa uğrayan insanlar yapılan zorbalıklardan sonra hayatlarında sosyal kaygı, depresyon ve düşük özsaygı gibi durumları yaşadıkları için intihar düşüncelerine girme olasılıklarıda yüksektir. Hatta bu zorbalık türü insanların düşük eğitim hayatı, yeme bozukluğu, alkol, uyuşturucu ve diğer farklı olaylarla karşı karşıya kalmasına neden olmaktadır. Sonuç olarak bütün bunlar siber zorbalığın ne kadar önemli bir konu olduğunu göstermektedir.

(Grom Social, 2022) kaynağında belirtilen bulgulara göre 2017 Pediatrik Akademik Dernekler Toplantısın yaptığı araştırmalarda 2008-2015 yıllarında intihar düşüncesi olan veya intihar girişiminde bulunan gençlerin sayısı iki katına çıkmış ve bu vakaların çoğu siber zorbalar yüzünden olmuştur.

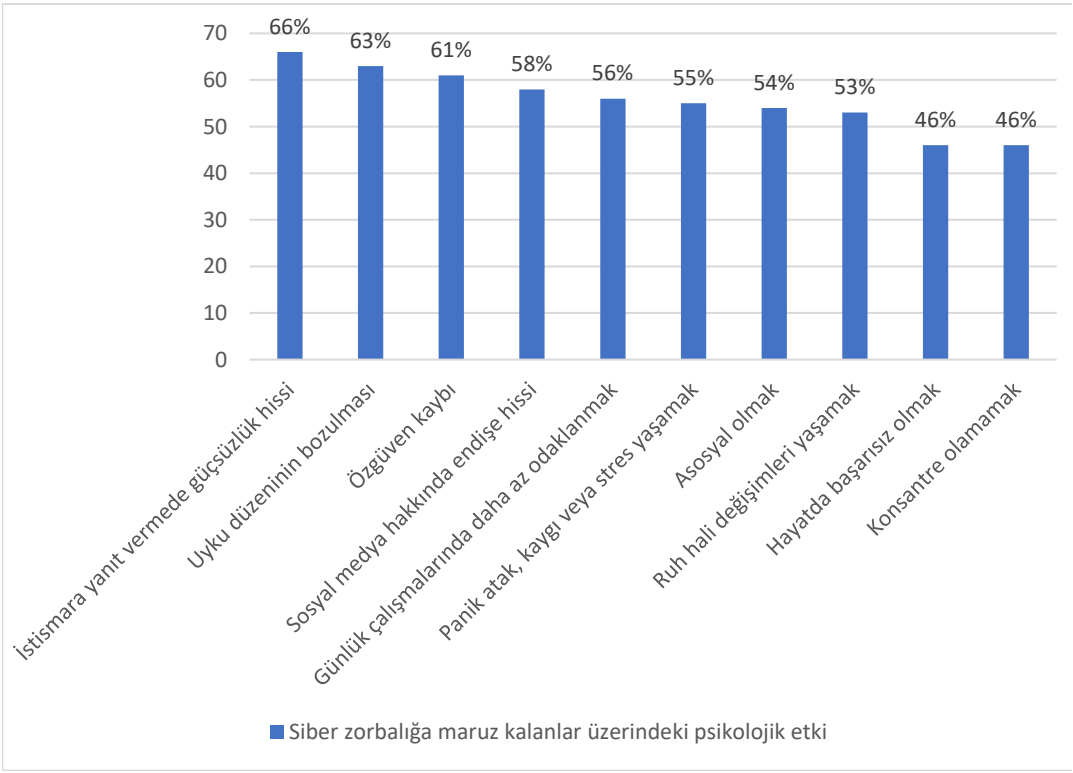
(Ditch the Label, 2022)'nın siber zorbalığının etkileri hakkında yaptığı araştırmalara göre sosyal kaygı, depresyon, intihar düşüncesi ve diğerlerinin yüzdesel olarak oranları Şekil 3.7'da gösterilmiştir.

Çevrimiçi tacizlerden sonra özellikle kadınlar üzerinde psikolojik etkilerinin olduğu görüldü ve bunun hakkında dünya çapında araştırma yapıldı. (Joseph, 2022)'nın 2017 senesinde bu konuyla ilgili yaptığı anketde kadınların zihinsel sağlıkları ve refahlarıyla bağlı ne tür olumsuz durumlarla yüz yüze kaldıklarını öğrendiler. Özgüven kaybı, uyku düzeninin bozulması, yapılan zorbalığa karşı yanıt vermede kendini güçsüz hiss etme gibi olumsuz durumların oranları Şekil 3.8'de verilmiştir.



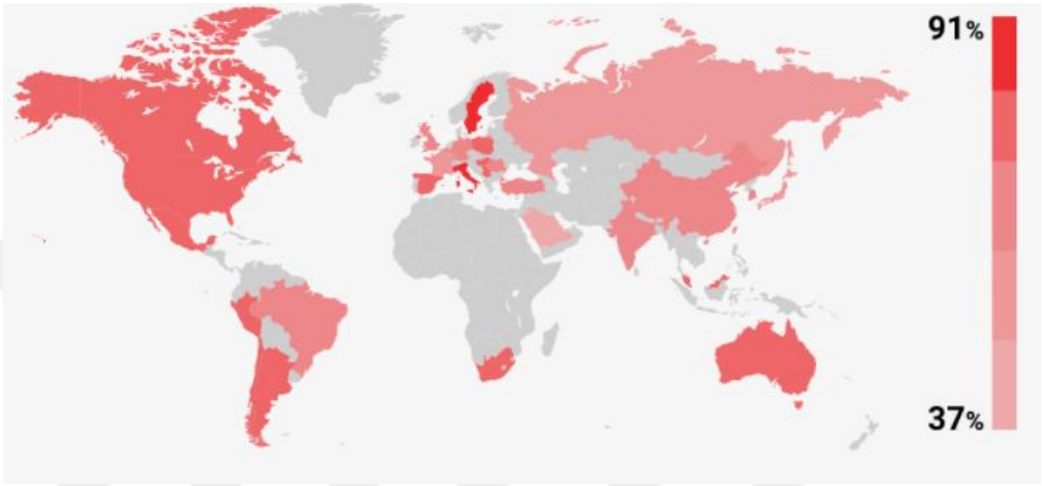
Şekil 3.7 Siber zorbalığın etkileri

Siber zorbalık artık tüm dünyada küresel sorun haline gelmiştir. Hindistan, Brezilya ve ABD bu konuda ön sıradadır ve son zamanlarda diğer ülkelerde de bu oran artmaya başlamıştır. Çocukları sosyal medya kullanan ebeveynlerin %65'inden fazlası siber zorbalık konusunda endişelidirler. Bu yüzden neredeyse tüm dünyada siber zorbalığa karşı önlemler alınmaya çalışılıyor.



Şekil 3.8 Siber zorbalığa maruz kalanlar üzerindeki psikolojik etki

(Joseph, 2022)'nın küresel siber zorbalıkla ilgili yaptığı araştırmalardan sonra belli oranlar elde edilmiştir, bu oran genel olarak tüm dünyada %75 seviyesinde ve ayrıca listenin en başında olan İsveç ve İtalyada ise %91 seviyesinde bulunduğu görülmüşdür. Bu oranların git gide artmasına rağmen en düşük oran olarak Suudi Arabistanda %37'dir. Genel olarak bu oranlarla ilgili bilgiler Şekil 3.9'da detaylı şekilde görselleştirilmiştir.



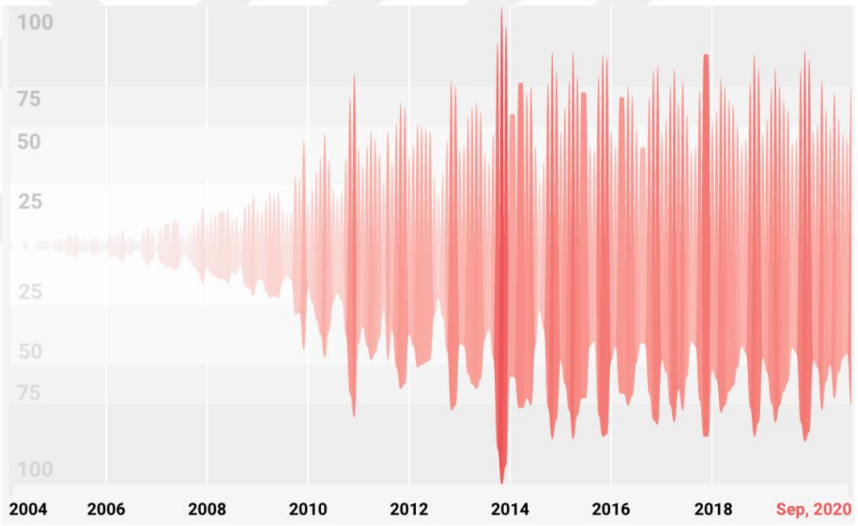
Şekil 3.9 Çocuklarının siber zorbalık kurbanı olduğunu bildiren ebeveynler

(Amarendra, 2022)'nın Hindistanda 2018 yılında yaptığı araştırmalara göre hintli ebeyevnlerin %37'den fazlası çocuklarının siber zorbalığa maruz kaldığını söylediler ve en son 2016'da yapılan araştırmalardan %5 daha fazla olduğu gözükmektedir.

(Helen, 2022)'nın yaptığı araştırmalara göre her 30 ülkeden biri her üç gençten birinin siber zorbalık kurbanı olduğunu ve beşde birininde zorbaların yüzünden okulu asdıkları gözlemlenmiştir. (Joseph, 2022)'nın 20793 ebeveyn arasında yaptığı anket sonucunda %65'nin çocuklarının İnstagram, Facebook ve Snapchat gibi platformalarında çevrimiçi mesajlaşma ve sohbet odalarında siber zorbalığa maruz kaldıkları tespitlenmiştir.

3.4 Siber Zorbalığa Tepkiler

Artık neredeyse herkes siber zorbalığın ne olduğunu bilmelerine rağmen çoğu bu zorbalıkla nasıl başa çıkacağını bilemiyorlar. İnsanlar siber zorbalığa tanık olsa bile dahil olmaktan korktukları için sessiz kalmayı tercih ederler. Ebeveynler ise çocuklarının siber zobalığa maruz kaldıkları bilemezler, çünki çocuklar genelde bunu ebeveynlerinden gizlerler. Sosyal medyada zorbalıkla karşı karşıya kalan çocuklar çözümü çoğu zaman zorbaları engellemekte görürler. Bu yüzden ABD'de 48 adet elektronik taciz yasası çıkarıldı ve bunun 44'ü siber zorbalık içeren cezai yaptırımdır. (Google, 2022)'un verilerinde siber zorbalığın bu kadar yaygın olmasıyla son zamanlara doğru aşırı şekilde artım olduğunu Şekil 3.10'da göre biliriz.

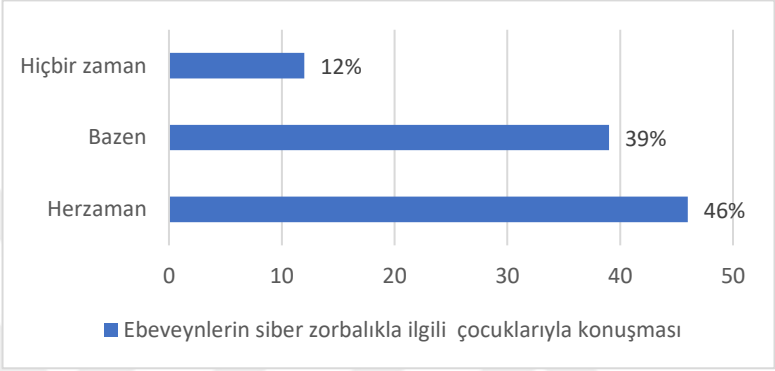


Şekil 3.10 Google'un 2004'ten 2020'ye kadar "siber zorbalık" arama terimiyle ilgili veri göstergeleri

(Office of Justice Programs, 2022)'in 2016 senesinde siber zorbalığa karşı güvenlikle alakalı yaptığı çalışmada destek için başvuru yapan kişilerin sayısı 9.3 milyonu geçmiştir. (Do Something, 2022)'nun gençlerle yaptığı anketlerde %83'ü sosyal medya zorbalıkla ilgili yapılanların az olduğunu ve sosyal medya şirketlerinin daha fazla şey yapmasının gerektiği söylemişlerdir.

Siber zobalığın aslında bu kadar hızlı gelişmesinin nedeni özellikle çocuklar ve gençler bu konuyla ilgili ebeveynlerine konuşmamalarıdır. Son zamanlarda artık ebeveynler bu konuda daha duyarlıdır ve (Telenor, 2022)'nın çalışanları tarafından bu konuyla yapılan araştırmaların sonuçları Şekil 3.11'de gösterilmiştir.

Birçok kişi isimlerinin bilinmesi endişesiyle siber zorbalık vakalarında sessiz kalmayı tercih ediyorlar. Bunun nedeni özellikle okul içerisinde yapılan zorbalıklarda hem mağdurları, hem de failleri tanıdıkları için olası bir müdahelede faillerle aralarının bozulması endişesini yaşarlar. Hatta (Do Something, 2022)'nun öğrenciler arasında yaptığı araştırma sonucunda %81'i eğer isimlerinin anonim kalacağı bir ortam olursa o zaman siber zorbalığa müdahale edeceklerini söylemişlerdir.

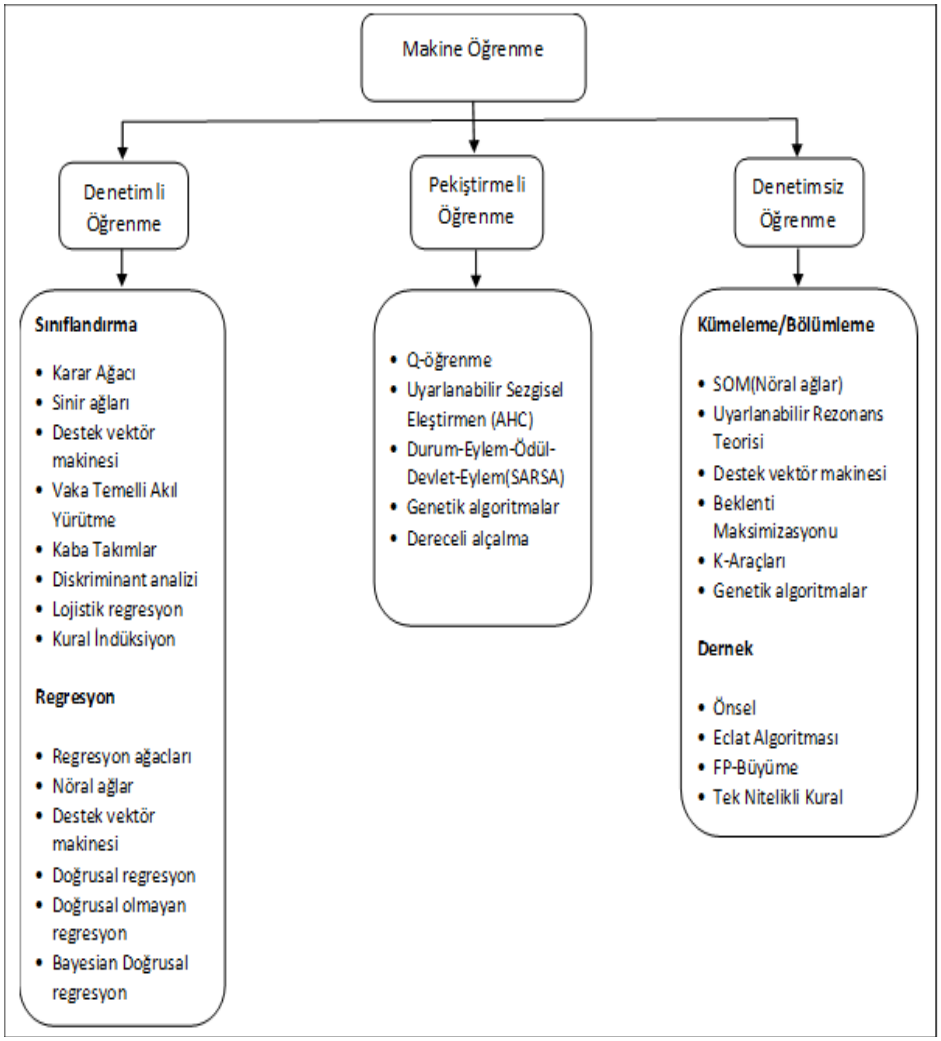


Şekil 3.11 Ebeveynlerin siber zorbalıkla ilgili çocuklarıyla konuşması

Artık gençler ve çocukların çoğu bu durumun normal bir hal olduğunu söylemektedirler ve ebeveynlerinin onlara karışmamasını istemekdedirler. Bu ise ebeveynleri endişelendirdiği için onlar siber zorbalıkla ilgili farkındalığı artırmaya çalışmalar yapılmasını istemekdedirler. Bunun için ise artık neredeyse tüm dünyada bulunan devletler hem yazılımsal olarak, hemde eğitimsel açıdan çalışmalar yaparak, çözümler üretmeye başlamışlartır.

BÖLÜM 4
MAKİNE ÖĞRENİMİ YÖNTEMLERİ

Bilgisayar biliminin bir alt dalı olan Makine öğrenmesi, 1959 yılından itibaren yapay zeka kavramının sayısal öğrenme ve model tanıma çalışmalarından sonra daha yaygın olmaya başladı. Bilgisayar bilim dalında genel olarak yapay zeka veya makine öğrenmesi birlikte düşünülmektedir. Bu iki yöntem çoğu çalışmada birlikte yada birbirlerinin yerlerine kullanılmasına rağmen farklı yönleride oldukları için aynı anlama gelmemektedirler. Onları bir birinden ayıran en büyük özelliyi karşılaşılan herhangi bir sorunun Yapay zekanın çözümlenmesi zaman bu sorunların hepsini Makine öğrenmesi ile çözmek mümkün değildir ama tam aksine Makine öğrenmesinin çözümlenmesi zamanı tüm sorunlar Yapay zekayla halledilebilinmektedir. Makine öğrenmesi çalışılan verileri ele alarak onlar üzerinde çeşitli tahminler yapan algoritmaların nasıl çalışdığını öğrenen ve yapımını araştıran bir yöntem olup, aynı zamanda yapısal bir fonksiyon gibide öğrenebilen bir sistemdir. Buradaki algoritmalar diğer algoritmalardan farklı olarak çalışdığı zaman programın verdiği istatistik talimatlara göre değilde örnek girdiler üzerinden veri tabanlı tahminler ve belirli kararlar almak için model oluşturarak çalışmaktadırlar (Şafak, 2022). Artık uzun yıllardır kullanılan Makine öğrenimini çeşitli bilim insanları ve araştırmacılar kendi yorumlarıyla tanımlamışdırlar. Bu araştırmacılardan birisi olan Kevin Murphy Makine öğrenmesini şu şekilde tanımlamıştır: “Makine öğrenmesinde kullanılan algoritmalar verilerdeki kalıpları otomatik tespit eder ve sonrasında bu öğrenme tespit edilen kalıpları gelecekteki işlemlerde veriler üzerinde tahmin etmek yada olası bir belirsizlik zamanı karar vermek için kullanan bir dizi yöntemidir” (Murphy, 2012). Birlikte bir çok çalışmada yer almış olan (Thomas ve ark., 2022) yaptıkları çalışmalarda bu öğrenimin tanımını yorumlamıştırlar: “Makine öğrenimi, kullanılan verileri modellere dönüştürmek için bazı hesaplama algoritmalarını ele alarak uyarlayan bir çalışma alanıdır.”



Şekil 4.1 Makine öğrenimi yöntemleri

Makine öğrenimi, verilere ve deneyime dayalı bağımsız bir bilgisayar çözümüdür. Makine öğrenmesinde verilecek kararlar genellikle tahmin veya sınıflandırmaya yöneliktir (Hutter ve ark., 2019; Sarkar, 2019). Eğitim verilerini etiketlemek için bazı eğitim yöntemleri vardır. Burada etiketleme yaparak, verilerin önce insanlar tarafından sınıflandırıldığı ve ardından makine öğrenimi için kullanıldığı kastedilmektedir (Aggarwal, 2018). Ayrıca bir modelin denetimli öğrenmesi, etiketlenmiş verilere bağlıdır ve denetimsiz öğrenme, öğrenme verilerinin etiketlenmediği durumdur. Denetimsiz öğrenmede makine, veriler arasındaki benzerlik ve farklılıklara göre kendini sınıflandırmayı öğrenir. Yarı denetimli öğrenme, etiketlenmiş ve etiketlenmemiş verilerin bir kombinasyonuna uygulanan öğrenmedir (Rosa ve ark., 2019).

4.1 Makine öğrenimi türleri

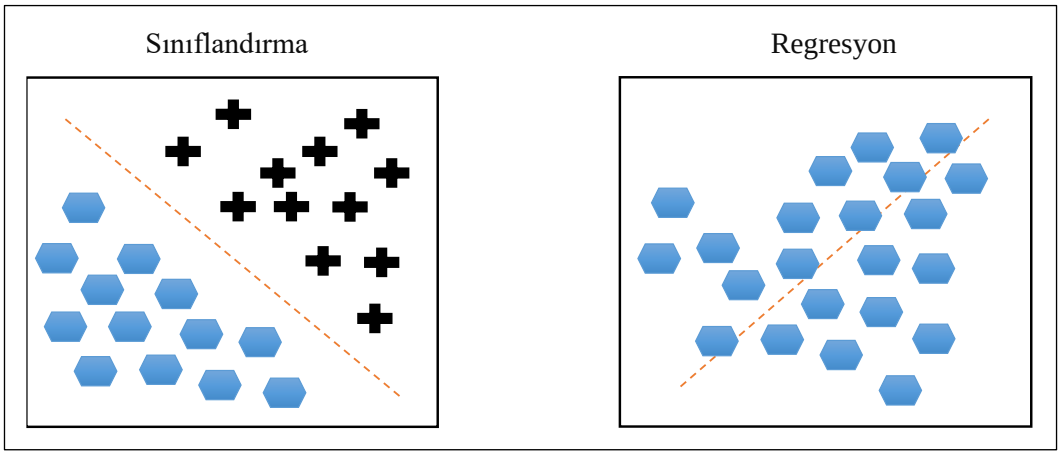
Makine öğrenimi türleri, Şekil 4.1'de gösterildiği gibi kullanım amacına ve hedef çıktılara göre üç ana kategorisi vardır (Potentia Analytics, 2022):

- Denetimli öğrenme (Supervised learning)
- Denetimsiz öğrenme(Unsupervised learning)
- Pekiştirmeli öğrenme(Reinforced learning)

4.1.1 Denetimli Öğrenme

Makine öğrenmenin bir türü olan Denetimli öğrenme, bir model eğitir ve bu model değişkenlerden ibaret olan bir veri kümesinden oluşur. Bu değişkenler ise girdi(X) ve etiketli çıktı(Y) olarak iki yere ayrılmaktadırlar. Denetimli öğrenmenin çalışma yöntemi bakıldığında burada eğitilen modele göre yeni girdi verileri kullanılarak sonda alınan çıktı verileri tahmin edilmektedir (Kaşgarlı, 2021). Şekil 4.2'de gösterildiği bu öğrenmenin iki çeşit sınıflandırılma yolu vardır (Kumar, 2022).

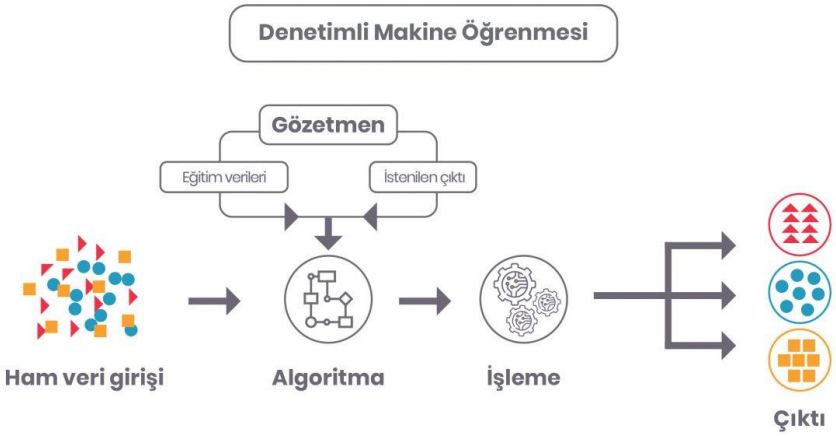
- Sınıflandırma (Classification)
- Regresyon (Regression)



Şekil 4.2 Sınıflandırma ve Regresyon

Sınıflandırma yöntemi modelin kesikli ve kategorik yanıt değerini tahmin eder. Regresyon ise sadece modelin yanıt değerini tahmin eder.

Şekil 4.3'de Denetimli Makine öğrenmesinin aşamalı şekilde nasıl çalışdığı gösterilmiştir.



Şekil 4.3 Denetimli Makine öğrenmesi

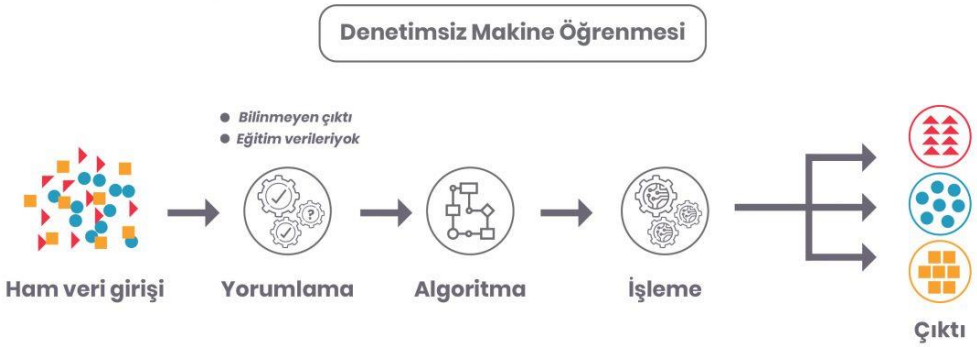
Denetimli öğrenmenin çalışmalarda daha yaygın kullanılan algoritmalarına bakıldığına aşağıdakiler söylenebilir:

RF; Gradyan Artırılmış Ağaçlar; DT; Doğrusal Regresyon; LR; SVM; Sinir ağlar; NB; KNN.

4.1.2 Denetimsiz öğrenme

Denetimsiz öğrenmenin denetimli öğrenmeden en büyük farkı bu yöntemin veri seti ve çıktısının olmamasıdır. Bu öğrenmenin amacına bakıldığında Denetimsiz öğrenmenin veriler üzerinde daha iyi bilgi elde etmek için verilerin altında bulunan dağılımı modellemesi görülmektedir. Burada kullanılan algoritmalar ise Tanımlayıcı modelleme olup, sadece etiketlenmemiş verileri kullanır. Bu yöntemin kümeleme mantığına bakıldığında, burada üzerinde çalışılan veriler arasındaki benzerliğe ve aynı verilerin bir biriyle olan ilişkilerine göre gruplandırıldığı tespit edilmektedir. Bütün bu işlemler yapıldığında Denetimsiz öğrenmenin belirli sayda daha fazla kullanılan algoritmaları bunlardır: Temel Bileşen Analizi; Birliktelik kuralı; K-means kümeleme; t – Dağıtılmış Stokastik Komşu Gömme.

Şekil 4.4'de Denetimsiz Makine öğrenmesinin aşamalı şekilde nasıl çalışdığı gösterilmiştir.



Şekil 4.4 Denetimsiz Makine öğrenmesi

4.1.3 Pekiştirmeli öğrenme

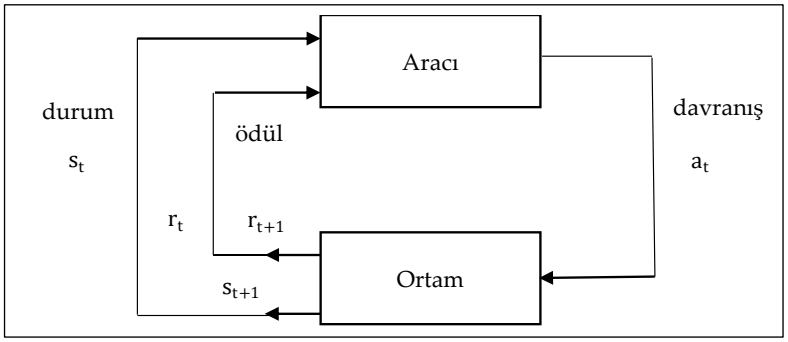
Pekiştirmeli öğrenme çalışılan ortamda en iyi şekilde sonucu almak için kullanıcıların yapması gereken işlere göre rotasını oluşturan bir makine öğrenme yaklaşımıdır. Bu yöntemin çalışma tarzı diğer öğrenmelerden farklı olduğu için genellikle bu sorun bilgi teorisi, kontrol teorisi, simülasyon tabanlı optimizasyon, oyun teorisi, yöneylem araştırması ve istatistik gibi diğer birçok farklı dalda da çalışılmaktadır. Pekiştirmeli öğrenme çalışma ortamı Markov Karar Süreci olarak modellendiği için, bu öğrenmenin algoritmaları statik değilde dinamik programlama tekniklerini kullanmaktadırlar.

Pekiştirmeli öğrenme ortamdaki belirsizlik olsa bile hedefine erişebilmek için aktif olarak karar veren bir aracının ortamla arasındaki ilişkisini inceler. Bu arac ise işlem zamanı kendi davranışlarını seçerken yararlanma-arama(exploit-explore) ikilemi ile karşılaşır. Pekiştirmeli öğrenmenin sistemi:

- π => yaklaşım(policy)
- r => ödül fonksiyonu(reward function)
- Q^{π}=> değer fonksiyonu(value function)
- s => ortam modeli

Şekil 4.5'de Pekiştirmeli Makine öğrenmesinin aşamalı şekilde etkileşim içinde olduğu ortam gösterilmiştir ve burada bu öğrenmenin amacı aşagıdaki hesaplamalarla π^* optimal yaklaşımını bulmaktır.

$$V^*(s) = \max_{\pi} V^{\pi}, Q^*(s,a) = \max_{\pi} Q^{\pi}(s,a) \quad (4.1)$$



Şekil 4.5 Pekiştirmeli Makine öğrenmesi

Eğitim verilerinin farklı eğitim yöntemleri ile etiketlenmesine bağlı olarak bu çalışmada kullanılan algoritmalar aşağıda sunulmuştur:

Q-öğrenme, Uyarlanabilir Sezgisel Eleştirmen, Durum-Eylem-Ödül-Devlet-Eylem(SARSA), Genetik algoritmalar, Dereceli alçalma.

4.2 K-En Yakın Komşular (KNN)

Makine öğrenmesinin yaygın kullanılan algoritmalarından biri olan KNN'nin kullanılma nedeni en basit sıralama mantığı ile çalışmasıdır. KNN sınıflandırma yaptığı zaman benzerlik metriği yöntemiyle çalışan bir tembel öğrenen (örnek tabanlı) algoritmadır (Peterson, 2009). Bu algoritmanın çalışma mantığı ele alındığında, burada yeni bir sınıfın bulunduğu veri setinde ona en yakın K komşusuna bakılarak belirlendiği görülmektedir. Sonrasında bu yeni sınıf en yakın K komşusunun çoğunluğunu ait olduğu sınıfa atanır. Burada veri setinin analiz edilmesi önemlidir. Bunun nedeni ise bulunacak değerin K ve benzerlik için en uygununun olmasıdır. KNN algoritmasının diğerlerinden avantajlı yapan özellikleri gürültülü verilere (çok miktarda ek anlamsız bilgilerden oluşan veriler) veya diğer adı ile bozuk verilere karşı daha iyi sonuçlar verebilmesi ve eğitim aşamasının olmamasıdır.

Çalışmalarda genellikle ızgara arama sürecinde kullanılan model için herhangi bir parametre kombinasyonu, kullanıcı tanımlı parametre değerlerinden oluşturulur. Bundan sonraki süreçin devamında kullanılan modelin paramatreye göre en iyi sonuçları hangi kombinasyonlarda veriyorsa onları seçer ve veri kümesine uygular.

(Bozyiğit ve ark., 2021) yaptığı çalışmada da KNN algoritması kullanılmış ve bu modelin parametreleri için değerler $K = \{1,2,3,\ldots,20\}$ ve metric = $\{euclidean, manhattan, minkowski\}$ olarak kullanılmıştır. Sonuç olarak ise ızgara aramasına göre en iyi yüzdeleri K = 6 ve metrik = euclidean kullanıldığında vermiştir.

KNN çalışma sürecinde hedef test kaydını ve bu hedefi etiketlemek için hedef test kaydına en yakın eğitim kaydını kullanır. Burada model yakınlık, benzerlik yada uzaklık kullanarak ölçülür ve en iyi sonuç için gereken en küçük mesafe veya maksimum benzerliyin alınmasıdır (Kanan ve ark., 2020).

Şekil 4.6'da KNN algoritmasının çalışma mantığının sözde kodu gösterilmiştir.

Input: eğitim seti D, test nesnesi x, kategori etiket seti C
Output: test nesnesi x'in c_x kategorisi, c_x C'ye aittir
begin
{
foreach y D'ye aittir **do**
 y ve x arasındaki D(y,x) mesafesini hesapla
end loop
D veri kümesinden N alt kümesini seçin
N, x örneğinin k en yakın komşusu olan k eğitim örneğini içerir

x kategorisini hesapla:

$$c_x = arg \max_{c \in C} \sum_{y \in N} I(c = class(y))$$

}
end

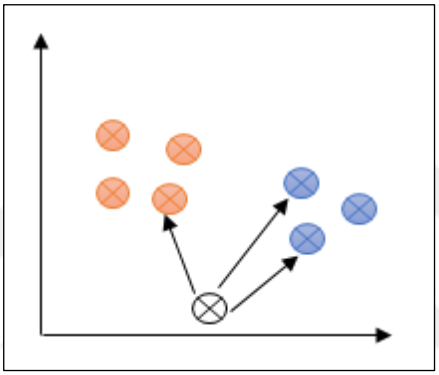
Şekil 4.6 KNN'in sözde kodu

KNN algoritması diğer makine öğrenimi algoritmalarına göre uygulaması kolaydır ve gözetimli algoritma olarak ta bilinmektedir. Bu algoritma sınıflandırma ve regresyon gibi sorunlarının çözümünde kullanıla bilinmektedir. Ama genellikle endüstride sınıflandırma sorunlarını çözmekte kullanılır. KNN ilk defa 1967'de T.M. Cover ve P.E. Hart isimli iki araştırmacının katkılarıyla önerilmiştir. Algoritma önceden sınıfları belirlenmiş olan kümelerdeki veriler üzerinde çalışmaktadır. Burada ele alınan veri setine eğer yeni bir veri katılıyorsa bu zaman onun diğer verilerle arasındaki uzaklığı hesaplanır ve k sayıda yakın komşuluğuna

bakılır. Bahsedilen veriler arasında olan uzaklığı hesaplamak için 3 farklı uzaklık fonksiyonu kullanılır:

- “Euclidean”
- “Manhattan”
- “Minkowski”

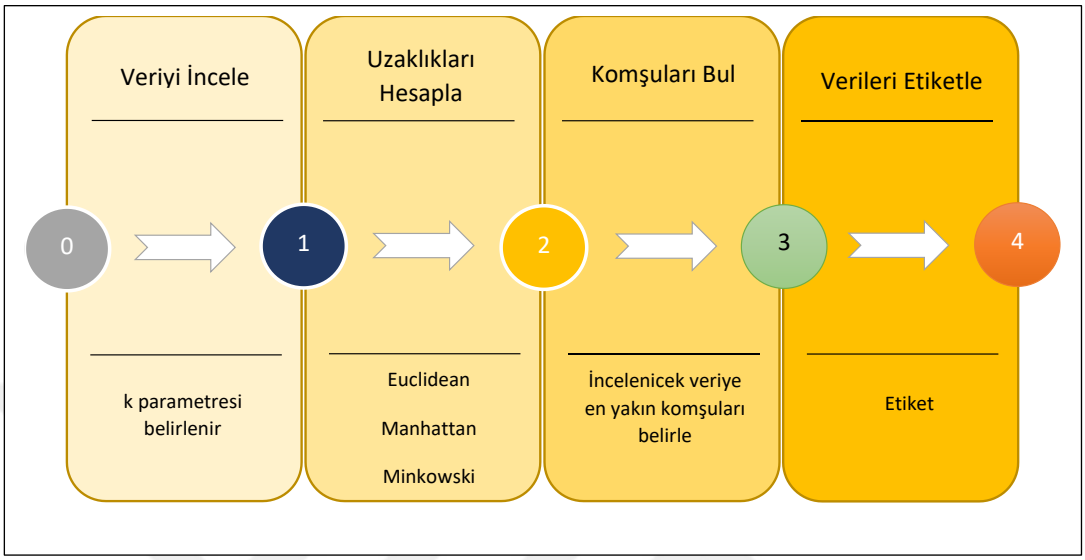
Bu fonksiyonlar arasında gennelikle en yaygın olarak kullanılan “Euclidean” uzaklık fonksiyonudur. Şekil 4.7 de KNN algoritmasının örnek çizimi verilmiştir.



Şekil 4.7 KNN örnek çizim

Bu algoritmanın dikkat edilmesi gereken özelliyi kullanıldığında uzaklık hesabı yaparken tüm durumları saklamaktadır ve bu yüzden büyük veriler üzerinde çalışıldığında bellek alanını çok kapsar. KNN algoritmasının adımlarının örnek olarak Şekil 4.8’de gösterilmiştir.

Şekil 4.8’de gösterildiği gibi KNN’nin adımlarına baktığımızda ilk olarak k parametresi belirlenir. k burada verilen noktaya en yakın komşuların sayısını bildirir. Mesela k=3 olursa o zaman en yakın 3 komşuya göre sınıflandırma yapılır. İkinci adımda 3 fonksiyondan biri kullanılarak veri setine yeni dahil olan verinin diğer eski verilerle arasındaki mesafeler hesaplanır. Üçüncü adımda yapılan hesaplamalardan sonra yeni verinin öznitelik değerleri komşularından en yakını veya yakınları hangisiyse onun sınıfına atanır. Son adımda ise seçilen sınıf, tahmin edilmesi beklenen gözlem değerinin sınıfı olarak kabul edilir. Yani sonuç olarak yeni veri etiketlenmiş (label) olur.



Şekil 4.8 KNN adımları

4.3 Çok Terimli Naif Bayes (MNB)

Son zamanlarda makine öğrenmesinde sıkça kullanılan NB sınıflandırıcısı, bilgi elde etme tekniği haline gelmiştir. NB modellerinin çeşitli varyantları, metin alma ve sınıflandırma için, belgelerdeki kelimelerin dağılımına odaklanarak kullanılmaktadır (Lewis, 1998). Bir metin sınıflandırılıcağı zaman, nihai yaklaşımlar elde edilecekse, kullanılan model MNB varsayımını yapar ve bu varsayımın iki farklı birinci mertebeden olasılığa sahip olması gerekebilir (McCallum ve Nigam, 1998). Bu konudaki ilk açıklamalardan biri (Duda ve Hart, 1973) tarafından yapılmıştır ve çalışmalarında NB sınıflandırıcı kullanmıştırlar. Bu sınıflandırıcının hızlı ve etkili olmasının yanı sıra kolay uygulanabilir olması da sıklıkla kullanılmasının önemli nedenlerinden biridir.

MNB'nin tercih edilme sebeplerinden biri metin planlamada işlemlerin kolay ve hızlı ilerlemesidir (Rennie ve ark., 2003). Tümevarım için MNB sınıflandırıcı kullanıldığında, basit bir matematiksel model uyarlanmalıdır. Bu model sınıflandırma süresi modelin ayırt edici özelliği olmasına rağmen, tahminin doğruluğu onu diğer daha karmaşık tekniklerden üstün kılabilir. Modelin matematiksel hesaplamalarında

kullanılan her sınıf için bir ağırlık doğrulanır ve bu ağırlık yapılan sınıflandırmada MNB olasılığı ile çarpılır. (Webb ve Pazzani, 1998) çalışmasında, önceden kullanılan olasılık yerine en son düzeltilen değer kullanılarak tahmin doğruluğu sonuçlarının daha optimal hale geldiği gözlemlenmiştir.

Mevcut çalışmamızda, MNB sınıflandırıcısı, bir gereksinimin belirli bir konuyla ilgili olma olasılığını hesaplamak için istatistiksel yöntemler kullanmaktadır. W kelimesinin bir R konusuyla ilgili olma olasılığı, Bayes kuralı kullanılarak aşağıdaki gibi hesaplanır:

$$P(R|W) = \frac{P(W|R)P(R)}{P(W)} \tag{4.2}$$

Şekil 4.9'ta MNB'nin sözde kodu addımlarla gösterilmiştir.

Input:

Eğitim veri kümesi T,

$F = (f_1, f_2, f_3, .., f_n)$ //test veri kümesindeki tahmin değişkeninin değeri

Output:

Bir test veri kümesi sınıfı

Steps:

1. Eğitim veri kümesi T'yi okuyun
2. Her sınıftaki tahmin değişkenlerinin ortalamasını ve standart sapmasını hesaplayın
3. **foreach**

   Her sınıfta gauss yoğunluk denklemini kullanarak f_i olasılığını hesaplayın

   Tüm tahmin değişkenlerinin $(f_1, f_2, f_3, .., f_n)$ olasılığı hesaplanana kadar
4. Her sınıf için olasılığı hesaplayın;
5. En büyük olasılığı elde et;

Şekil 4.9 MNB'nin sözde kodu

4.4 Rastgele Orman (RF)

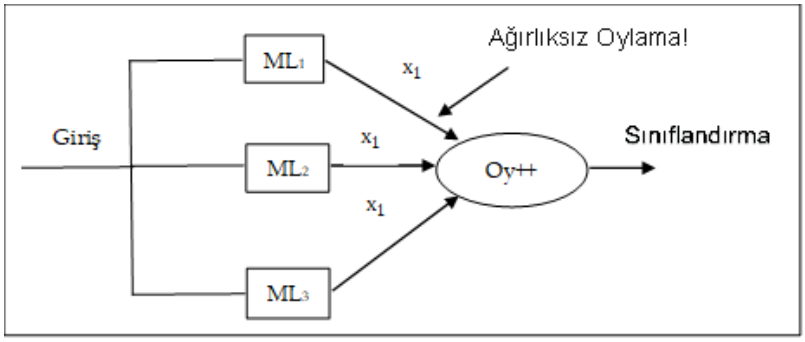
RF makine algoritması bir meta-algoritmadır yani, birçok bireysel ağaçlardan oluşur. RF, belirli bir girdi kümesi için genel sınıflandırmaya karar vermek için birden çok rastgele ağaç sınıflandırmasını kullanır. Genellikle, farklı kararlar eşit ağırlığa sahiptir. Orman, en çok oy alan kararı seçer. Şekil 4.11'da, ağırlıksız rastgele orman algoritmasının görsel bir temsili bulunmaktadır.

RF'ı sıra dışı yapan yönlerden biri, özelliklerinin ne kadar önemli olduğunu göstermesidir. Rastgele orman algoritmasında gerektiği kadar x özniteliği sağladıktan sonra, bunlardan y tanesi rastgele seçilir ve farklı bir modelde kullanılır. RF'ı daha anlaşılır kılmak için akış şeması sembolleri Şekil 4.12'de verilmiştir.

RF algoritmasının sözde kodu Şekil 4.10'da adımlarla gösterilmiştir.

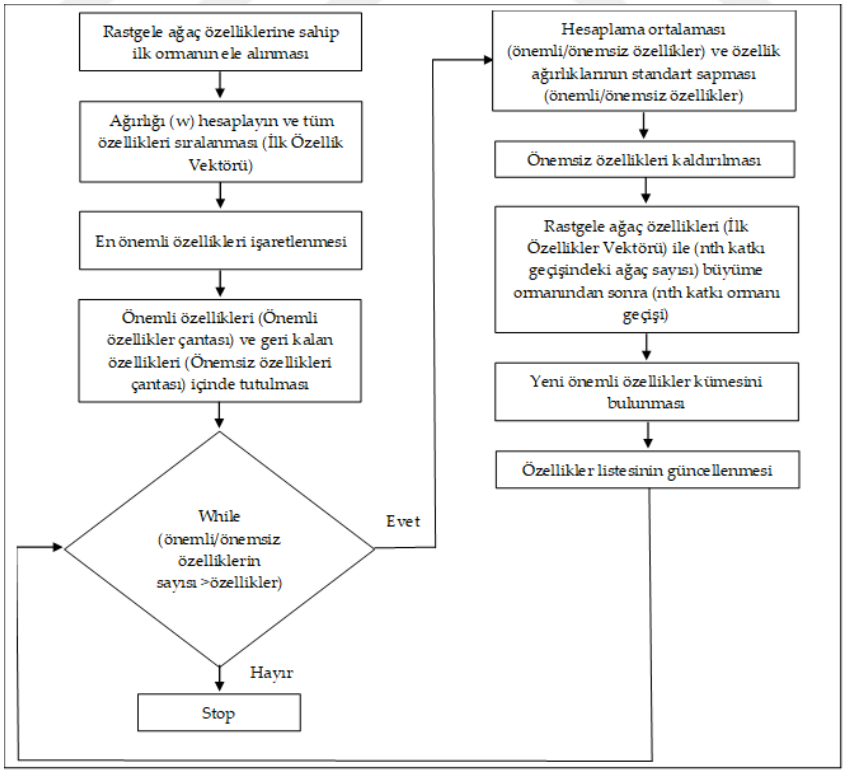
```
c sınıflandırıcıları oluşturmak için:
for i = 1 to c do
  D'yi üretmek için eğitim verilerini D_i yerine  rastgele örnekleyin
  D içeren N_i bir kök düğüm oluşturun
  BuildTree(N_i) ziyaret edin
end loop
buildTree(N):

if(N, yalnızca bir sınıfın örneklerini içerirse)
  return;
else
  N'deki olası bölme özelliklerinin %x'ini rastgele seçin
  Bölünecek en yüksek bilgi kazancına sahip F özelliklerini seçin
  F'nin f olası olduğunda (F_1,...,F_f) N,N_1,...,N_f'nin f alt düğümünü
  oluşturun
 i = 1 to f do
  D_i'nin tüm örneklerin N ile eşleştiği yerde N_i'den D_i'ye içeriğini
ayarlayın
  N_i'nin içeriğini D_i'ye ayarlayın, burada D_i, N'de F_i ile eşleşen tüm
  örneklerdir
  BuildTree(N_i) ziyaret edin
 end loop
 end if
```

Şekil 4.10 RF'in sözde kodu.



Şekil 4.11 Meta Öğrenenler

RF, çeşitli alanlarda yaygın olarak kullanılan bir sınıflandırıcıdır. Örnek olarak, (Ozigis ve ark., 2019) makalesinde, Nijerya'nın Nijer Deltası bölgesindeki çevrenin petrolden ne kadar etkilendiğini ölçmek için kullanılmıştır. (Lin ve Jeon 2002) çalışmasında, potansiyel en yakın komşular (k-PNN'ler) kavramını tanıtılıyor ve rastgele ormanların uyarlanabilir ağırlıklı k-PNN yöntemleri olarak görülebileceği gösteriliyor. Sonuç olarak, rastgele ormanlar kullanıldığında, performansı en üst düzeye çıkarmak için mümkün olan en büyük ağaçların kullanılması gerektiği vurgulanıyor.



Şekil 4.12 RF'nın akış şeması sembolleri

6.4.5 Karar Ağaçları (DT)

Ana fikir, bir kümeleme algoritması kullanarak girdi verilerinin gruplara çoklu bölümlenmesine dayanmaktadır. Entropi Sınıflandırma Ağaçları (ID3, C4.5) ve Regresyon Ağaçları (CART) olmak üzere iki kategoride birçok algoritma önerilmiştir. Önce entropi karar ağaçlarına bakılmıştır. Karar ağaçlarının sınıflandırma zamanı, örneğin herhangi bir ağaç başka bir ağacın ondan daha iyi veya daha kötü olup olmadığına nasıl karar verilir sorusunun ortaya çıkmasının nedeni, bu ağaçların aynı verileri doğru bir şekilde temsil edebilmesinden kaynaklanmaktadır. Bunun için, bu ağaçların kalitesini ölçmek için çeşitli yollar önerilmiştir. Moret (1982), tahmin edilen test maliyetine ve en kötü durum test maliyetine dayalı ağaç boyutu ölçümleri için çalışmalar incelemiştir. Bu ölçüler ise ikili olarak çok uyumlu gösterilmez, bu nedenle kullanılan bir algoritmanın ölçüyü en aza indiren olduğunu ve kendisinden başka ağaçların bazı ağaçlardan daha yüksek olduğunu garanti eder.

DT'ı, özellik tabanları için öğrenme ve akıl yürütme süresi için popüler seçenekler söz konusu olduğunda akla gelen ilk algoritmalardan biridir. Bu uygulama zamanı ölçüm belirsizliklerine karşı koymak için bir dizi değişiklik gereklidir. Sunulan değişikliklerin nedeni, bulanık temsilin sunduğu yürütülen zihnin sembolik karar ağaçlarını göstermektir. Burada faydalı olan, her iki tarafın da birbirini tamamlayıcı avantajlara sahip olmasıdır. Karar ağaçları kolayca anlaşılır ve bulanık gösterimin kesin ve belirsiz bilgilerle karşı karşıya kaldığında bile bunları kolayca aşabilir ve bu özelliği onu örnekler arasında daha çok kullanılan bir uygulama haline getirir (Janikow, 1998).

Rastgele bir değişkenin belirsizliğinin bir ölçüsü olarak bilinen entropi, bir süreç için tüm örneklerde yer alan bilgilerin beklenen değeridir. Eşit olasılıklı durumlar yüksek belirsizliği temsil eder. Shannon'a göre, sistemin durumu değiştiğinde, entropideki değişim elde edilen bilgiyi belirler. Buna göre, maksimum belirsizlikle durum değişikliğinin maksimum sonucu sağlaması muhtameldir. Karar ağaçlarının hesaplanmasında aşağıdaki özdeyişler kullanılır:

$$(x) = log\frac{1}{P(x)} = -logP(x) \tag{4.3}$$

$$H(X) = E\big(I(X)\big) = \sum_{1\le i\le n} P(x_i) * I(x_i) = \sum_{i=1}^{n} P(x_i) log_2 \frac{1}{P(x_i)} = -\sum_{i=1}^{n} P_i log_2 P_i \tag{4.4}$$

(Biggs ve ark., 1991) çalışmasında sınıflandırma veya karar ağacı analizlerinde k-yollu bölümlerin seçilmesi için bir yöntem verilmiştir. Bu yöntem her zaman istatistiksel olarak en iyi parçaları seçer ve Bonferroni eşitsizliğini kullanarak anlamlılığı hesaplar.

Şekil 4.13'te Karar ağacı öğrenmenin sözde kodu adımlarla gösterilmiştir.

```
Ağaç-Öğrenimi (E, H, A)
E - Eğitim örnekleri
H - Hedef özellik
A - Açıklayıcı nitelik kümesi
{
Ağaç için bir Kök düğüm oluşturun
If(E aynı hedef öznitelik t_j değerine sahipse)
 Return(Hedef özniteliği = t_j olan tek düğümlü ağacı, yani
 Kökü döndürün)
If(A = boşdursa(yani açıklayıcı nitelikler mevcut değilse))
 Return(E'deki en yaygın Hedef değerine sahip tek  düğümlü
ağacı, yani Kökü döndürün)
Else
Entropi tabanlı bir ölçüye dayalı olarak E'yi en iyi
sınıflandıran A'dan B niteliğini seçin.
B'yi Kök için öznitelik olarak ayarlayın
For each yasal değer in B, v_i
Kökün altına B = v_i'ye karşılık gelen bir dal ekleyin
E'nin B = v_i olan alt kümesi E_{v_i} olsun
If(E_{v_i} boşdursa)
 Hedef değere sahip o dalın altına eklenen bir yaprak
 düğümü = E'deki Hedefin en yaygın değeri
Else
 Dalın altına, Ağaç-Öğrenimi(E_{v_i}, H, A-{B})tarafından
 öğrenilen alt ağacı ekleyin
End loop
Return(Kök)
}
```

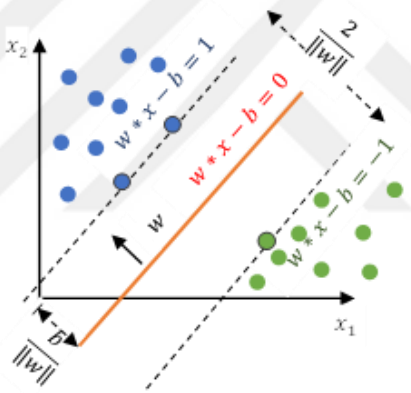
Şekil 4.13 DT'ı öğrenmenin sözde kodu.

4.6 Lineer Destek Vektör Makinesi (SVM)

Lineer Destek Vektör Makine Sınıflandırıcı, verileri sınıflandırma için analiz eden denetimli bir öğrenme modelidir. LSVM, siber zorbalık tespitleri için sık kullanılan algoritmalardan biridir.

(Severyn ve Moschitti, 2012) çalışmasında DAG'ler tavsiye edilmesinin nedeni, SVM algoritması kullanıldığında burada ağaçları daha parçalı bir biçimde temsil etmesidir. DAG kullanmanın bir diğer önemli nedeni, kesme düzlemi algoritmasının CPA yineleme zamanında bir dizi ağaçtan elde edilen modeli kodlamasıdır.

İki sınıfında söz konusu olduğu örneklerle eğitilmiş bir Lineer SVM için maksimum kenar boşluklar hiper düzlem olup ve kenar boşluğundaki örnekler, Şekil 4.14'teki destek vektörleri olarak adlandırılır.



Şekil 4.14 Olası hiper düzlemler

Kullanılan veri sınıflarının, hiperdüzlemle arasında ayırt edici bir faktörü ve maksimum marjı vardır. Buradaki çiftler$(x_i, y_i), x_i \in R^n, y_i \in \{1, -1\}, i = 1, \dots, l,$ SVM aşağıdaki kısıtlı optimizasyon problemini çözmek için kullanılır:

$$min_{w,b} \ \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{l}\varepsilon(\omega, b; x_i, y_i) \qquad (4.5)$$

burada $\varepsilon(\omega, b; x_i, y_i)$ eksik bir fonksiyon olduğunda, $C \geq 0$ eğitim hatasına dayalı bir ceza parametresidir. En çok kullanılan iki hata fonksiyonu:

$$max(1 - y_i)(\omega^T \varphi(x_i) + b),0) \text{ and } \max\,(1 - y_i(\omega^T \varphi(x_i) + b), 0)^2 \quad (4.6)$$

burada φ fonksiyonu, eğitim verilerini gelişmiş boyut uzayına eşdeğer yapmak için kullanılır. Buradaki amaç, karşılaştırma yaptıktan sonra her iki kayıp fonksiyonun arasındaki farkı göstermektir. x'e göre rastgele alınan bir test örneği olarak karar fonksiyonu:

$$f(x) = sgn(\omega^T\, \varphi(x) + b) \quad (4.7)$$

$K(x_i, x_j\,) \;=\; \varphi(x_i)^T\, \varphi(x_j)$ SVM'yi eğitmek için kullanılan bir çekirdek işlevidir. Çekirdek işlevi $K(x_i, x_j\,) \;=\; {x_i}^T\, x_j$'dirse, burada SVM'de $\varphi(x) = x$'dır. Daha çok kullanılan farklı bir çekirdek de RBF'dur:

$$K(x_i, x_j) = exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \text{ where } \gamma > 0 \quad (4.8)$$

Çalışma zamanı Lineer SVM, eş zamanlı özellik sıralama ve sınıflandırma için kullanılmalıdır. Bu arada, sınıflandırıcı görevi olarak RBF çekirdeği ve SVM aynı anda kullanılmalı ve ardından deneyler yapılmalıdır. Lineer SVM algoritmasında C parametresini ve RBF çekirdekli Lineer SVM algoritmasında ise C ve γ parametrelerini belirlemek için izgara tarama yaklaşımı kullanılır. Son olarak, her bir C veya C, γ değerine göre eğitim seti için doğrulukta en iyi performans gösteren parametreler seçilir ve buna göre beş kat daha fazla çapraz doğrulama yapılır.

Özellik sıralama, veriler hakkında bilgi elde etmek ve alınan ilgili özellikleri belirlemek gerektiğinde kullanışlıdır. (Chang ve Lin, 2008) çalışmasında Lineer SVM kullanmanın amacı, farklı öznitelik dizileri ile birleştirerek performansını araştırmak ve deneylerden sonra elde edilen sonuçları rapor etmekdir. Deneylerde kullanılan Lineer SVM modellerinden alınan ağırlıklar kullanıldığında, eğitim ve test süresi verilerinin eşit dağılmadığı durumlarda bile performans açısından iyi sonuçlar verdiği gözlemlenmiştir.

Şekil 4.15'te Lineer SVM algoritmasının nasıl çalıştığını göstermek için sözde kodu adımlarla detaylı şekilde gösterilmiştir.

```
RandomizedSearchCv, en iyi tahmin edici ile tahmin ve/veya
skor kullanarak parametreler için en iyi değerleri seçin
Input: X: train özelliklerine sahip matris, y: etiket
vektörü, hiper parametreler (n_komşular, ağırlıklar,
metrikler)
{
function RSearchCv(n_ komşular, ağırlıklar, metrikler)
 Denemek için gama, çekirdek, karar işlevi değerleriyle bir
sözlük oluşturun
 Hiper parametrelerde Rastgele aramayı kullanın
 Return (en iyi tahmin ediciyi döndür)
Procedure SVM_sınıflandırıcı(x_train, y_train, x_test,
y_test)
En iyi tahmin ediciyi elde etmek için RSearchCV kullanın
Modeli x_train, y_train ile doldurun
x_test için etiketleri tahmin et
Return (tahmin değerleri)
}
```

Şekil 4.15 SVM algoritmasının sözde kodu (Bag ve ark., 2020).

BÖLÜM 5
KULLANILAN ARAÇLAR

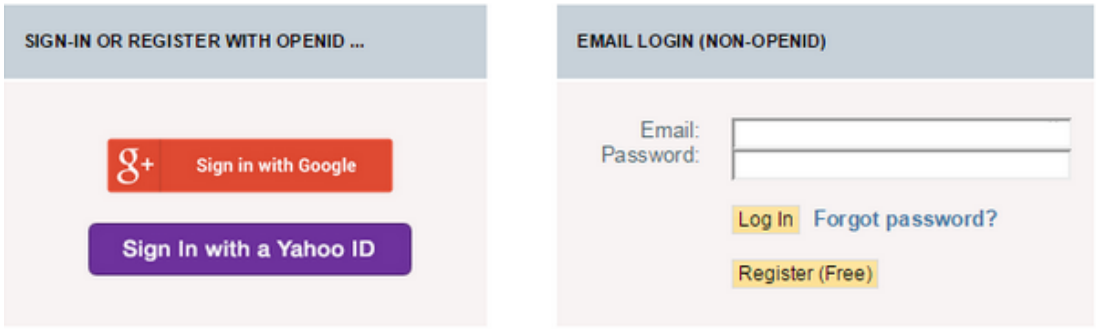
5.1 Netlytic

Netlytic, sosyal medyada halka açık çevrimiçi yazılmış yorumları otomatik olarak özetleyebilen ve aynı zamanda görselleştirebilen, topluluk tarafından desteklenmiş metin ve sosyal ağ analizcisi olarak Dr.Anatoliy ve Philip Mai, M.A., J.D birlikte kurmuştur. Netlytic, büyük hacimli metinleri otomatik olarak özetleyebilir ve Twitter, Youtube, blog yorumları, çevrimiçi forumlar ve sohbetler gibi sosyal medya sitelerindeki konuşmalardan sosyal ağları keşfedebilir ve görselleştirebilmektedir. Twitter, Youtube gibi sosyal medya platformalarında bulunan verileri kolaylıkla CSV dosyası gibi indirilebilinir ve kolaylıkla kullanılabilinir. Bu sitenin en iyi taraflarından biri onun sayesinde sosyal medyadan verilerin çekilmesi için kullanıcıların programlama yada API becerileri gerekmezsizin kullanabilmesidir.

Netlytic, kullanmamız gereken durumlar:

- Sosyal medya sitelerinden ve diğer kaynaklardan(Youtube, Twitter, Google Sheets, Text File, Rss, Reddit) herkese açık yayınları üzerinde çalışmaların yapılması
- Popüler konuların izlenilerek keşfedilebilmesi
- Güncel tartışma konular üzerinde çalışma yapılabilmesi
- Sosyal ağ analizini kullanarak çevrimiçi iletişim ağların oluşturulabilmesi, görselleştirilebilmesi ve son olarak analiz edilebilmesi
- Sosyal medyalarda bulunan verileri coğrafi olarak kodlanmış şekilde kullanılabilmesi

Sisteme giriş yapmak için Şekil 5.1’de belirtildiği gibi mevcut Gmail veya Yahoo!'nuzu kullanılabilinir veya alternatif olarak, başka bir e-posta adresi ile kayıt olunabilinir. Netlytic, kullanıcılara şimdilik sadece birkaç tane sosyal medya platformunda bulunan verilerin keşfedilebilmesi ve görselleştirilebilmesi imkanını yaratıyor.

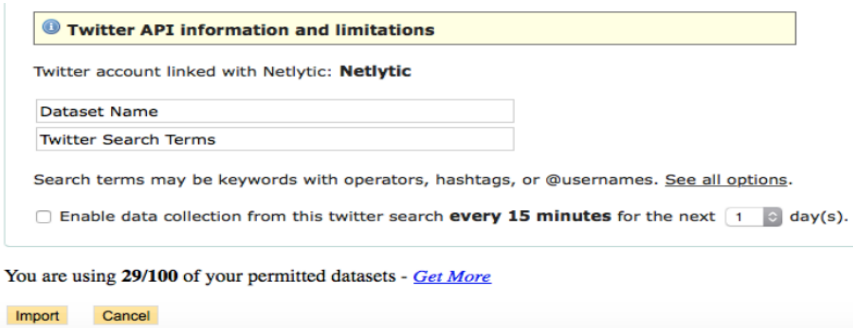


Şekil 5.1 Oturum Açma/Kaydolma

Burada veriler kullanıcıların taleplerine ve ne kadar kayıt istemesine göre farklılık göstermektedir. Bahsedilen sosyal medya plaformalarından verilerin nasıl çekildiği ve kullanılabildiği hakkında güncel bilgiler aşağıda belirtilmiştir:

- **Twitter**

Şekil 5.2'de gözüktüğü gibi operatörler, hashtag'ler veya @kullanıcı adları ile anahtar kelimeleri kullanarak bir Tweet kolleksiyonunu içe aktarmalıdır. Bunun için ise önce kullanılan Twitter hesabını "Twitter Feed" de gerektiği gibi Netlytic ile eşleştirme gerekmektedir.



Şekil 5.2 Twitter içe aktarma

Burada Twitter sekmesi altında doldurulması gereken veri kümesinin isimlendirilmesi ve Twitter'de verilerin toplanması için en iyi anahtar kelime hangisiyse onun yazılması gerekmektedir. Eğer bir kaç tane anahtar kelime girilirse o zaman onlar "ve" veya "veya" kelimeleri ile kullanılmalıdır. Kullanıcının isteğine göre verilerin güncel olarak toplanılması 1,3,7,14 yada 31 gün boyunca her 15 dakikada bir twitter aramasından etkinleştirilir. Son olarak girilen bilgilere göre Twitter'den

verilerin çekilmesini başlatmak için “Import” düymesine basılır ve işlem başlatılır. Girilen bilgiler doğruysa o zaman Şekil 5.3’deki görüntünü görünür.

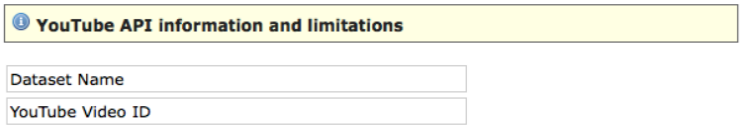
Şekil 1.3 İçe aktarma başarılı

Ama girilen bilgiler yanlışsa yada girilen etiketle ilgili bilgiler yoksa o zaman karşımız Şekil 5.4’deki görsel çıkar. Bu zaman önceki sayfaya geri dönüp bilgileri deyişmek gerekmektedir.

Şekil 5.4 İçe aktarma başarısız

- **Youtube**

Youtube’da olan videoların altında yazılan yorumları içe aktarabiliriz. Bunun için ise iki alanı doldurmak gerekmektedir. Şekil 5.5’de gösterildiği gibi ilk olarak “Veri Kümesi Adı” yazılmalı, ikinci olarak ise “Youtube Video Kimliği” alanı doldurulmalıdır.



Şekil 5.5 Youtube içe aktarma

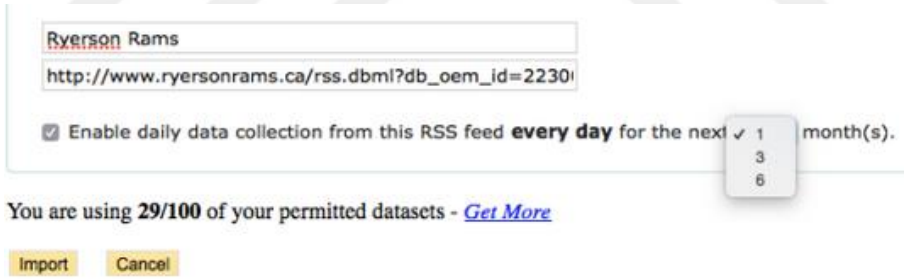
Örnek olarak bir videonun linkini ele alırsak:

“https://www.youtube.com/watch?v=**1CLJwYW6Ao**&index=3&list=PLIPIuyyGllel jTAUpvJ3MzPIxC6GU5Vy2” ve buradan yorumları içe aktarmak için “Youtube Video Kimliği” alanına “v=” ’den sonraki tarafı &’a kadar alıp doldurulması gerekir. Gerekenleri yaptıktan sonra içe aktar butonuna basarak Şekil 5.3’deki görseli görebiliriz.

Eğer girilen linkde bir sorun varsa Şekil 5.4’deki görsel çıkar ve yeniden bilgilerin kontrol edilerek tekrar denenmesi gerekmektedir. Bunun için herhangi bir limit yok ama sunulan çeşitli sosyal medya platformalarından ücretsiz şekilde en fazla 3 adet farklı veriyi içe aktarma işlemi yapılabilir ve bu veriler en fazla 2500 yorum olabilir. Gerektiğinde ücret ödenilerek bu sayı daha büyük çalışmalarda çoğaltılabilinir.

- **RSS/Feeds**

Bu seçenekde Gerçekten Basit Dağıtım(RSS)’da bulunan kayıtları ele alarak içe aktarmaktır. Burada da RSS akışını içe aktarma işlemini yapmak için “Veri Kümesi Adı” ve “Feed Url” alanlarının doldurulmasını gerekmektedir. Netlytic, kullanıcılara 1,3 ve 6 aylık seçeneklerinden birini seçerek RSS akış verilerini toplama seçeneğini sunar ve sonra Şekil 5.6’de görülen “İmport” düğmesine basılması gerekmektedir.



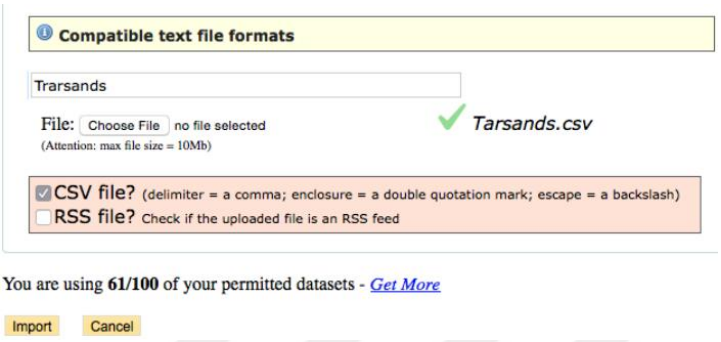
Şekil 5.6 RSS içe aktarma

Sonuç olarak işlem başarılı olursa Şekil 5.3’deki ve eğer başarılı olmazsa o zaman Şekil 5.4’deki sonuç görülür.

- **Text File**

Bu seçenekde bir metin dosyasından metinleri Netlytic’e yüklemek için kullanılmaktadır. Kullanılan veri kümesi birden fazla metin dosyasından oluşuyorsa ve bu dosyaların hepsini aynı anda sisteme yüklememiz gerekiyorsa “Text File” seçeneği kullanılabilinir. Yüklenen dosya iki tür metin dosyasından oluşabilir. Şekil 5.7’de gözüktüğü gibi “CSV File” ve “RSS File” olarak iki seçenekden birinin

seçilmesi gerekmektedir. Burada gözüken kutucuklardan biri seçilerek dosya türlerinden birini seçtikten sonra içe aktarmak butonuna tıklanması gerekir. En sonda sonuç başarılı olursa Şekil 5.3'deki ve eğer başarılı olmazsa o zaman Şekil5.4'deki sonuç görülür.



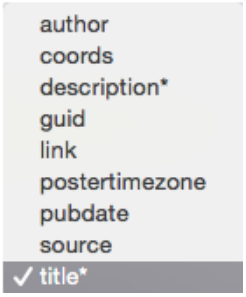
Şekil 5.7 Metin dosyası içe aktarma

Netyltic sayesinde yapılan bu işlemlerden sonra sonuç olarak ele alınan verilerdeki kelimelerin ve sözlüklerin analizi sağlanabilinmektedir. Şekil 5.8'de gösterildiği gibi verilerin analizi zamanı veri setindeki ortak metinsel söylemlerin kısa özetleri görülmektedir.

Şekil 5.8 Metin Analizi için kategorilerin oluşturulması/düzenlenmesi

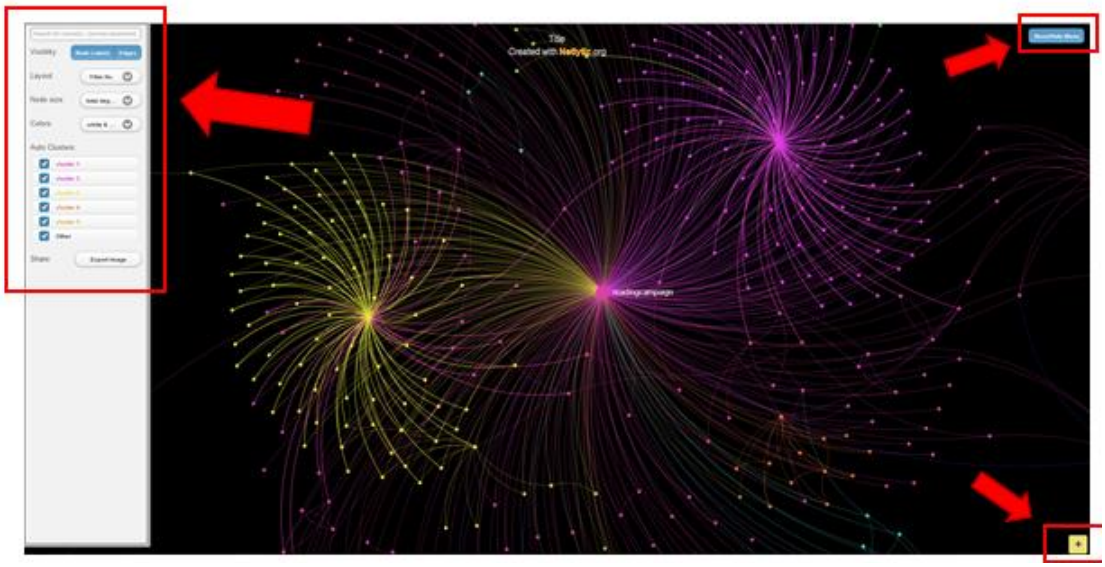
Kırmızıyla işaretlenen kısımda kullanıcılar için çeşitli seçenekler sunulmaktadır ve onlar Şekil 5.9'de görüldüğü mevcut seçenekler yazar, açıklama, kılavuz, bağlantı, yayın tarihi, kaynak veya başlıktır. Özellikle burada Netlytic kullanıcılara anahtar

kelimelerle ilgili olarak mesajlardaki tek tek kelimelerin ve konuların arasındaki ilişkiler hakkında bilgi sunmaktadır.



Şekil 5.9 Metin Analizi işleme seçenekleri

Sonraki adımda, Twitter'den çekilmiş yorumları örnek olarak gösterirsek buradaki kullanıcıların ad veya zincir ağlarına bakılmaksızın, Netlytic'in görselleştirme özelliği ile veri kümesi geniş bir şekilde keşfedilebilir. Kullanıcı Şekil 5.10'de gözükdüğü gibi "Zincir Ağı" veya "Ağı adlandır" kısmında bulunan "Görselleştir" düğmesine basarak HTML5 Ağ Görüntüleyiciye erişebilmektedir. Netyltic Görselleştirme Ekranının üç ana ilgi alanı vardır. Solda Menü Çubuğu, sağ üstte Menüyü Göster/Gizle işlevi ve sağ altta Not işlevi bulunmaktadır. HTML5 Ağ görselleştiricisinde özelleştirilebilir düzen, düğüm boyutu ve küme görünürlüğü özelliklerine sahip olup ve Netyltic kullanıcılara ağın bir görüntüsünü etiketleyebilmek ve dışa aktarabilmek imkanı sunmaktadır.



Şekil 5.10 Ağ görselleştirme, öne çıkan ilgi alanları

Şekil 5.10'de sol yukarı köşede kırmızı okla gösterilen Menü Çubuğu'nun çeşitli görselleştirme seçenekleri vardır ve onlar 9 yere ayrılmaktadır:

- Arama
- Görünürlük
- Düzenleme
- Düğüm Boyutu
- Renkler
- Otomatik Kümeler
- Paylaşma

Arama

Bu özellik sayesinde kullanıcıların adlarını istenilen şartlara göre virgülle ayırarak bir grup düğümü veya belirli düğümleri aramasını sağlar ve bunun sayesinde ağdaki kişiler daha hızlı bir şekilde belirlenilebilir.

Görünürlük

Bu seçenekle kullanıcılar belirli Şekil 5.11'de gösterildiği gibi seçilmiş olan "Node Labels" veya "Edges" seçimlerini değişe bilirler. Yani örnek olarak metinin "noise" denilen bölümü olmadan ağın görüntüsünü dışa aktarmak gerekiyorsa, "Node Labels" seçeneğini kaldırabilir yada kullanıcı daha net bir görüntü istiyorsa o zaman "Edges" seçimini de kaldırabilir.

Şekil 5.11 Görünürlük özelliği

Düzenleme

Ağın düzenlenmesi zamanı kullanıcılar ağdaki kalıpları oluştururken Şekil 5.12'de gözüktüğü gibi Yifan Hu, OpenOrd, ForceAtlas seçeneklerinden birini seçerek bunu yapabilmektedirler.

Şekil 5.12 Düzenleme özellikleri

Düğüm Boyutu

Bu seçenekle düğüm boyutunun çeşitli özelliklerini kullanarak, metinlerdeki yorumları yazan kullanıcılar arasındaki etkileşimlerin nasıl göründüğü sunulmaktadır. "Indegree", "Outdegree" ve "Totaldegree" olarak sunulan 3 derecede farklı anlamlar taşımaktadırlar ve kullanıcının ağdaki kalıpları ayırt etmesine yardım olmaktadırlar. Her seçeneğin ne anlama geldiğinin tam açıklamasına bakılırsa:

- *Indegree*

"Indegree" seçildiğinde görsel olarak derece bazında merkeziliği daha yüksek olan düğümler daha büyük gözükmektedir. Düğümün büyüklülüğü ona yönlendirilen veya alınan bağlarının sayısı ile belirlenmektedir.

- *Outdegree*

"Outdegree" seçildiğinde görsel olarak "Indegree" de olduğu gibi derece bazında merkeziliği daha yüksek olan düğümler daha büyük gözükmektedir. "Indegree" den farkı düğümün büyüklülüğü ona yalnızca yönlendirilen bağlarının sayısı ile belirlenmektedir.

- *Totaldegree*

“Totaldegree” derece sıralaması düğümü oluştururken hem “Indegree”, hemde “Outdegree” sayılarını birleştirir veya yönlendirilmemiş bir ağ olduğunda, belirli sayda düğümün bağlantı sayısı tarafından belirlenir.

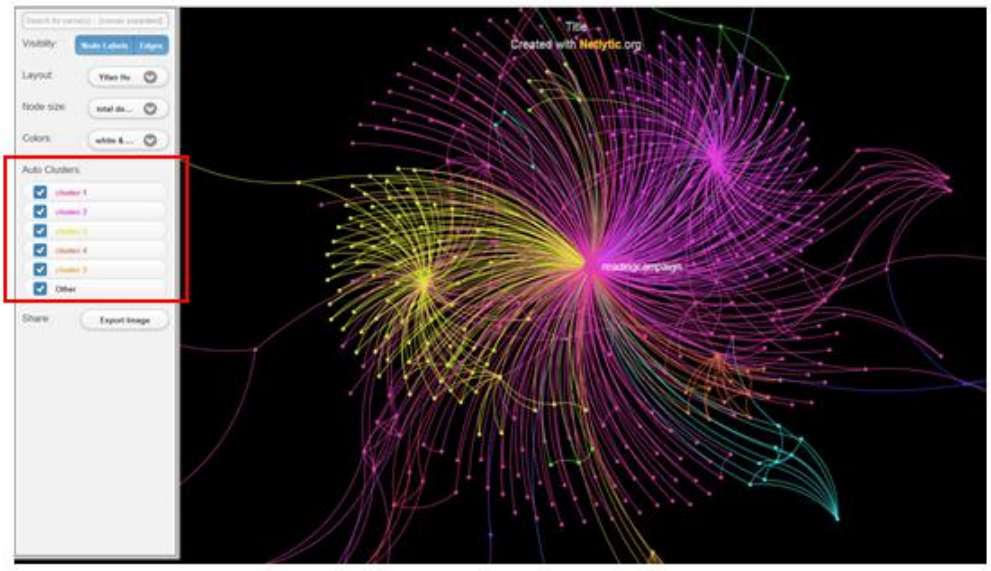
Renkler

Menüdeki renkler kısmında kullanıcıların etiketlerinin renklerini ve görselleştirmenin arka planını sunulmuş çeşitli seçimlerle değiştirile bilinmektedir. Şekil 5.13’de gözüktüğü gibi her seçenekte iki tane renk bulunmaktadır ve bunlardan ilki düğüm etiketini, ikincisi ise ana arka plan rengini belirlemektedir. Burada karışık olarak bakıldığında beyaz, koyu gri, siyah, kırmızı, yeşil, mavi, sarı renkleri vardır ve son olarak da özel renk seçimleri yapılabilinmektedir. Netlytic tarafından varsayılan renk olarak beyaz ve koyu gri olarak ayarlanmıştır, ama genelde kullanıcılar tarafından kullanılan renkler beyaz ve siyahdır. Bunun nedeni siyah ve beyazın daha yüksek kontrast oluşturması ve kullanıcıların düğümleri ve kenarları daha net görmesini sağlamasıdırki, bu da onların analiz aşamasında zorluk çekmemesine olanak sağlar.

Şekil 5.13 Renk özelliği

Otomatik Kümeler

Ağ içinde ele alınan bir kümeyi daha yakından incelemek gerekiyorsa yada çeşitli kümeleri tanımak gerekiyorsa Şekil 5.14’da kırmızı çerçeve ile gösterilen kısım kullanılarak yapılır. Bu kısım kullanıldığında kümeleri FastGreedy adlı bir algılama algoritmasıyla otomatik olarak tanımlanmaktadır.



Şekil 5.14 Otomatik kümeler

Paylaşma

Bu kısmı kullanarak kullanıcılar ağ görüntülerini dışa aktara bilmektedirler. Dışa aktarılan bu görseller veri kümesinin görselleştirilmesinin bir parçası olarak kullanıcılar tarafından tutulabilir veya belgelenerek bir resim olarak kullanıcın bilgisayarında kaydedilerbilir yada çevrimiçi olarak paylaşılabilir.

5.2 Python Programlama Dili

90'lı yıllarda Hollandada Guido Van Rossum tarafından geliştirilen ve çoğu dilden farklı olarak derlenmeden çalıştırılabilen bir programlama dilidir. İsmini MonthyPython isimli bir komedi grubundan alan ve Nesne Yönelik Programlamayı destekleyen bu programlama dilinin Sınıf açma gibi zorunluluğunun olmaması onu özel kılan taraflarından biridir. Kolay uyarlanılabilen olan Python programlama dili bir çok alanda yazılımcılar tarafından kullanılmaktadır. Özellikle zengin kütüphanelerinin de olması onun Makine Öğreniminde de yaygın kullanılmasının en büyük nedenlerindendir.

Çalışmamızda Pythonun Pandas, Numpy, Nltk, Mathplotlib, SciKit-Learn gibi kütüphaneleri ele alınarak kullanılmıştır. Burada Pandas veriler üzerinde ön işleme

gibi işlemleri yapmak için kullanılmaktadır. Pandasın Birleştirme, Sıralama, Yeniden İndeksleme, Yineleme, Toplama, gibi özelliklere sahiptir. Numpy kütüphanesi matrisler ve çok boyutlu diziler üzerinde çalışmakta olup genellikle matematiksel işlemlerde yazılımcılara yardım etmektedir. Nltk Doğal Dil İşleme'de kulanılan bir kütüphane olup, veriler üzerinde etiketlenme ve sınıflandırma gibi çalışmalar yapmak için kullanılmaktadır. SciKit-Learn denetimli ve denetimsiz problemlerde çözüm üreten bir moduldur ve genellikle metin yada resimden özellik çıkarmak için kullanılmaktadır. Son olarak Mathplotlib ise farklı formatlarda çıktılar diğer adıyla grafikler çıkarmak için kullanılır. Ancak çıkarılan görseller kalite olarak yüksek seviyeli değildir.

BÖLÜM 6
UYGULAMA

6.1 Değerlendirme Yöntemleri

Siber zorbalık tespiti bir sınıflandırma görevi olduğu için doğruluk puanı da ilk derecelendirme ölçeği olarak kabul edilir. Bir metrik olan doğruluk (accuracy) daha çok modelin başarısını ölçmek için kullanılır ama malesef tek başına yetersiz kalmaktadır. Doğruluğun değeri modelde tahmin edilen alanların toplam veri kümesine oranı ile hesaplanmaktadır. Doğruluğun eksik kalması nedeniyle onun yerine F1-skor kullanılmasının en büyük sebeblerinden biri ele alınan veri kümesinin eşit dağılmamasıdır. Ayrıca F1-skoru bizim için önemli yapan bir başka nedenide yalnız False Negative veya False Positive değil tüm hata maliyetlerini de içerecek bir ölçme metriği olmasıdır. Duyarlılık (recall) pozitif durumların ne kadar başarılı tahmin edildiğini ve kesinlik (precision) ise pozitif olarak tahmin edilen bir durumdaki başarıyı gösterir. Tablo 6.1'de karışıklık (confusion) matrisi kullanılarak öğrenme performansı hesaplanmıştır. F1-skoru, doğruluk (accuracy), kesinlik (precision) ve duyarlılık (recall) aşağıdaki denklemlerde verilmiştir.

Tablo 6.1 Karmaşıklık Matrisi

	Gerçek Değer		
Model tarafından tahmin edilen		**pozitifler**	**negatifler**
	pozitifler	**TP**(True Positive)	**FP**(False Positive)
	negatifler	**FN**(False Negative)	**TN**(True Negative)

$$Doğruluk = \frac{TP+TN}{TP+FP+TN+FN} \tag{6.1}$$

$$Duyarlılık = \frac{TP}{TP+FN} \tag{6.2}$$

$$Kesinlik = \frac{TP}{TP+FP} \tag{6.3}$$

$$F1 - skor = \frac{2*Duyarlılık*Kesinlik}{Duyarlılık+Kesinlik} \quad (6.4)$$

6.2 Özellik Çıkarma

Bu aşamada sosyal medyalarda paylaşılmış metinler ele alınarak, makine öğrenmesinin üzerinde çalışabileceği özellikler vektör formatına dönüştürülmekte ve belirli çalışmalar yapılmaktadır.

Birçok uygulamada yaygın olarak kullanılan bir yöntem olan n-gram uygulamalarda kullanılmasının nedeni, verilerin içeriklerinin özelliklerini yakalamaktır. Paket yükü analizi, metin analizi gibi birçok alanda kullanılan bir tekniktir. Özellikle ağ yükü analizinde çeşitli yaklaşımlarla birlikte bu kavram çeşitli şekillerde kullanmaktadır (Hadžiosmanović ve ark., 2012).

Siber zorbalık tespiti ile ilgili diğer türkçe çalışmalardan farklı olarak, veri setinde n-gram yöntemi (n=1,2,3,..) kullanılarak belge terim matrisinde sayısal olarak ifade edilmiştir. Böylece argo kelime dizilerini sınıflandırma sürecinde bir bütün olarak değerlendirmek amaçlanmışdır. Örnek olarak "çok gerizekalı birisin" mesajından aşağıdaki terimler alınmıştır.

- 1- gram terimleri: {“çok”, “gerizekalı”, “birisin”},
- 2- gram terimleri: {“çok gerizekalı”, “gerizekalı birisin”},
- 3- gram terimleri: {“çok gerizekalı birisin”}

CountVectorizer Python programlama dilinde bulunan scikit-learn kütüphanesi tarafından sağlanan ve yaygın kullanılan bir teknikdir. Bu tekniğin kullanılmasının amacı ele alınan bir metni, tüm metinde geçen her kelimenin sıklığına yani sayımına dayanarak bir vektöre dönüştürmekdir. Bu yöntemde belgelerdeki her kelimenin sayısı kullanılarak vektörleştirilir, yani kısaca CountVectorizer sadece korpusta bulunan kelimelerin sıklığına odaklanır. Bu yüzden birazda olsa Tf-IdfVectorizer’den geri kalmaktadır.

CounVectorizer özellikle yazılmış yorumlar ve metinler üzerinde uygulandığında örneğin yazılan yorumdaki tüm kelimelerin eşit uzunluğuna sahip kodlanmış bir

vektörü ve her kelimenin bu yorumda görünme sayısı için bir tamsayını döndürür (Alsubait and Alfageh, 2021).

(Kaşgarlı, 2021) çalışmasında da belirtildiği gibi bu teknik her kelimenin sıklığına göre metin dosyasını vektöre dönüştürür. Metin madenciliği ve özellik çıkarma analizi için kullanışlıdır ve söz konusu olan CountVectors(), sütunların bir metinde benzersiz bir kelime olarak temsil edildiği ve ham metnin her metin için örnek olarak temsil edildiği bir matris oluşturur.

Tf-Idf bir terimin doküman içerisindeki önemini gösteren istatistiki yöntem ile hesaplanmış ağırlık faktörüdür. Bir Tf-Idf teriminin ağırlığı, terim frekansının (terimin bir belgede bulunma sayısı) ters terim frekansıyla (veri kümesindeki belge sayısının terimi içeren belge sayısına log oranı) çarpılmasıyla elde edilir.

Daha açık olarak, d belgesindeki t kelimesi için Tf-Idf değeri, C belgeler kümesi dikkate alınarak (korpusundan) aşağıdaki gibi hesaplanır:

$$tf\_idf(t,d,C) = tf(t,d) * idf(t,C) \tag{6.5}$$

burada

$$tf(t,d) = \log(1 + freq(t,d)) \tag{6.6}$$

ve

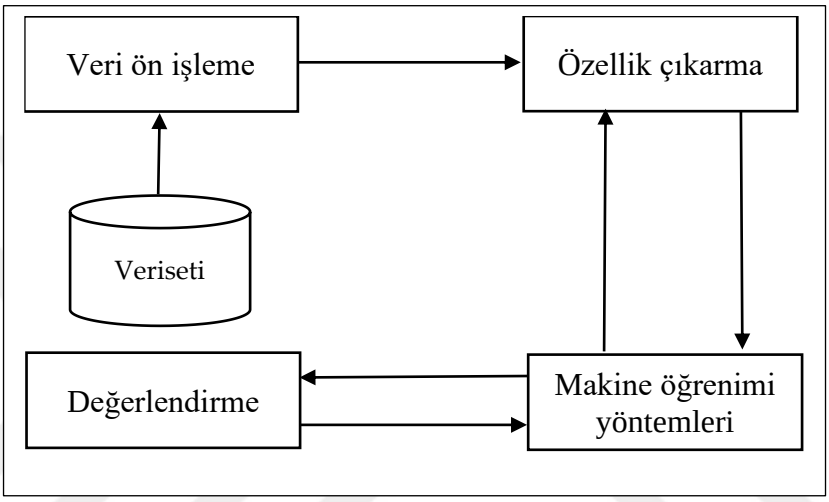
$$idf(t,d,C) = \log\left(\frac{N}{count(d \in C : t \in d)}\right) \tag{6.7}$$

6.3 Çapraz Doğrulama

Çapraz doğrulama, sınıflandırma modellerini değerlendirmek ve modeli eğitmek için veri kümesini parçalara ayırma yöntemlerinden biridir. Örneğin, bin kayıt içeren bir veri setimiz olduğunu varsayalım. Modelimizi bu veri setinin bir parçası ile eğitmek ve modelimizin başarısını bunun diğer bir parçası ile değerlendirmek istiyoruz. Basit bir yaklaşım olarak, eğitim için %75 ve test için %25 sağlanabilir. Ancak veriler parçalı olduğundan, verilerin dağılımına bağlı olarak modelin eğitiminde ve test edilmesinde bazı yanılgılar ve hatalar olabilir. Burada çapraz doğrulama, verileri belirli bir k-sayısına göre eşit parçalara böler ve her parçanın hem eğitim hem de test için kullanılmasını sağlayarak saçılma ve parçalanmadan kaynaklanan sapmaları ve hataları en aza indirir.

6.4 Veri Seti

Bu çalışmamızda, farklı makine öğrenmesi algoritmaları kullanarak türk siber zorbalık mesajlarını tespit etme etkinliğini karşılaştırmayı amaçlamaktadır. Önerilen çalışma, ön işleme ve özellik çıkarma gibi makine öğrenmesi yöntemlerinden oluşmaktadır. Çalışmada gerçekleştirilen adımlar ve aralarındaki akış Şekil 6.1 ile görselleştirilmiştir.

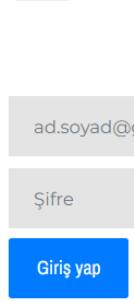


Şekil 6.1 Önerilen işin akışı

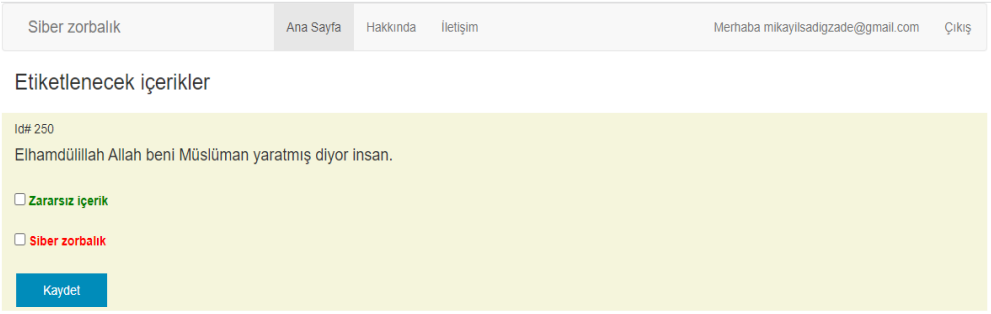
Bu çalışmamızda, toplamda 4400 Youtube paylaşımından oluşan iki farklı veri seti kullanılmıştır. Verisetlerinden birinde 2200 adet cümleden oluşan Azerbaycan dilinde yazılmış metinler kullanılmıştır. Diğer veri setinde ise 2200 adet cümleden oluşan Türkçe yazılmış metinler kullanılmıştır. Bu kümelerdeki mesajların yarısı manuel olarak olumlu (siber zorbalık içeren) ve diğer yarısı olumsuz (siber zorbalık içermeyen) olarak işaretlenmiştir. Çalışmada kullanılan veri seti için daha iyi sonuçlar elde edebilmek için bazı ön işleme adımları gerçekleştirilmiştir. İlk adımda, veri kümesindeki paylaşılan klasörlerden sayısal karakterler, noktalama işaretleri ve web bağlantıları kaldırılarak küçük harfe dönüştürülmüştür. Bu doğrultuda kullanılan veri setinin oldukça dengeli bir dağılıma sahip olduğu dikkat çekmektedir.

Manuel olarak verilerin siber zorbalık içerip içermemesini yapmak için toplamda 3 öğrenci tarafından yapılmıştır. Bu öğrenciler Hem Azerbaycan dilini hemde Türkçeni iyi derecede bilen ve ayrıca uzun süredir Türkiyede yaşayan Azerbaycanlı

öğrencilerdir. Bu işlemin yapılması için frontend olarak Html ve Css, backend olarak ise C# kullanılarak bir websitesi yazılmıştır. Burada Şekil 6.2'de gösterilen kısımdan siteye sadece her kullanıcı belirlenen 2 farklı mail ve şifre ile giriş yapılabilinmektedir. Bu maillerden biriyle (https://github.com/mikayil1551/Tezde-kullan-lan-veri-setleri/tree/main) linkinde bulunan Azerbaycan veri seti, diğeriyle ise Türkçe veri seti üzerinde yapılacak işlemleri göreceklerdir. Kullanıcılar giriş yaptıktan sonra karşılarına Şekil 6.3'deki gibi veri tabanından veriler gelicek. Kullanıcı eğer cümle siber zorbalık içeriyorsa "Siber zorbalık" seçeneğini, yok değilse o zaman "Zararsız içerik" seçeneğini seçer ve en sonda "Kaydet" butonuna basar. Cevap alındıktan sonra otomatik olarak veri tabanına gider. Sonuç olarak 3 kullanıcıdan 2 si cümleni siber zorbalık olarak görüyorsa o zaman veri tabanında yaratılmış yeni tabloda siber zorbalık olarak etiketlenir.



Şekil 6.2 Veri açıklaması için geliştirilen web uygulamasının giriş sayfası



Şekil 6.3 Veri açıklaması için geliştirilen web uygulamasından bir anlık görüntü

6.4.1 Azerbaycan Veri Seti

Veri kümesindeki bazı cümleler ve bunların etiketleri örnek olarak Tablo 6.2'te gösterilmiştir. Bu tabloda toplamda örnek olarak 6 adet cümle verilmiştir. Cümlelerin karşısında "Olumlu" yani siber zorbalık içermeyen yada "Olumsuz" ise siber zorbalık içeren olmasıyla ilgili etiketleri verilmiştir.

Tablo 6.2 Azerbaycan dilinde toplanmış veri kümesindeki bazı cümleler

Paylaşım #	Paylaşım	Etiket
1	İt kimi dilinidə çıxarır çölə	Olumsuz
2	Neçenci defedir baxıram,her defe zövq alıram	Olumlu
3	Ala petux besdide sen bilmirsen axiri is verenide gayirilar	Olumsuz
4	Həqiqətən çox gözəl videodu . Çox bəyəndim	Olumlu
5	Allah hec senin qapini acmasin	Olumsuz
6	Yaziq duz danisir deye haminin acigi gelir. Bu dunyada duz danisanlara yer yoxdu	Olumlu

Azerbaycan dilinde hazırlanmış Youtube'da bulunan videonun altında yazılmış yorumlar Netlytic sitesinin vasıtasıyla çekilmiştir ve burada kullanılmış kelimeler hakkında bilgiler Şekil 6.4'de verilmiştir. Bu şekilde gözüktüğü gibi kullanılan kelimenin sayısı ne kadar fazladırsa o zaman onun görselide büyük olur. Örnek olarak yorumlarda "abdullayev", "allah", "ata" gibi kelimeler daha fazla geçmiştir.

Şekil 6.4 Kelimelerin görsel olarak yorumlarda ne kadar geçtiği

Şekil 6.4'de gösterilen herhangi bir kelimenin hangi yorumda geçdiğini görmek için bu kelimenin üzerine tıklamak yeterlidir. Burada örnek olarak "şərəfsiz" kelimesinin hangi yorumlarda geçtiği Şekil 6.5'de gösterilmektedir.

özü üçün şərəfsiz ələ ən

1 2 3 NEXT

Date	User	Posts (n = 26, including partial matches)
	7	Sənin zatına kökünə sortuna lənət ay **şərəfsiz** biqeyrət pilis adın yavərdi xavərdi nəköpəyin oglusan bilmirəm a biqeyrət a **şərəfsiz** adam vətənfaşla şəhif ailələriylə belə rəftar eliyir a boynu yogun sabah sənində evladın yetim qalacaq onda sənin balanıda
	9	O **şərəfsiz** mayor hələdə işləyirsə daha sözüm yoxdur Daxili işlər nazirliynə
	41	Bir insanda bu qədərdə **şərəfsiz** olarmı
	44	**Şərəfsiz** oğlu **şərəfsizdi** bu atasız.tfu

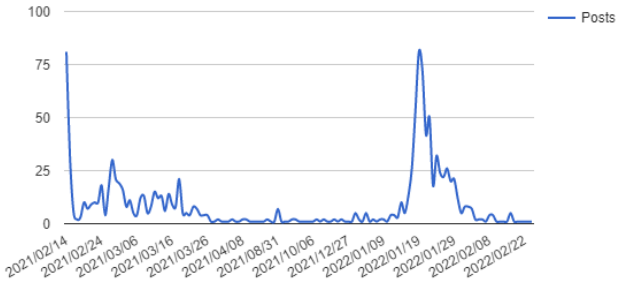
Şekil 6.5 Kelimelerin hangi cümlelerde kullanıldığı

Bazı kullanıcılar yorumlar yazdığında kullanılan kelimelerin üst üste düşme olasılığı görülmektedir. Şekil 6.6'da Azerbaycan dilinden oluşan veri setindeki kelimelerin yorumlarda kaç kez geçtiği gösterilmektedir ve bu kelimelerin siber zorbalık içerip içermemeside belirtilmiştir. Genelde siber zorbalık içeren cümleler tabloda gösterilen siber zorbalık içeren kelimelerle oluşur. Ama bu durum değişiklik gösterebilmektedir, çünki bazen cümlede siber zorbalık içeren herhangi bir kelime olmamasına rağmen cümle siber zorbalık içerebilir.

Kelimeler	Kelimelerin cümlelerde geçme sayısı	Siber zorbalık içerip(1) içermemesi(0)
video	21	0
yaltaq	17	1
allah	84	0
azerbaycan	16	0
şərəfsiz	28	1
petux	15	1
yaxşı	24	0
oğraş	14	1
uğurlar	12	0
peyser	11	1

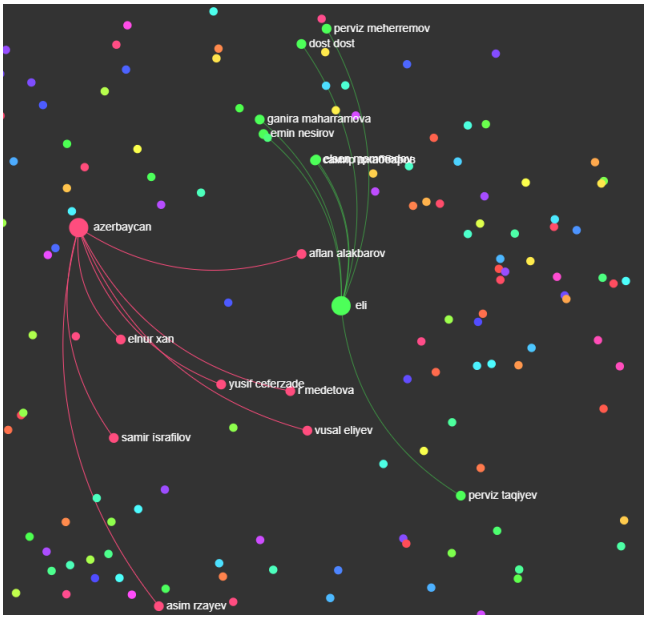
Şekil 6.6 Azerbaycan dilinden oluşan veri setindeki kelimeler

Şekil 6.7'de Azerbaycan dilinden oluşan veri setinde kullanıcılar tarafından yazılmış yorumlar hangi tarihler arasında daha sık yazıldığı görülmektedir. Burada genelde yorumların Ocak ve Şubat aylarında daha fazla yazıldığı tespit edilmiştir.



Şekil 6.7 Yorumların yazılma tarihleri

Youtube veya diğer sosyal medya platformalarında bulunan paylaşımlar altında kullanıcılar sadece video yada görseller hakkında düşündüklerini yazmıyorlar aynı zamanda bir birlerinede yorumlar yazmaktadırlar. Şekil 6.8'de kullanıcılar düğümler şekilde gösterilmiştir. Eğer kullanıcıların ortak kullandığı kelimeler varsa onlar arasında ilişkiler bağlarla gösterilir. Örnek olarak Şekil 6.8'de gözüktüğü gibi "azerbaycan" ortak kelime olarak bir çok kullanıcı tarafından kullanıldığı için onlar arasında ilişkiler tespit edilmiştir.



Şekil 6.8 Yorumlarda ortak kelimler kulannanlar arasındaki ilişki

6.4.2 Türkçe veri seti

Türkçe kelimelerle yazılmış cümlelerden oluşan bu veri kümesindeki bazı cümleler ve bunların etiketleri örnek olarak Tablo 6.3'te gösterilmiş ve bu tabloda toplamda 6 adet cümle verilmiştir. Burada tablonun son sütununda yazılmış "Olumlu" ve "Olumsuz" etiketler cümlelerin sırasıyla siber zorbalık içermeyen ve siber zorbalık içeren olmasını göstermektedir.

Tablo 6.3 Türkçe yazılmış yorumlardan toplanmış veri kümesindeki bazı cümleler

Paylaşım #	Paylaşım	Etiket
1	Allah belanızı versin adi pislikler	Olumsuz
2	Türkiye de çok var ihracatı olur	Olumlu
3	Yemin ederim ağlıyorum	Olumsuz
4	Bunların kalbi yok gerçekten	Olumlu
5	Siz insan değilsiniz cehennem zebanileri	Olumsuz
6	Bok yiyin emi booook	Olumlu

Türklerden hazırlanmış Youtubeda bulunan videonun altında yazılmış yorumlarda Türkçe olan bu video Netlytic sitesinin vasıtasıyla çekilmiştir ve burada kullanılmış kelimeler hakkında bilgiler Şekil 6.9'da verilmiştir. Bu şekilde Netlytic sitesinin özelliği kullanılarak en fazla kullanılan 100 adet kelimenin görseli verilmiştir. Burada gözüktüğü gibi mesela yorumlarda daha çok "köpek", "allah", "belanızı" gibi kelimeler daha fazla geçmiştir.

Şekil 6.9 Türkçe kelimelerin görsel olarak yorumlarda ne kadar geçtiği

Şekil 6.9'de kullanıcılar tarafından ortak olarak daha fazla kullanılan kelimelerden "yazık" kelimesinin hangi cümlelerde geçtiği Şekil 6.10'da gösterilmiştir.

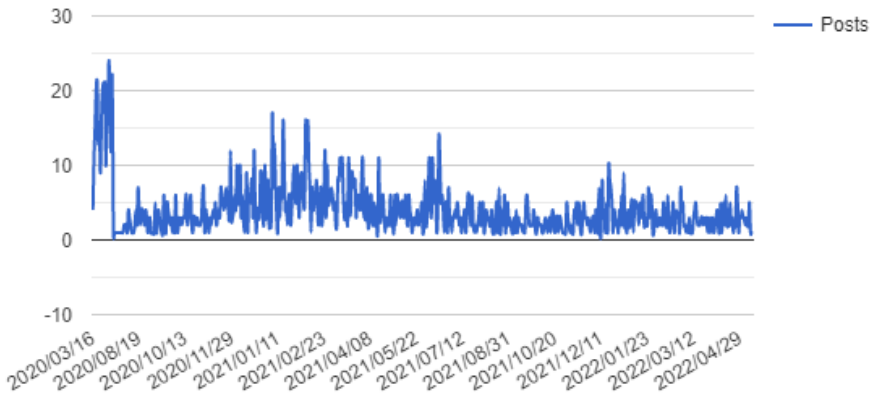
Şekil 6.10 Türkçe veri setinde kelimelerin hangi cümlelerde kullanıldığı

Şekil 6.11'de cümlelerde kullanılma yerine göre fark etmezsizin siber zorbalık içerip içermeyen ve ayrıca bir çok kullanıcı tarafındanda kullanılan kelimelerden örnek olarak 10 adeti seçilerek verilmiştir. Bu tabloda Türkçede de yaygın olarak kullanılan bu kelimelerin veri setinde kaç kez geçtiği gösterilmiş ve son sütunda etiketlenerek belirtilmiştir.

Kelimeler	Kelimelerin cümlelerde geçme sayısı	Siber zorbalık içerip(1) içermemesi(0)
allah	402	0
köpek	43	1
insan	64	0
orospu	38	1
çocuk	20	0
şerefsizler	44	1
güzel	33	0
pislik	28	1
versin	218	0
gerizekalı	25	1

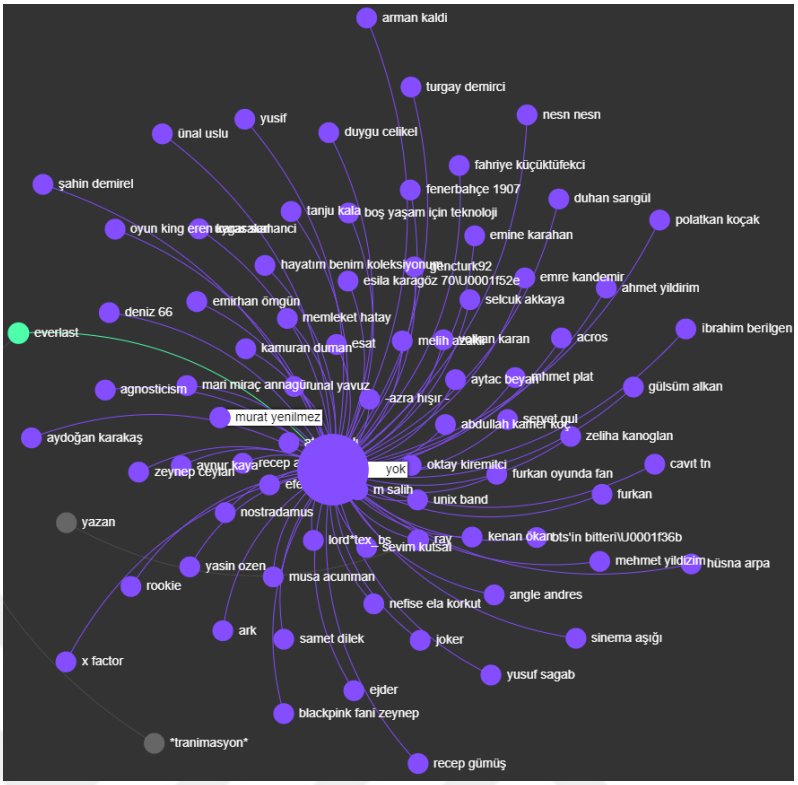
Şekil 6.11 Türkçe veri setindeki kelimeler

Şekil 6.12'de Türkçeden oluşan veri setinde bulunan yorumların 16.03.2020 den 29.04.2022 tarihine kadar hangi sıklıkla yazıldığının bir istatistikası gösterilmiştir. Burada da yorumların daha çok 2020 yılının Mart ayı ve 2021'nin Ocak ve Şubat aylarında yazıldığı tespit edilmiştir.



Şekil 6.12 Türkçe yorumların yazılma tarihleri

Bir çok sosyal medya platforması gibi Youtube da kullanıcılar açık kaynaklı paylaşımların altında kendi düşüncelerini açık bir şekilde yazmakta özgürdürler. Yazılan yorumların yazılma amaçları farklılık göstermektedirler. Örnek olarak çoğu zaman kullanıcılar Youtube da paylaşılan bir video hakkında kendi yorumunu yazar bazen sadece başka birinin yazdığı yoruma yanıt olarak bir şeyler yazarlar. Bu durumda siber zorbalığın olma olasılığı daha yüksek olduğu öngörülmüştür. Şekil 6.13'de mor renkde düğümler gözükmektedir ve onların yanlarında bir kullanıcı yorum yazmak için kullandığı ismi gösterilmiştir. Burada onların bir birlerinin yazdığı yorumlara karşı agresif ve siber zorbalık içeren yorumların olduğu tespit edilmiştir. Bu tarz siber zorbalık içeren yorumlar yazan kullanıcılar genelde ya takma ad, yada sahte isim ve soyisim kullanmaktadırlar.



Şekil 6.13 Yorumlarda bir birlerine yanıt veren kullanıcılar

6.5 Veri Ön İşleme

Veri seti yazılı programda çalıştırılmadan önce bazı ön işleme süreçleri gözlemlenmiştir. Bu işlemlerin anlaşılır olması için sözde kodu Şekil 6.14'de gösterilmiştir.

```
Metin Ön İşleme (Siber zorbalık cümlelerinin listesi)
{
Girişten sayısal karakterleri, noktalama işaretlerini, web
bağlantılarını kaldırın.
Küçük harfli konuşma girişi uygulayın.
For each kelime in cümleler
Türkçe sözcük kökü kullanılarak kelimeler köklerine zeyrek
ile ayrılmıştır. MorphAnalyzer lemmatization kullanılarak
yanlış kelimeler düzeltildi.
End loop
}
```

Şekil 6.14 Metin ön işleme yönteminin sözde kodu

BÖLÜM 7
HESAPLAMA SONUÇLARI VE DEĞERLENDİRMELER

Tablo 7.1, Tablo 7.2, Tablo 7.3 ve Tablo 7.4'te gösterildiği gibi, başarısız tahminler için Makine öğrenimi modellerinin doğruluk sonuçları düşük olarak elde edilmiştir. 5 modelinde doğruluk ve F1-skor değerleri %80 ile %95 arasında değişmektedir.

Tablo 7.1 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları (Azerbaycan dili ile hazırlanmış veri seti)

Modeller	F1-skor	Kesinlik	Duyarlılık	Doğruluk
Lineer SVM	92.84	94.32	94.07	83.31
MNB	92.11	93.62	92.66	83.19
DT	90.55	92.98	91.04	81.39
RF	91.62	96.46	95.85	80.29
KNN	92.08	94.19	93.37	80.82

Tablo 7.2 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenim Modellerinin Değerlendirme Sonuçları (Azerbaycan dili ile hazırlanmış veri seti)

Modeller	F1-skor	Kesinlik	Duyarlılık	Doğruluk
Lineer SVM	93.91	95.49	95.87	84.52
MNB	93.01	94.38	93.17	84.10
DT	91.79	93.41	92.35	82.63
RF	90.32	97.88	96.02	81.61
KNN	92.66	94.34	94.84	81.70

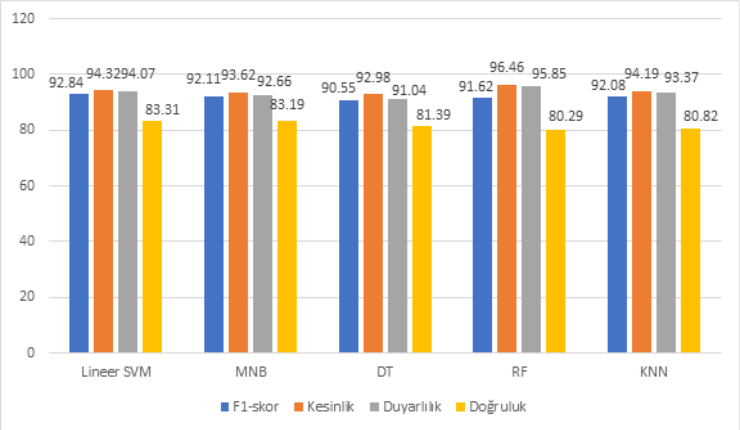
Tablo 7.3 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları (Türkçe ile hazırlanmış veri seti)

Modeller	F1-skor	Kesinlik	Duyarlılık	Doğruluk
Lineer SVM	96.94	95.96	97.97	85.98
MNB	96.50	97.60	96.70	85.90
DT	94.64	98.58	97.54	85.78
RF	93.23	94.66	95.81	82.36
KNN	95.14	96.86	96.91	83.96

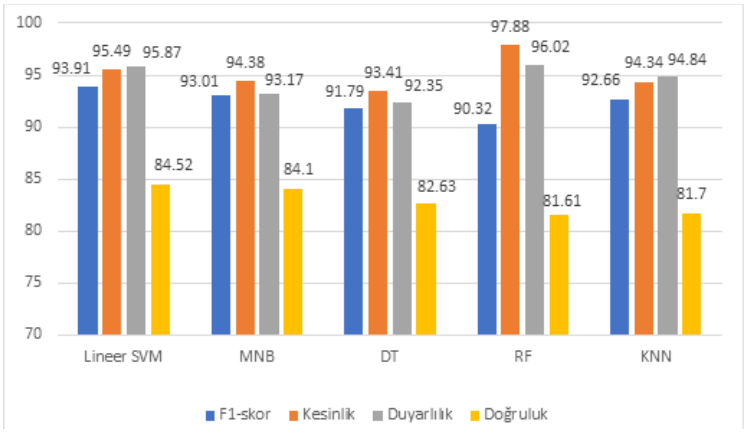
Tablo 7.4 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçları (Türkçe ile hazırlanmış veri seti)

Modeller	F1-skor	Kesinlik	Duyarlılık	Doğruluk
Lineer SVM	97.85	98.06	98.26	85.77
MNB	95.53	96.36	95.44	86.91
DT	94.94	94.88	95.04	84.08
RF	91.23	96.70	96.99	81.21
KNN	95.33	97.21	96.31	84.11

Şekil 7.1 ve Şekil 7.2'de Azerbaycan dilinde hazırlanmış veri seti, Şekil 7.3 ve Şekil 7.4'de ise Türkçede hazırlanmış veri seti kullanıldıktan sonra 5 modelinde sonuçları arasındaki fark gösterilmektedir.

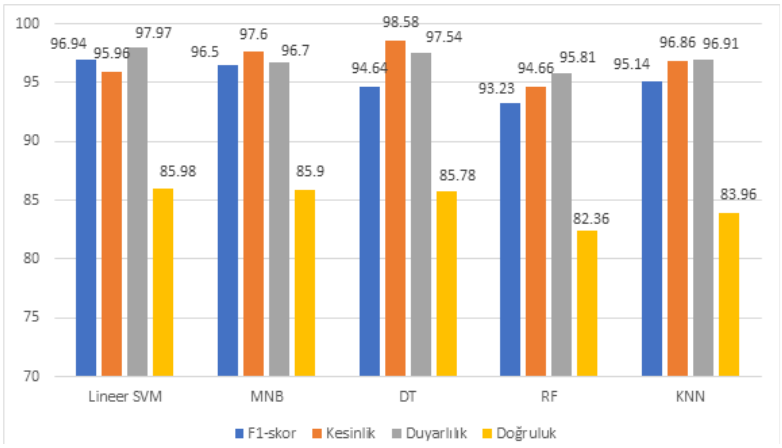


Şekil 7.1 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Azerbaycan dili ile hazırlanmış veri seti)



Şekil 7.2 Tf-Idfectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Azerbaycan dili ile hazırlanmış veri seti)

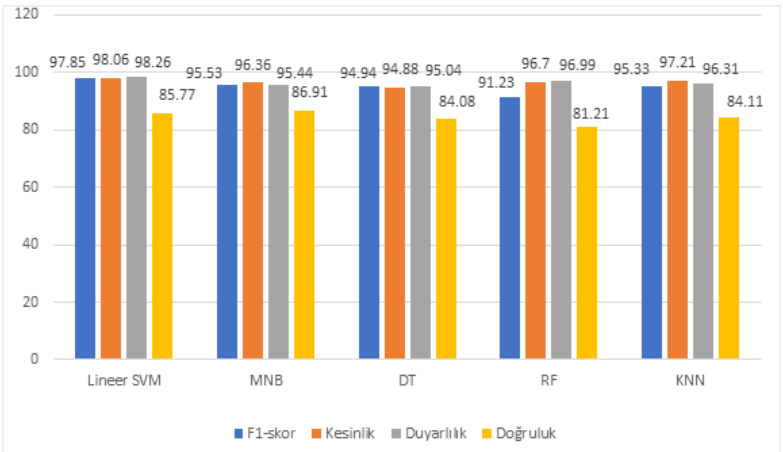
Çalışmada, Azerbaycan dili ile hazırlanmış veri setine göre CountVectorizer için %83.31 doğruluk ve %92.84 F1-skor ile Lineer SVM modeli en yüksek, %81.61doğruluk ve %90.32 F1-skor ile Rastgele Orman modeli en düşük sonuçlar vermiştir. Tf-IdfVectorizer için %84.52 doğruluk ve %93.91 F1-skor ile Lineer SVM modeli en yüksek, %81.61doğruluk ve %90.32 F1-skor ile Rastgele Orman modeli en düşük sonuçlar vermiştir.



Şekil 7.3 CountVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Türkçe ile hazırlanmış veri seti)

Türkçe ile hazırlanmış veri setine göre CountVectorizer için %85.98 doğruluk ve %96.94 F1-skor ile Lineer SVM modeli en yüksek, %82.36 doğruluk ve %93.23 F1-skor ile Rastgele Orman modeli en düşük sonuçlar vermiştir. Tf-IdfVectorizer için %85.77 doğruluk ve %97.85 F1-skor ile Lineer SVM modeli en yüksek, %81.21 doğruluk ve %91.23 F1-skor ile Rastgele Orman modeli en düşük sonuçlar vermiştir. Türkçe metinlerde siber zorbalığın tespitine yönelik geliştirilen sınıflandırma bağlamında;

- Araştırma bölümünde farklı dillerde yapılan araştırma sonuçları ile güncel sonuçlar karşılaştırıldığında türkçe olarak da anlamlı sonuçlar elde edilebilir.



Şekil 7.4 Tf-IdfVectorizer için Eğitim Modellerinde Makine Öğrenimi Modellerinin Değerlendirme Sonuçlarının yüzdesi (Türkçe ile hazırlanmış veri seti)

BÖLÜM 8
SONUÇ

Son zamanlarda sosyal yaşantımız insanların birçok konuda kendilerini anlatmalarına yardımcı olmuştur. Fakat ne yazık ki bazı kullanıcılar sosyal medyayı siber zorbalık benzeri eylemlerle iyi niyetle kullanmamakta, tanıdıkları veya tanımadıkları kullanıcılara karşı bazı küfürlü ve tehdit edici sözler kullanmaktadırlar. Bunun için çalışmamızda bu tür kötü niyetli eylemleri engellemek için makine öğrenme modelleri kullanıldı. Siber zorbalık tespiti için uluslararası alanda bizim de kullandığımız makine öğrenmesi algoritmaları kullanılarak, hatta sadece Türkçe olarak pek çok çalışma sunulmuştur. Bu çalışmada, Youtube'da Türkçe ve Azerbaycan dilinde hazırlanmış videolar için yazılmış yorumlar dikkate alınarak 5 farklı makine öğrenmesi algoritması kullanılarak siber zorbalık tespit edilmiştir. Bu bağlamda gelecekte yapılacak çalışmalara göre farklı sosyal platformalarından toplanan Azerbaycan dilinde ve Türkçede yapılan paylaşımlara yer verilerek kötü kelimeler için tanıma modelinin performansının artırılması aynı zamanda, çeşitli Makine öğrenimi yöntemleri ve Yapay sinir ağ modellerinin değerlendirilmesi hedeflenmektedir.

KAYNAKLAR

Aggarwal, C. C. (2018). Information retrieval and search engines. In *Machine Learning for Text* (pp. 259-304). Springer, Cham.

Agrawal, S., & Awekar, A. (2018). Deep learning for detecting cyberbullying across multiple social media platforms. In *European Conference on İnformation Retrieval* (pp. 141-153). Springer, Cham.

Aind, A. T., Ramnaney, A., & Sethia, D. (2020). Q-bully: a reinforcement learning based cyberbullying detection framework. In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-6). IEEE.

Akca, E. B., & Sayımer, İ. (2017). Cyberbullying, it's tyeps and related factors: an evaluation through the existing studies. *Online Academic Journal of Information Technology*, *8*(30), 7.

Altay, E. V., & Alatas, B. (2018). Detection of cyberbullying in social networks using machine learning methods. In *2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)* (pp. 87-91). IEEE.

Alsubait, T., & Alfageh, D. (2021). Comparison of machine learning techniques for cyberbullying detection on youtube arabic comments. *International Journal of Computer Science & Network Security*, *21*(1), 1-5.

Amarendra, B. D. (25 Mayıs 2022). *Countries where Cyber-bullying was Reported The Most In 2018.* https://ceoworld.biz/2018/10/29/countries-where-cyber-bullying-was-reported-the-most-in-2018/

Amanda., L. (25 Mayıs 2022). Cyberbullying. https://www.pewresearch.org/internet/2007/06/27/cyberbullying/.

Amy., M. (25 Mayıs 2022). *Cyberbullying Statistics Everyone Should Know.* https://www.verywellfamily.com/cyberbullying-statistics-4589988.

Arroyo-Fernández, I., Forest, D., Torres-Moreno, J. M., Carrasco-Ruiz, M., Legeleux, T., & Joannette, K. (2018). Cyberbullying detection task: the ebsi-lia-unam system (elu) at coling'18 trac-1. In *Proceedings of the first workshop on trolling, aggression and cyberbullying (TRAC-2018)* (pp. 140-149).

Aslan, A., & Doğan, B. Ö. (2017). Çevrimiçi şiddet: Bir siber zorbalık alanı olarak "Potinss" örneği. *Marmara İletişim Dergisi*, (27), 95-119.

Bag, A., Sahoo, M., & Mohanty, A. K. (2020). An assessment of Diagnosis and Prognosis of Breast Cancer Using Image Mining. *Journal of Engineering Research and Application*, 10(2), 2248-9622.

Biggs, D., De Ville, B., & Suen, E. (1991). A method of choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics*, *18*(1), 49-62.

Bozyiğit, A., Utku, S., & Nasiboğlu, E. (2018). Sanal zorbalık içeren sosyal medya mesajlarının tespiti. In *3rd International Conference on Computer Sciences and Engineering UBMK*.

Bozyiğit, A., Utku, S., & Nasiboğlu, E. (2019). Cyberbullying detection by using artificial neural network models. In *2019 4th International Conference on Computer Science and Engineering (UBMK)* (pp. 520-524). IEEE.

Bozyiğit, A., Utku, S., & Nasibov, E. (2021). Cyberbullying detection: Utilizing social media features. *Expert Systems with Applications*, *179*, 115001.

Chang, Y. W., & Lin, C. J. (2008). Feature ranking using linear SVM. In *Causation and prediction challenge* (pp. 53-64). PMLR.

Di Capua, M., Di Nardo, E., & Petrosino, A. (2016). Unsupervised cyber bullying detection in social networks. In *2016 23rd International Conference on Pattern Recognition (ICPR)* (pp. 432-437). IEEE.

Ditch the Label. (25 Mayıs 2022). *Cyberbullying Statistics: What They Tell Us.* https ://www.ditchthelabel.org/cyber-bullying-statistics-what-they-tell-us/

Do Something. (25 Mayıs 2022). *11 Facts About Cyberbullyıng.* https://www. dosomething.org/us/facts/11-facts-about-cyber-bullying.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis* (Vol. 3, pp. 731-739). New York: Wiley.

Enough Is Enough. (25 Mayıs 2022). *Cyberbullying Statistics.* https://enough.org/stats _cyberbullying.

Eroğlu, Y. (2014). *Ergenlerde siber zorbalık ve mağduriyeti yordayan risk etmenlerini belirlemeye yönelik bütüncül bir model önerisi* [Doktora Tezi]. Bursa Uludag University.

Eroğlu, Y., & Güler, N. (2015). Koşullu öz-değer, riskli internet davranışları ve siber zorbalık/mağduriyet arasındaki ilişkinin incelenmesi. *Sakarya University Journal of Education, 5*(3), 118-129.

Florida Atlantic University. (25 Mayıs 2022). *Nationwide teen bullying and cyberbullying study reveals significant issues impacting youth.* https:// May .com/releases/2017/02/170221102036.htm

Google. (25 Mayıs 2022). *Cyberbullying.* https://trends.google.com/trends/explore? q=cyberbullying

Grom Social. (25 Mayıs 2022). *Protect your kids from cyberbullying.* https://gromsocial.com/2019/07/22/protect-your-kids-from-cyberbullying/.

Hadžiosmanović, D., Simionato, L., Bolzoni, D., Zambon, E., & Etalle, S. (2012). N-gram against the machine: On the feasibility of the n-gram network analysis for binary protocols. In *International Workshop on Recent Advances in Intrusion Detection* (pp. 354-373). Springer, Berlin, Heidelberg.

Haidar, B., Chamoun, M., & Serhrouchni, A. (2019). Arabic cyberbullying detection: Enhancing performance by using ensemble machine learning. In *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 323-327). IEEE.

Hani, J., Nashaat, M., Ahmed, M., Emad, Z., Amer, E., & Mohammed, A. (2019). Social media cyberbullying detection using machine learning. *International Journal of Advanced Computer Science and Applications*, *10*(5), 703-707.

Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). *Automated machine learning: methods, systems, challenges* (p. 219). Springer Nature.

Isa, S. M., & Ashianti, L. (2017). Cyberbullying classification using text mining. *1st International Conference on Informatics and Computational Sciences* (ICICoS) 241-246.

Janikow, C. Z. (1998). Fuzzy decision trees: issues and methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *28*(1), 1-14.

Joseph., J (25 Mayıs 2022). *U.S. internet users who have experienced cyber bullying 2021*.https://www.statista.com/statistics/333942/us-internet-online-harassment-severity/.

Justin., W. P., & Sameer ., H. (25 Mayıs 2022). *Tween Cyberbullying In 2020*. https://cyberbullying.org/tween-statistics.

Karcioğlu, A. A., & Aydin, T. (2019). Sentiment analysis of Turkish and english twitter feeds using Word2Vec model. In *2019 27th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

Kanan, T., Aldaaja, A., & Hawashin, B. (2020). Cyber-bullying and cyber-harassment detection using supervised machine learning techniques in Arabic social media contents. *Journal of Internet Technology*, *21*(5), 1409-1421.

Kaşgarlı, K. M. (2021). *Twitter sentiment analysis via machine learning/Makine öğrenimi yoluyla twitter duygu analizi*. [Yüksek Lisans Tezi]. Kadir Has University.

Kumar., GN. C. (25 Mayıs 2022). *Machine Learning Types and Algorithms | Towards Data Science*. https://towardsdatascience.com/machine-learning-types-and-algorithms-d8b79545a6ec.

Kumari, K., Singh, J. P., Dwivedi, Y. K., & Rana, N. P. (2021). Multi-modal aggression identification using convolutional neural network and binary particle swarm optimization. *Future Generation Computer Systems*, *118*, 187-197.

Lin, Y., & Jeon, Y. (2002). Random Forests and Adaptive Nearest Neighbors, Technical Report No. 1055. *University of Wisconsin*.

Lewis, D. D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. In *European Conference on Machine Learning* (pp. 4-15). Springer, Berlin, Heidelberg.

McCallum, A., & Nigam, K. (1998, July). A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (Vol. 752, No. 1, pp. 41-48).

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Novalita, N., Herdiani, A., Lukmana, I., & Puspandari, D. (2019). Cyberbullying identification on twitter using random forest classifier. In *Journal of Physics: Conference Series* (Vol. 1192, No. 1, p. 012029). IOP Publishing.

Nasiboglu R., & Gencer, M. (2022). Comparison of Spacy and Stanford libraries' pre-trained deep learning models for named entity recognition, , *Journal of Modern Technology And Engineering* , 6(2), 104-111.

National Voices For Equality, Education and Enlightenment. (25 Mayıs 2022). *Bullying Statistics*. https://www.nveee.org/statistics/.

Office of Justice Programs. (25 Mayıs 2022). *Information and resources to curb the problem of cyberbullying*. https://www.ncpc.org/resources/cyberbullying/.

Ozigis, M. S., Kaduk, J. D., & Jarvis, C. H. (2019). Mapping terrestrial oil spill impact using machine learning random forest and Landsat 8 OLI imagery: a case site within the Niger Delta region of Nigeria. *Environmental Science and Pollution Research*, *26*(4), 3621-3635.

Öztürk, C. (2017). 8. Sınıf öğrencilerinin siber zorbalık eğilimlerinin çeşitli değişkenler açısından incelenmesi: sakarya ili adapazarı ilçesi örneği. *Journal of Interdisciplinary Education*: Theory and Practice 1(1):42-58.

Paul, S., & Saha, S. (2020). CyberBERT: BERT for cyberbullying identification. *Multimedia Systems*, 1-8.

Pawar, R., & Raje, R. R. (2019). Multilingual cyberbullying detection system. In *2019 IEEE İnternational Conference on Electro İnformation Technology (EIT)* (pp. 040-044). IEEE.

Pervan, N., & Keleş, Y. (2019). Derin öğrenme yaklaşımları kullanarak Türkçe metinlerden anlamsal çıkarım yapma. *Ankara Üniversitesi, Ankara, Türkiye*.

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, *4*(2), 1883.

Polat, Z. D., & Bayraktar, S. (2016). Ergenlerde siber zorbalık ve siber mağduriyet ile ilişkili değişkenlerin incelenmesi. *Mediterranean Journal of Humanities*, *5*(1), 115-132.

Potentia Analytics (25 Mayıs 2022). What is Machine Learning: Definition, Types, Applications And Examples | Potentia Analytics. https ://www .potentiaco .com/what-is-machine-learning-definition-types-applications and-examples/.

Rachid, B. A., Azza, H., & Ghezala, H. H. B. (2020). Classification of cyberbullying text in arabic. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-7). IEEE.

Ravichandran, K., & Arulchelvan, S. (2017). The model of multilayer perceptron analysed the crime news awareness in India (quantitative analysis method). In *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 1-6). IEEE.

Reportlinker. (25 Mayıs 2022). *Cyberbullying Statistics: For Young Generations, Standing Up to the bully is Now a Life Skill.* https:// www.reportlinker.com/insight /americas-youth-cyberbully-life-skill.html.

Rennie, J. D., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of naive bayes text classifiers. In *Proceedings of The 20th İnternational Conference on Machine Learning (ICML-03)* (pp. 616-623).

Reynolds, K., Kontostathis, A., & Edwards, L. (2011). Using machine learning to detect cyberbullying. In *2011 10th International Conference on Machine learning and applications and workshops* (Vol. 2, pp. 241-244). IEEE.

Rezvani, N., Beheshti, A., & Tabebordbar, A. (2020). Linking textual and contextual features for intelligent cyberbullying detection in social media. In *Proceedings of the 18th International Conference on Advances in Mobile Computing & Multimedia* (pp. 3-10).

Rosa, H., Pereira, N., Ribeiro, R., Ferreira, P. C., Carvalho, J. P., Oliveira, S., ... & Trancoso, I. (2019). Automatic cyberbullying detection: A systematic review. *Computers in Human Behavior*, *93*, 333-345.

Sadigzade, M., & Nasibov, E. (2022) Comparative analysis of count vectorization vs Tf-Idf vectorization for detecting cyberbullying in Turkish twitter messages, *Journal of Modern Technology And Engineering* , 7(1), 5-19.

Sam., C. (25 Mayıs 2022). *This list of cyberbullying statistics from 2018-2022 is regularly updated with the latest facts, figures, and trends.* https://www. comparitech.com/internet-providers/cyberbullying-statistics/

Sarkar, D. (2019). *Text analytics with Python: a practitioner's guide to natural language processing*. Bangalore: Apress.

Severyn, A., & Moschitti, A. (2012). Fast support vector machines for convolution tree kernels. *Data mining and knowledge discovery*, *25*(2), 325-357.

Shukan, A., Abdizhami, A., Ospanova, G., & Abdakimova, D. (2019). Issues of information technology crime control in the republic of Turkey. *Informatologia*, *52*(1-2), 65-73.

Song, T. M., & Song, J. (2021). Prediction of risk factors of cyberbullying-related words in Korea: Application of data mining using social big data. *Telematics and Informatics*, *58*, 101524.

Soni, D., & Singh, V. K. (2018). See no evil, hear no evil: Audio-visual-textual cyberbullying detection. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1-26.

Sugandhi, R., Pande, A., Agrawal, A., & Bhagat, H. (2016). Automatic monitoring and prevention of cyberbullying. *International Journal of Computer Applications*, *8*, 17-19.

Şafak, H. İ. (25 Mayıs 2022). *Makine Öğrenmesi Nedir? (What is Machine Learning?)|Medium.com.* https://medium .com/t% C3%B Crkiye/makine-%C3 %B6 %C4 %9 Frenmesi-nedir-20dee450b56e

Telenor. (25 Mayıs 2022). *Asia's parents speak up on cyberbullying*. https://www. telenor.com/asias-parents-speak-up-on-cyberbullying/.

Thomas W. Edgar, David O. Manz, in Research Methods for Cyber Security, (25 Mayıs 2022). Machine Learning | Science Direct. https://www.sciencedirect.com/ topics/computer-science/machinelearning#:~:text=Machine%20learning%20is%20 a%20 field,statistics%20and%20artificial%20intelligences%20communities.

Helen., W. (25 Mayıs 2022). *UNICEF poll: More than a third of young people in 30 countries report being a victim of online bullying*. https://www.unicef.org/press-releases/unicef-poll-more-third-young-people-30-countries-report-being-victim-online-bullying.

Ünver, H., & Zihni, K. O. Ç. (2017). Siber zorbalık ile problemli internet kullanımı ve riskli internet davranışı arasındaki ilişkinin incelenmesi. *Türk Eğitim Bilimleri Dergisi*, *15*(2), 117-140.

Webb, G. I., & Pazzani, M. J. (1998, July). Adjusted probability Naive Bayesian induction. In *Australian Joint Conference on Artificial Intelligence* (pp. 285-295). Springer, Berlin, Heidelberg.

Weru, T., Sevilla, J., Olukuru, J., Mutegi, L., & Mberi, T. (2017). Cyber-smart children, cyber-safe teenagers: Enhancing internet safety for children. In *2017 IST-Africa Week Conference (IST-Africa)* (pp. 1-8). IEEE.

Yazğili, E., & Baykara, M. (2021). Siber Zorbalık Tespit Yöntemleri Potansiyel Uygulama Alanları ve Zorluklar. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, *12*(1), 23-35.

Yuvaraj, N., Srihari, K., Dhiman, G., Somasundaram, K., Sharma, A., Rajeskannan, S., & Masud, M. (2021). Nature-inspired-based approach for automated cyberbullying classification on multimedia social networking. *Mathematical Problems in Engineering*, *2021*.

Zhao, Z., Gao, M., Luo, F., Zhang, Y., & Xiong, Q. (2020). LSHWE: improving similarity-based word embedding with locality sensitive hashing for cyberbullying detection. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.