

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**TÜRKÇE ÜRÜN YORUMLARI VERİSİ İLE DUYGU
ANALİZİ**

Buğra POLAT

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**ANKARA
2021**

Her hakkı saklıdır

ÖZET

Yüksek Lisans Tezi

TÜRKÇE ÜRÜN YORUMLARI VERİSİ İLE DUYGU ANALİZİ

Buğra POLAT

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr.Öğr.Üyesi Bülent TUĞRUL

Günümüzde sosyal medya platformları, alışveriş siteleri, forum ve blog gibi ortamlarda kullanıcılar her türlü ürün, hizmet veya durum hakkında düşüncelerini kolaylıkla paylaşabilmektedir. Paylaşılan bu içerikler analiz edilmesi gereken bir veri kaynağı oluşturmaktadır. Duygu analizinin amacı bu veri kaynağından anlamlı bilgilerin çıkarılmasını sağlamaktır.

Yapılan çalışmada, bir e-ticaret sitesindeki farklı kategorilerden ürünlere ait Türkçe yorumları içeren veri seti kullanılmıştır. Her bir yorumun duygu durumunun olumlu veya olumsuz olarak belirlenmesi hedeflenmiştir. Son yıllarda duygu analizinde sıklıkla kullanılan Konvolüsyonel Sinir Ağı (CNN), Uzun Kısa Süreli Bellek (LSTM) ve Geçitli Tekrarlayan Birim (GRU) derin öğrenme yöntemleri ile modeller geliştirilmiştir. FastText tarafından Skip-gram ve Continuous Bag of Words (CBOW) yöntemleri ile oluşturulan önceden eğitilmiş kelime vektörlerinden faydalanılarak bu yaklaşımların modellerin performansına etkisi karşılaştırılmıştır. Deneysel sonuçlar, Türkçe metinlerin duygu analizinde GRU yöntemi ile oluşturulan modellerin diğer modellerden daha başarılı sonuç verdiğini göstermiştir.

Haziran 2021, 48 Sayfa

Anahtar Kelimeler: Duygu Analizi, Fikir Madenciliği, Veri Madenciliği, Metin Sınıflandırma, Metin Öznitelikleri, Derin Öğrenme

ABSTRACT

Master Thesis

SENTIMENT ANALYSIS USING TURKISH PRODUCT REVIEW DATA

Buğra POLAT

Ankara University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Dr.Öğr.Üyesi Bülent TUĞRUL

Nowadays, users can easily share their thoughts about all kinds of products, services or situations in social media platforms, shopping sites, forums or blogs. These shared contents create a data source that needs to be analyzed. The purpose of sentiment analysis is to extract meaningful information from this data source.

In this study, a dataset containing Turkish reviews of products from different categories on an e-commerce website was used. It was aimed to determine the sentiment of each review as positive or negative. Models have been developed with Convolution Neural Network (CNN), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) deep learning methods which are frequently used in sentiment analysis in recent years. Pre-trained word vectors created by FastText with Skip-gram and Continuous Bag of Words (CBOW) methods were used and the effects of these approaches on the performance of the models were compared. The experimental results show that the models created with the GRU method in the sentiment analysis of Turkish texts achieved more effective results than other models.

June 2021, 48 Pages

Key Words: Sentiment Analysis, Opinion Mining, Data Mining, Text Classification, Text Features, Deep Learning

TEŐEKKÖR

Bilgi ve önerileri ile alıőmalarımı yönlendiren danıőman hocam sayın Dr.Öğr.Üyesi Bülent TUĞRUL'a, alıőmalarım süresince desteklerini esirgemeyen aile ve arkadaşlarıma içtenlik ile teşekkür ederim.

Buğra POLAT
Ankara, Haziran 2021

İÇİNDEKİLER

| | |
|---|-------------|
| TEZ ONAY SAYFASI | |
| ETİK | i |
| ÖZET | ii |
| ABSTRACT | iii |
| TEŞEKKÜR | iv |
| KISALTMALAR DİZİNİ | vi |
| ŞEKİLLER DİZİNİ | vii |
| ÇİZELGELER DİZİNİ | viii |
| 1. GİRİŞ | 1 |
| 2. KAYNAK ÖZETLERİ | 3 |
| 3. KURAMSAL TEMELLER | 6 |
| 3.1 Duygu Analizi | 6 |
| 3.2 Duygu Analizi Yöntemleri | 8 |
| 3.3 Kelime Temsili | 11 |
| 3.4 Derin Öğrenme | 13 |
| 3.4.1 Aktivasyon fonksiyonları | 15 |
| 3.4.2 Kayıp fonksiyonları | 17 |
| 3.4.3 Optimizasyon fonksiyonları | 20 |
| 3.4.4 Aşırı uyum ve yetersiz uyum | 21 |
| 3.4.5 Konvolüsyonel sinir ağları (CNN) | 22 |
| 3.4.6 Tekrarlayan sinir ağları (RNN) | 24 |
| 4. MATERYAL VE YÖNTEM | 28 |
| 4.1 Veri Seti | 28 |
| 4.2 Ön İşlemler | 29 |
| 4.3 Kelime Gömme (Word Embedding) | 30 |
| 4.4 Uygulama Modelleri | 30 |
| 4.4.1 CNN modeli | 31 |
| 4.4.2 LSTM ve GRU modeli | 31 |
| 4.5 Performans Metrikleri | 32 |
| 5. BULGULAR VE TARTIŞMA | 34 |
| 6. SONUÇ | 42 |
| KAYNAKLAR | 43 |
| ÖZGEÇMİŞ | 48 |

KISALTMALAR DİZİNİ

| | |
|----------|---|
| BERT | Bidirectional Encoder Representations from Transformers |
| BOW | Bag of Words (Kelime Çantası) |
| CBOW | Continuous Bag of Words (Sürekli Kelime Çantası) |
| CharSCNN | Character to Sentence Convolutional Neural Network |
| CNN | Convolution Neural Network (Konvolüsyonel Sinir Ağı) |
| DBN | Deep Belief Network (Derin İnanç Ağı) |
| DNN | Deep Neural Network (Derin Sinir Ağı) |
| FAIR | Facebook AI Research |
| Forex | Foreign Exchange (Döviz) |
| GRU | Gated Recurrent Unit (Geçitli Tekrarlayan Birim) |
| GloVe | Global Vectors for Word Representation (Kelime Temsili için Global Vektörler) |
| IMDb | Internet Movie Database |
| KNN | K-Nearest Neighbour (K-En Yakın Komşu) |
| LR | Logistic Regression (Lojistik Regresyon) |
| LSTM | Long Short-Term Memory (Uzun Kısa Süreli Bellek) |
| MAE | Mean Absolute Error |
| ME | Maximum Entropi (Maksimum Entropi) |
| MNB | Multinomial Naive Bayes |
| MSE | Mean Squared Error |
| NB | Naive Bayes |
| NLTK | Natural Language Toolkit (Doğal Dil Araç Seti) |
| NN | Neural Network (Sinir Ağı) |
| RELU | Rectified Linear Units (Doğrultulmuş Doğrusal Birimler) |
| RF | Random Forest (Rastgele Orman) |
| RNN | Recurrent Neural Network (Tekrarlayan Sinir Ağı) |
| RecNN | Recursive Neural Network (Özyinelemeli Sinir Ağı) |
| SCNN | Sentence Convolutional Neural Network |
| SGD | Stochastic Gradient Descent |
| SMO | Sequential Minimal Optimization (Sıralı Minimal Optimizasyon) |
| SVM | Support Vector Machine (Destek Vektör Makinesi) |
| TC-LSTM | Target-Connection LSTM |
| TD-LSTM | Target-Dependent LSTM |
| TF-IDF | Term Frequency-Inverse Document Frequency |
| TÜİK | Türkiye İstatistik Kurumu |

ŞEKİLLER DİZİNİ

| | |
|--|----|
| Şekil 3.1 Duygu analizi yöntemleri..... | 8 |
| Şekil 3.2 Yapay zeka, makine öğrenmesi ve derin öğrenme zaman çizelgesi..... | 10 |
| Şekil 3.3 Makine öğrenmesi ve derin öğrenme karşılaştırması | 10 |
| Şekil 3.4 CBOW ve Skip-gram modelleri..... | 12 |
| Şekil 3.5 Bir yapay sinir ağı modeli..... | 14 |
| Şekil 3.6 Bir nöron modeli | 14 |
| Şekil 3.7 Sigmoid | 15 |
| Şekil 3.8 Tahn | 15 |
| Şekil 3.9 ReLU | 16 |
| Şekil 3.10 SoftMax | 16 |
| Şekil 3.11 Huber kayıp fonksiyonu..... | 18 |
| Şekil 3.12 Binary Cross Entropy kayıp fonksiyonu..... | 19 |
| Şekil 3.13 Yetersiz uyum, ideal uyum ve aşırı uyum | 22 |
| Şekil 3.14 Konvolüsyonel sinir ağı modeli | 22 |
| Şekil 3.15 Maksimum havuzlama ve ortalama havuzlama yöntemi..... | 23 |
| Şekil 3.16 Bir tam bağlı katman modeli..... | 23 |
| Şekil 3.17 Bir tekrarlayan sinir ağı yapısı | 24 |
| Şekil 3.18 RNN modelleri..... | 24 |
| Şekil 3.19 LSTM hücresi | 25 |
| Şekil 3.20 GRU | 26 |
| Şekil 4.1 Veri setinin kelime bulutu olarak gösterimi | 29 |
| Şekil 4.2 CNN model mimarisi | 31 |
| Şekil 4.3 LSTM ve GRU model mimarisi | 32 |
| Şekil 4.4 Karar matrisi | 32 |
| Şekil 5.1 Duygu sınıflandırması karar matrisi (dengesiz dağılımlı veri)..... | 34 |
| Şekil 5.2 Duygu sınıflandırması karar matrisi (dengeli dağılımlı veri) | 35 |
| Şekil 5.3 (A) CNN+FastText (Skip-gram) doğruluk grafiği, (B) CNN+FastText (Skip-gram) kayıp grafiği, (C) LSTM+FastText (Skip-gram) doğruluk grafiği, (D) LSTM+FastText (Skip-gram) kayıp grafiği, (E) GRU+FastText (Skip-gram) doğruluk grafiği, (F) GRU+FastText (Skip-gram) kayıp grafiği..... | 38 |
| Şekil 5.4 (A) CNN+FastText (CBOW) doğruluk grafiği, (B) CNN+FastText (CBOW) kayıp grafiği, (C) LSTM+FastText (CBOW) doğruluk grafiği, (D) LSTM+FastText (CBOW) kayıp grafiği, (E) GRU+FastText (CBOW) doğruluk grafiği, (F) GRU+FastText (CBOW) kayıp grafiği..... | 40 |

ÇİZELGELER DİZİNİ

| | |
|--|----|
| Çizelge 2.1 Duygu analizi çalışmaları ve kullanılan yöntemler tablosu | 5 |
| Çizelge 4.1 Veri setinde yer alan yorumların skor dağılımı | 28 |
| Çizelge 5.1 Paket boyutunun CNN, LSTM ve GRU modellerinin performansına etkisi | 36 |
| Çizelge 5.2 Tur sayısının FastText (Skip-gram) kelime vektörleri kullanan modellerin performansına etkisi | 37 |
| Çizelge 5.3 Tur sayısının FastText (CBOW) kelime vektörleri kullanan modellerin performansına etkisi | 39 |
| Çizelge 5.4 CNN, LSTM ve GRU eğitim süreleri | 41 |

1. GİRİŞ

Günümüzde internet kullanımının yaygınlaşması ile her türlü işlem internet üzerinden yapılır duruma geldi. Kullanıcılar istedikleri zaman istedikleri yerde interneti diğer kullanıcılar ile iletişime geçmek, alışveriş yapmak, sosyal medya platformlarında zaman geçirmek, oyun oynamak, bankacılık işlemleri vb. pek çok farklı işlem için kullanabiliyor ve kendi deneyimlerini bu platformlardan paylaşabiliyor. Bu şekilde bir kullanıcının paylaşımı başka bir kullanıcı için bilgi kaynağı durumuna geliyor. Örnek olarak bir kullanıcı alışveriş sitesinden alışveriş yapmak istediğinde, almak istediği ürün hakkında o ürünü daha önce kullanmış kişilerin düşünceleri bilmek istiyor ve bu durum karar verme aşamasında önemli bir rol oynuyor (Pang ve Lee 2008). İstatistikler de bu durumu doğrular niteliktedir, TÜİK verilerine göre 2018'in son üç ayında kullanıcıların %67,8'i interneti mal ve hizmetler hakkında bilgi almak için kullanmıştır (Anonim 2019b).

Kullanıcı yorumlarının analiz edilerek faydalı bilginin elde edilmesi sadece bireysel kullanıcılar için değil aynı zamanda kuruluş ve organizasyonlar için de önemlidir. Büyük firmalar kullanıcı yorumlarından yola çıkarak ürün veya hizmetleri hakkında değiştirme veya geliştirme yaparlar. Mobil uygulama geliştiricileri, kullanıcı yorumlarından uygulama hataları hakkında bilgi sahibi olurlar ve gerekli düzeltmeleri yaparlar. Siyasi partilerin seçim stratejisi belirlemek için seçmenlerin görüşlerini öğrenmek istemesi de sosyal medya platformlarındaki kullanıcı yorumlarının analizini gerektirmektedir. Tüm bu örnekler veri analizinin büyük bir kitle tarafından ihtiyaç duyulan bir yöntem olduğunu göstermektedir.

Verinin analizi ile metinlerdeki ana düşüncenin bulunması ve sınıflandırılması işlemi duygu analizinin konusunu oluşturur. Yapılan çalışmada e-ticaret sitelerindeki ürünler hakkında yapılan yorumlar ile duygu analizi yapılmış ve sonuçlar paylaşılmıştır. Duygu analizinde derin öğrenme yöntemlerinden faydalanılmıştır.

Tezin genel organizasyonunu şu şekilde özetlenebilir. Bölüm 2'de daha önce duygu analizi ile yapılan çalışmalar ve yöntemleri açıklanmıştır. Bölüm 3'te duygu analizi

tanımı yapılmıř ve duygu analizinde gnmze kadar kullanılan yntemlerden bahsedilmiřtir. Blm 4’de yapılan alıřmada kullanılan yntemler aıklanmıř ve elde edilen deneysel sonular tartıřılmıřtır. Son blmde ise duygu analizi alıřmasından elde edilen sonu bilgisi zetlenmiř ve ileride yapılabilecek alıřmalar iin nerilerde bulunulmuřtur.

2. KAYNAK ÖZETLERİ

Duygu analizi ile ilgili geçmişten günümüze birçok farklı yöntem ile çalışma yapılmıştır. Önceden yapılan çalışmaların birçoğunda makine öğrenmesi teknikleri kullanılmaktaydı ancak derin öğrenmenin uygulanabilir duruma gelmesi ile günümüzde birçok çalışmada derin öğrenme teknikleri kullanılmaktadır. Bu bölümde her iki yaklaşıma da ait incelemiş olduğumuz çalışma örneklerinden bahsedilmektedir. Bu çalışmalarda kullanılan yöntemler ile elde edilen sonuçlar Çizelge 2.1’de yer almaktadır.

Makine öğrenmesi yöntemleri ile çalışma yapan Pang vd. (2002) Naive Bayes (NB), Maksimum Entropi (ME) ve Destek Vektör Makinesi (SVM) sınıflandırmalarını kullanarak film yorumları veri setinde duygu analizi yapmış, iyi sonucu SVM ile elde etmişlerdir. Bu çalışmanın sonucunda, film yorumlarının temel konu ile ilişkili olup olmadığının tespit etmenin duygu analizinde faydalı olacağını düşünmüşlerdir. Bu düşünce ile birlikte Pang ve Lee bir diğer çalışmalarında metinlerin öznel bölümlerine odaklanan bir yöntem geliştirmişlerdir (Pang ve Lee 2004).

Go vd. (2009) tarafından Twitter veri seti üzerinde NB, ME ve SVM algoritmalarını kullanarak duygu analizi gerçekleştirilmiştir. Diğer yöntemlerden farklı olarak duygu ifadesi içeren tweetler toplanmış ve içerdikleri ifadelere göre etiketlenerek eğitim verisi oluşturulmuştur. Elde edilen verinin analizi sonucunda ise %80 üzerinde doğruluk değeri elde edilmiştir.

Dos Santos ve Gatti (2014) Twitter veri seti üzerinde Character to Sentence Convolutional Neural Network (CharSCNN) ve Sentence Convolutional Neural Network (SCNN) yöntemini kullanmış ve karakter düzeyindeki bilginin veri analizinde büyük bir etkisinin olduğunu tespit etmişlerdir.

Türkmenoğlu ve Tantuğ (2014) kısa metinlerden oluşan Twitter ve film yorumları veri setini kullanarak makine öğrenmesi temelli ve sözlük temelli yaklaşımların

karşılaştırmasını yapmışlar. Her iki veri setinde de makine öğrenmesi temelli yaklaşımın daha iyi sonuç verdiği gözlemlenmiştir. Her iki veri setinde de en iyi sonuç SVM algoritması ile elde edilmiştir.

Çoban vd. (2015) sosyal medya veri setinde N-Gram ve Kelime Çantası (BOW) yöntemleri ile öznitelik çıkarımı modellerini kullanmışlar. Her iki model için NB, Multinomial Naive Bayes (MNB), SVM ve K-En Yakın Komşu (KNN) algoritmaları ile duygu analizi yapmış ve MNB algoritması ile en iyi sonucu elde etmişlerdir.

Dehkharghani vd. (2015) Türkçe film yorumları veri setinde Lojistik Regresyon (LR) yöntemi ile cümle ve doküman seviyesinde duygu analizi yapmış ve doküman seviyesinde daha yüksek başarı elde etmişlerdir.

Tang vd. (2015) hedefe bağlı iki yeni Uzun Kısa Süreli Bellek (LSTM) modeli geliştirmiş ve yöntemlerini Twitter veri seti üzerinde uygulamışlardır. Elde ettikleri sonuçlar ile hedefe bağlı LSTM modellerinin LSTM'den daha başarılı olduğunu ortaya koymuşlardır.

Ayata vd. (2017) SVM ve Rastgele Orman (RF) yöntemlerini Twitter veri seti üzerinde uygulamış ve duygu analizinde SVM'nin RF'den daha iyi sonuç verdiğini gözlemlemişlerdir. Bunun yanı sıra veri setinin sektör temelli olmasının elde edilen başarı oranını arttırdığı sonucunu ortaya koymuşlardır.

Karakuş vd. (2018) Türkçe film yorumları veri setini kullanarak Konvolüsyonel Sinir Ağı (CNN), LSTM ve bu iki yöntemin kombinasyonları ile duygu analizi yapmışlar ve sonuçlarını paylaşmışlardır. Ek olarak veri setleri üzerinde Word2Vec tekniğini uygulamış ve elde ettikleri sonuçlar ile kelime gömme kullanımının derin öğrenme modelleri üzerindeki etkilerini tartışmışlardır.

Pervan (2019) e-ticaret sitelerinden elde ettiği Türkçe veri setinde RF, LSTM ve D-CNN yöntemlerini uygulamıştır. En iyi sonucu 94.21% doğruluk değeri ile LSTM yönteminde elde etmiştir.

Yang vd. (2020) bir e-ticaret sitesinden elde ettikleri Çince veri seti ve çalışmalarında ileri sürdükleri SLCABG modelini kullanarak 93.5% doğruluk değeri elde etmişlerdir. SLCABG modelini duygu sözlüğü, derin öğrenme ve dikkat mekanizmasının güçlü yönlerinden faydalanarak oluşturmuşlardır.

Çizelge 2.1 Duygu analizi çalışmaları ve kullanılan yöntemler tablosu

| Referans | Veri Seti | Dil | Kullanılan Modeller | En iyi Performans | |
|------------------------------|-------------------|-----------|-----------------------------|-------------------|--------------|
| | | | | Model | Doğruluk (%) |
| Pang vd. (2002) | Film Yorumları | İngilizce | NB, ME, SVM | SVM | 82.9 |
| Pang ve Lee (2004) | Film Yorumları | İngilizce | NB, SVM | SVM | 87.15 |
| Go vd. (2009) | Twitter Mesajları | İngilizce | NB, ME, SVM | ME | 83.0 |
| Dos Santos ve Gatti (2014) | Twitter Mesajları | İngilizce | CharSCNN, SCNN | CharSCNN | 86.4 |
| Türkmenoğlu ve Tantuğ (2014) | Film Yorumları | Türkçe | SVM, NB, J48 | SVM | 89.5 |
| | | | | NB | 89.5 |
| | Twitter Mesajları | Türkçe | SVM, NB, J48 | SVM | 85.0 |
| Tang vd. (2015) | Twitter Mesajları | İngilizce | LSTM, TD-LSTM, TC-LSTM | TC-LSTM | 71.5 |
| Çoban vd. (2015) | Twitter Mesajları | Türkçe | SVM, NB, MNB, KNN | KNN | 65.79 |
| Dehkharghani vd. (2015) | Film Yorumları | Türkçe | LR | LR | 79.06 |
| Ayata vd. (2017) | Twitter Mesajları | Türkçe | SVM, RO | SVM | 74.60 |
| Karakuş vd. (2018) | Film Yorumları | Türkçe | CNN, LSTM, BILSTM, CNN-LSTM | CNN-LSTM | 98.07 |
| Pervan (2019) | Ürün Yorumları | Türkçe | RO, LSTM, D-CNN | LSTM | 94.21 |
| Yang vd. (2020) | Ürün Yorumları | Çince | CNN, BiGRU, SLCABG | SLCABG | 93.5 |

3. KURAMSAL TEMELLER

3.1 Duygu Analizi

Günümüzde sosyal medyanın, blogların, forumların kullanımının artması ile internet ortamında çeşitli kaynakları içeren veri topluluğu oluşturmuştur. Gittikçe büyüyen ve geleneksel yazılım teknolojileri kullanılarak yönetilmesi ve analiz edilmesi zor olan bu dijital veri topluluğu büyük veri olarak adlandırılır (Sohangir vd. 2018). Büyük veri, veriden anlamlı kalıpların çıkarılması için özel yöntemlere ihtiyaç duyar, duygu analizi de bu tekniklerin bir bölümü kapsayan çalışma alanıdır.

Duygu analizi insanların bir konu, olay, durum vb. hakkındaki yorumlarından görüş, duygu, düşünce, değerlendirme ve tutumları analiz eden bir çalışma alanıdır. Aynı zamanda fikir madenciliği olarak da adlandırılır (Liu 2012) ve birçok alanda hedeflenen bilgiye ulaşmaya yardımcı olur. Bu alanlardan bazıları aşağıda açıklanmıştır:

Borsa tahmini: Borsa tahmini ile ilgili çalışmalar genel olarak önceki piyasa verileri ile yapılmaktadır. Ancak son yıllarda bu sayısal ve düzenli veri dışında haber metinleri gibi veriler de kullanılarak piyasa tahmini çalışmaları yapılmıştır. Nassirtoussi vd. (2015) tarafından haber başlıklarını metin madenciliği ile analiz edilmiş ve Forex piyasası tahmini yapılmıştır.

Gişe tahmini: Asur ve Huberman (2010) çalışmalarında Twitter verisini kullanmış ve duygu analizi yardımıyla gişe hasılatı tahmininde bulunmuşlardır.

Öneri sistemleri: Meehan vd. (2013) turizm alanıyla ilişkili, kullanıcıların ziyaret etmesi gereken mekanları öneren bir sistem oluşturmuşlardır. Bu öneri sisteminde konum, zaman, hava durumu gibi verilerin yanı sıra mekan ile ilgili turistlerin güncel tutumunun belirlenmesi için duygu analizinden faydalanılmıştır.

Siyasi konular: Duygu analizi genellikle siyasi hareketlere ilişkin kamuoyunun ruh halini gözetmede kullanılmaktadır. Wang vd. (2012) Twitter üzerinden gerçek zamanlı olarak 2012 U.S. başkanlık seçimine yönelik kamuoyu tutumunu gösteren bir sistem geliştirmişlerdir.

Soru Cevaplama: Somasundaran vd. (2007) soru cevaplama üzerine denemelerde bulunmuşlardır.

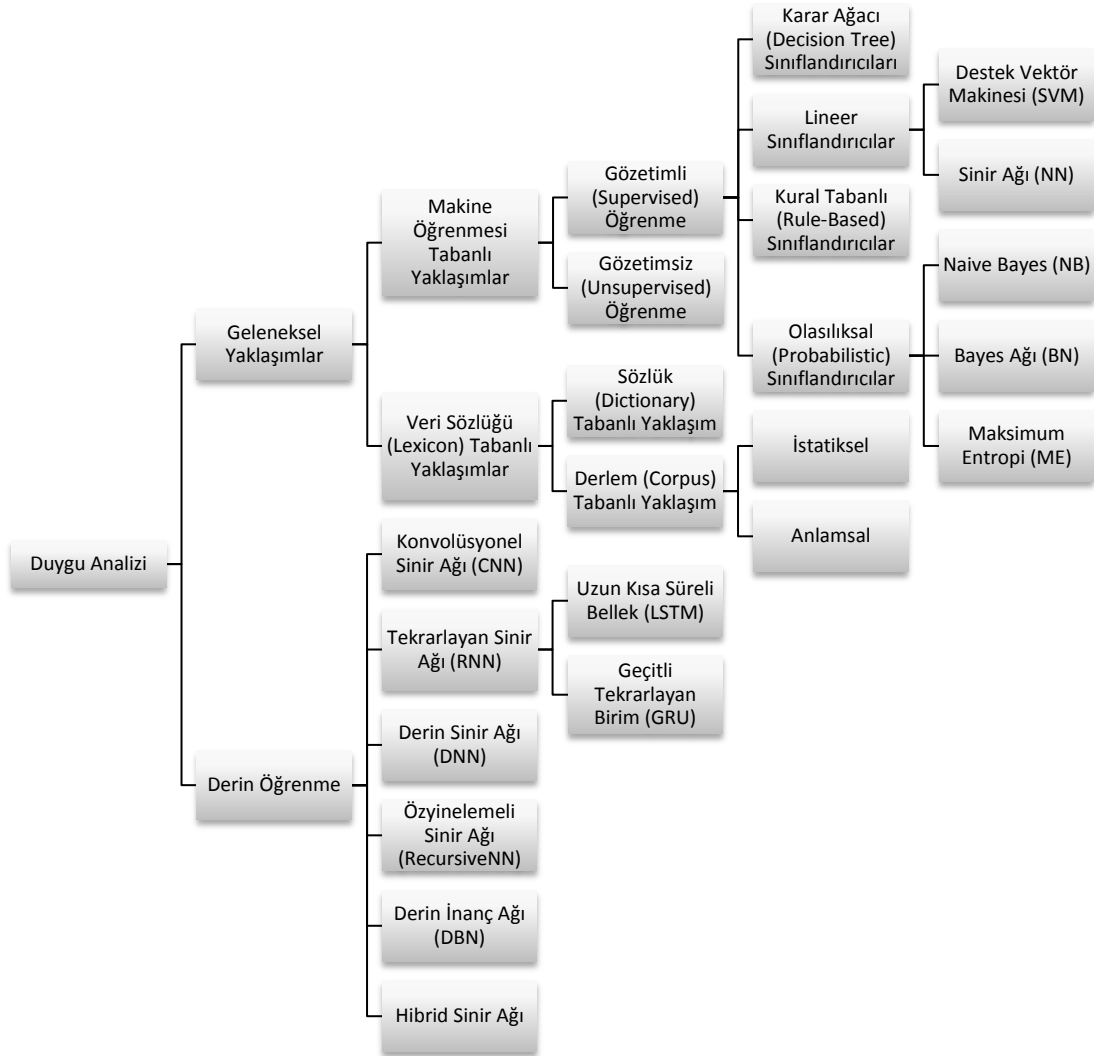
Özetleme: Zhuang vd. (2006) film yorumları veri seti ile metin madenciliği ve özetlenmesi üzerine bir sistem geliştirmişlerdir.

Duygu analizi doküman, cümle ve hedef tabanlı olmak üzere üç farklı seviyede incelenebilir. Doküman seviyesinde duygu analizinde tüm metin tek bir girdi olarak sisteme gönderilir ve olumlu veya olumsuz sonuç elde edilir. Metin içerisinde konudan bağımsız olarak yer alan görüşlerin de analiz sonucunu etkilemesi doküman seviyesindeki analizin problemlerinden biridir. Bir ileri seviyedeki inceleme biçimi ise cümle seviyesidir. Cümle seviyesinde daha kısa metinler üzerinden analiz yapılır. Ancak bazı karmaşık cümlelerin analizi problem olabilmektedir. “Yemekler güzeldi fakat servis kötüydü.” bu cümlede yemekler ve servis için iki farklı duygu ifadesi yer almaktadır. Burada hangi duygunun ele alınacağıın tespit edilmesi gerekir ve bu aşamada hedef tabanlı duygu analizi devreye girer. Hedef tabanlı duygu analizinde temel konu ile bağlantılı olan kelime veya kelime gruplarının bulunması amaçlanır.

Bir sonraki bölümde günümüze kadar yapılan çalışmalarda kullanılan temel yöntemler açıklanacaktır.

3.2 Duygu Analizi Yöntemleri

Duygu analizi yöntemlerini geleneksel yaklaşımlar ve derin öğrenme tabanlı yaklaşımlar olarak iki gruba ayrılabiliriz (Şekil 3.1).



Şekil 3.1 Duygu analizi yöntemleri (Medhat vd. 2014, Dang vd. 2020, Yadav ve Vishwakarma 2020)

Geleneksel yaklaşımlar makine öğrenmesi ve sözlük tabanlı yaklaşımlar olarak gruplandırılır. Makine öğrenmesi yapay zekanın bir alt kümesidir ve eldeki veriyi kullanarak öğrenme tabanlı sistemler geliştirmeye odaklanır. Temel yöntemi verileri

ayrıştırmak, çıkarılan veriden öğrenmeyi gerçekleştirmek ve daha sonra tespit veya tahmin yapmak olarak tanımlanabilir.

Makine öğrenmesi yöntemleri gözetimli ve gözetimsiz olarak iki gruba ayrılır. En yaygın kullanılanı gözetimli öğrenmedir. Gözetimli öğrenmede önceden sınıflandırılmış büyük miktarda veri kullanılarak algoritma eğitilir ve daha sonra test verisi üzerinden sonuç elde edilir. Naive Bayes, Maksimum Entropi ve Destek Vektör Makineleri gözetimli öğrenme algoritmalarına örnektir. Pang ve Lee tarafından bu yöntemleri Internet Movie Database (IMDb) veri kümesine uygulanmış ve duygu analizinde performans açısından SVM yönteminin en iyi sonucu gösterme eğiliminde olduğu gözlemlenmiştir (Pang vd. 2002).

Gözetimsiz makine öğrenmesi önceden sınıflandırılmış veriye ihtiyaç duymaz, veriler benzer yönleri bulunarak kümelenir ve buna göre işaretlenir. Bu yaklaşım Turney (2002) tarafından kullanılmıştır. Duygusal yönelimi gösteren sıfat ve zarf grupları çıkarılmış ve cümlede semantik yönelim bu kelime grupları ile “mükemmel” ve “zavallı” kelimeleri arasındaki ilişki doğrultusunda hesaplanmıştır.

Sözlük temelli yaklaşımlarda pozitiflik ve negatiflik değerleri belirli olan sözcüklerin manuel veya otomatik oluşturulmuş listesi hazırlanır ve daha sonra metinlerde kontrol edilecek her sözcüğün karşılığı bu listeden aranarak bu sözcükler için duygu değeri belirlenir (Taboada vd. 2011).

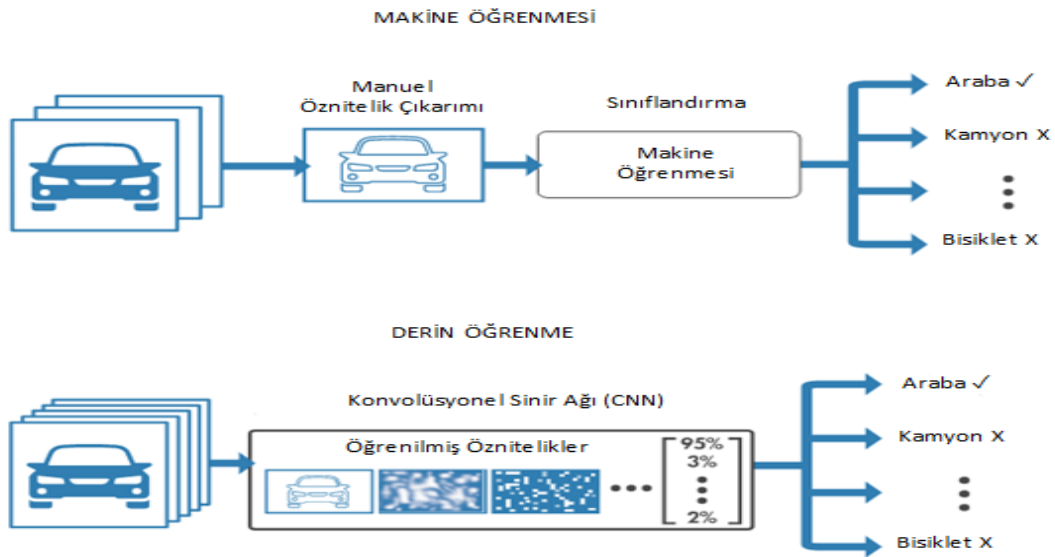
2010’lu yıllara kadar yapılan duygu analizi çalışmalarının çoğunluğunda makine öğrenmesi yöntemleri uygulanmıştır. Ancak sonrasında grafik işlemci birimi (GPU) etkisi ile paralel komutlar yüksek verim ile işlenebilir duruma gelmiş ve derin öğrenme algoritmaları popüler olmuştur (Şekil 3.2).



Şekil 3.2 Yapay zeka, makine öğrenmesi ve derin öğrenme zaman çizelgesi

Makine öğrenmesi yöntemlerinde performans doğru öznitelik çıkarımına bağlıdır ve bu işlem manuel olarak öznitelik seçme algoritmaları aracılığıyla yapılır. Akba vd. (2014) duygu analizi performansını iyileştirmek amacıyla öznitelik seçme yöntemleri üzerinde çalışmışlardır. Türkçe film yorumları veri setinde Information Gain ve Chi-Square öznitelik seçme yöntemlerinden faydalanmışlardır. Sınıflandırmada az sayıda öznitelik kullanımının başarıyı arttırdığını göstermişlerdir.

Makine öğrenmesinden farklı olarak derin öğrenme yöntemleri önceden belirlenmiş kurallara ihtiyaç duymaz ve resim, ses veya metin gibi verilerden otomatik olarak öğrenme gerçekleştirilir (Anonymous 2019a). Şekil 3.3'te makine öğrenmesi ve derin öğrenme yaklaşımının karşılaştırılması gösterilmektedir.



Şekil 3.3 Makine öğrenmesi ve derin öğrenme karşılaştırması (Anonymous 2019g)

Duygu analizi yöntemlerinin temelini sınıflandırma oluşturmaktadır ve sınıflandırmanın yapılabilmesi için metinlerin sayısal hale dönüştürülmesi gerekmektedir. Bir sonraki bölümde kelimelerin sayısal olarak temsil edilmesi için kullanılan yöntemlerden bahsedilecek ve sonrasında derin öğrenme ayrıntılı bir şekilde açıklanacaktır.

3.3 Kelime Temsili

Duygu analizi çalışmalarında sınıflandırmanın yapılabilmesi için metinlerin sayısal hale dönüştürülmesi gerekmektedir. Kelime temsili ile metin içerisindeki kelimeler sayısal veriler içeren vektörlere dönüştürülürler. Bu bölümde en popüler kelime temsil yöntemleri açıklanacaktır.

One-Hot Encoding kategorik değişkenlerin ikili kod ile temsil edilmesidir ve vektör temsili için kullanılan en basit yöntemlerden bir tanesidir. Metin içindeki her bir kelimeye karşılık gelen bit “1” değerlerinin ise “0” olarak kodlanması ise elde edilir. Vektör boyutu derlemdeki kelime sayısına bağlıdır. Bu yöntem metinde yer alan öğelerin sırası ve birbirleri ile olan bağlantısını yansıtmaz.

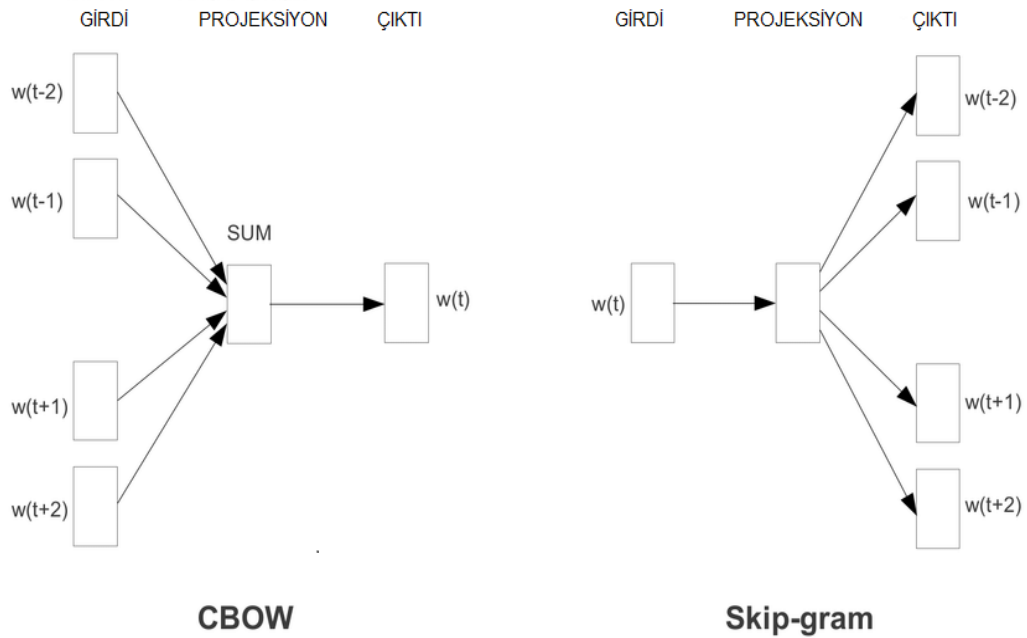
Bag of Words (BOW) modelinde metin içerisinde yer alan her kelimenin gözlenme sıklığı dikkate alınarak sayısal değeri hesaplanır. Bu yöntemde kelimelerin sırası ve yapısı göz ardı edilir.

Term Frequency-Inverse Document Frequency (TF-IDF) modeli bir kelimenin metin içerisindeki önemini ölçmek için kullanılan istatistiksel bir yöntemdir. Metinde bir terim ne kadar çok geçiyorsa o kadar az bilgi içerdiğini varsayılır ve bu nedenle durak kelimeler (stop words) gibi duygu analizine etkisi olmayan kelimelerin metinden çıkarılması amacı ile kullanılabilir.

Bu bölüme kadar açıkladığımız yöntemlerdeki ortak problem kelime temsiliinde anlamsal bir ilişki kuramıyor olmalıdır ve bu problem kelime gömme (word embedding) yaklaşımı ile çözümlenmiştir. Kelime gömme yönteminde metin içerisinde yer alan

kelimeler düşük boyutlu yoğun vektör uzayında temsil edilir ve benzer kelimeler benzer pozisyonlarda yer alırlar (Yang 2020). Word2Vec (Mikolov vd. 2013), Glove (Pennington vd. 2014) ve FastText (Joulin vd. 2016) sıklıkla kullanılan kelime gömme yaklaşımlarındandır.

Word2Vec Mikolov vd. (2013) tarafından geliştirilmiş her bir kelime için bağlamsal ve anlamsal benzerliği de ele alarak vektörsel bir gösterim oluşturmayı amaçlayan derin öğrenme tabanlı bir vektör temsil yöntemidir. Continuous Bag of Words (CBOW) ve Skip-gram olmak üzere iki modeli bulunmaktadır (Şekil 3.4).



Şekil 3.4 CBOW ve Skip-gram modelleri (Mikolov vd. 2013)

CBOW modelinde geçerli kelime etrafını çevreleyen kelimelere bağlı olarak tahmin edilir, Skip-gram modelinde ise geçerli kelimedden etrafını çevreleyen kelimelerin tahmini yapılır. Skip-gram modeli genel olarak küçük eğitim verisinde ve nadir kelimelerde daha iyi çalışırken, CBOW modeli Skip-gram modeline göre daha hızlı ve sık kullanılan kelimelerde daha iyi çalışmaktadır (Anonymous 2019d).

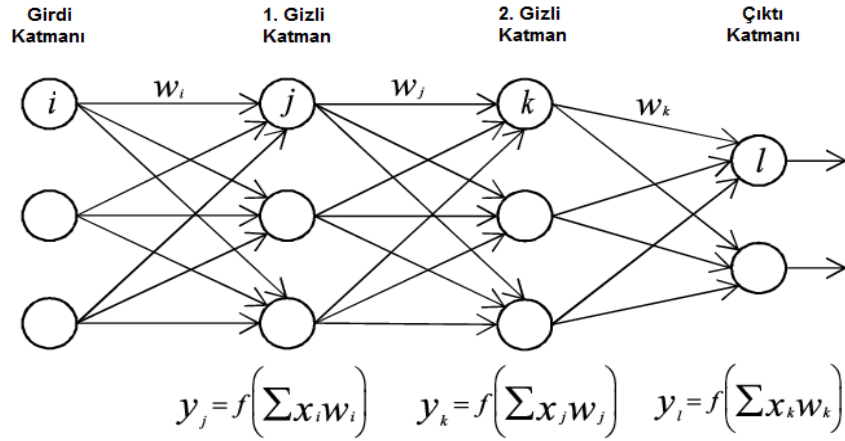
GloVe (Global Vectors for Word Representation), Pennington ve diğeri (2014) tarafından kelime vektör temsillerini elde etmek için word2vec yaklaşımına alternatif olarak sunulan bir yöntemdir. Kelime temsillerinin oluşturulmasında birlikte kullanılan kelimelerin istatistiklerinden faydalanır.

FastText Facebook AI Research (FAIR) tarafından geliştirilen bir kelime temsil yöntemidir. Bu model kelimeleri tek tek ele almak yerine n-gramlarına ayırır ve kelime vektörü n-gram vektörlerinin toplamı olacak şekilde temsil edilir. Geliştiriciler için birden fazla dilde CBOW veya Skip-gram yöntemleri ile önceden eğitilmiş kelime vektör modelleri sunulmaktadır.

3.4 Derin Öğrenme

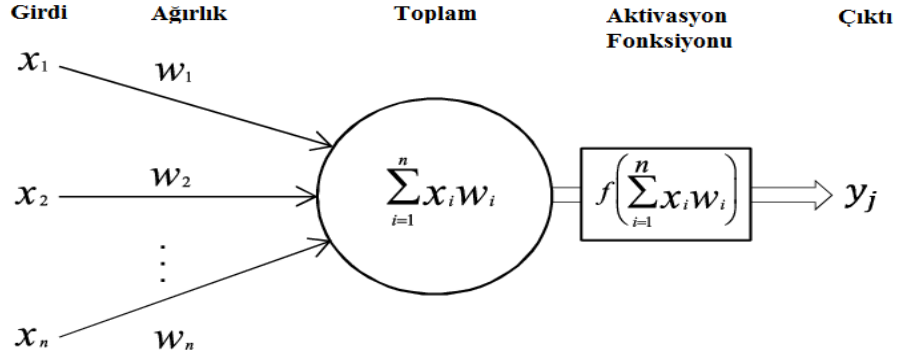
Derin öğrenme çok katmanlı sinir ağlarını kullanarak öğrenmenin gerçekleşmesini sağlayan bir yapay sinir ağı uyarlamasıdır ve makine öğrenmesinin bir alt kümesidir. Günümüzde nesne algılama, konuşma tanıma, dil çevirisi vb. birçok alanda derin öğrenme yöntemleri kullanılır. Gitgide büyüyen veri kümelerinin etkili bir şekilde analiz edilmesine duyulan ihtiyaç, işlemci becerilerinin artması ve maliyetinin düşmesi derin öğrenmenin tercih edilme nedenleri arasındadır. Konvolüsyonel sinir ağları (CNN) ve tekrarlayan sinir ağları (RNN) duygu analizinde en sık kullanılan ve en etkili sonucu veren derin öğrenme yöntemleri arasında gösterilebilir.

Derin öğrenme süreci çok sayıda bilgi işlem biriminden oluşan ve katmanlar halinde organize edilmiş bir yapay sinir ağı modeli (Şekil 3.5) ile yönetilir (Zhang vd. 2018). Katmanlı yapılar girdi katmanı, çıktı katmanı ve bu iki katman arasında bulunan gizli katman(lar) olarak sınıflandırılır. Her ardışık katman, önceki katmandaki çıktıyı girdi olarak alır ve sonraki katmanın kullanabileceği bilgilere dönüştürür.



Şekil 3.5 Bir yapay sinir ağı modeli (Anonymous 2017c)

Yapay sinir ağında her bir birim yapı taşı olan nöronu temsil eder (Şekil 3.6). Nöronda her X girdisi, W ağırlık değeri ile ilişkilendirilir. Tüm ağırlıklandırılmış girdilerin toplamı aktivasyon fonksiyonundan geçirilir ve çıktı bilgisi oluşturulur.



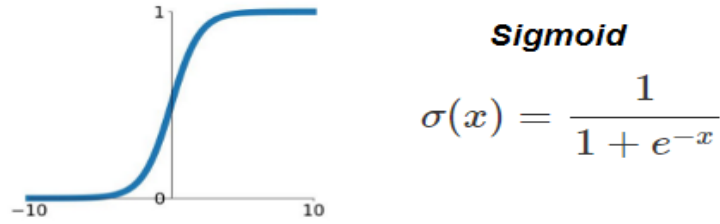
Şekil 3.6 Bir nöron modeli (Anonymous 2017c)

Oluşturulan çıktı değeri ile hedef değer karşılaştırılır ve kayıp fonksiyonu (loss function) yardımıyla hata değeri ölçülür. Hata değerinin düşürülmesi için optimizasyon fonksiyonundan faydalanılır. Elde edilen sonuç değerine göre her bir nörona ait parametreler tekrar ayarlanır ve geri yayılım (backpropagation) yöntemi kullanılarak hedef değere en yakın değere bulunmaya çalışılır.

3.4.1 Aktivasyon fonksiyonları

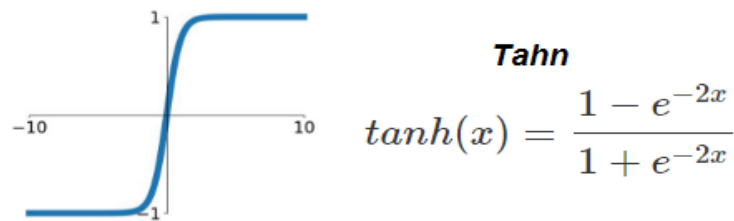
Toplama fonksiyonu sonucunda oluşan net girdiyi işleyerek hücre çıktısını belirleyen ve genellikle doğrusal olmayan fonksiyonlardır. Sigmoid, Tahn, RELU ve SoftMax en sık kullanılan aktivasyon fonksiyonlardır.

Sigmoid fonksiyonu $[0, 1]$ aralığında değer üreten, doğrusal olmayan bir aktivasyon fonksiyonudur (Şekil 3.7). Anlaşılır olması ve uygulanabilirliği ile popülerdir. Sigmoid fonksiyonunun kullanılmasının dezavantajı ise kaybolan gradyan problemine yol açmasıdır. Kaybolan gradyan problemi geri yayımlı ağ eğitimi sırasında hata fonksiyonunun değerinin zamanla küçülerek yok olması sonucunda meydana gelir. Eğitimi zorlaştıran bir durumdur ve çok katmanlı sinir ağlarında sıklıkla görülür.



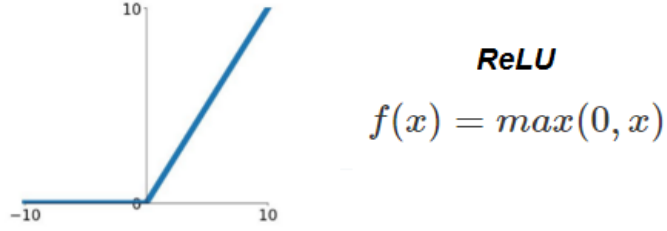
Şekil 3.7 Sigmoid (Anonymous 2018b)

Tahn fonksiyonu da doğrusal olmayan bir fonksiyondur ve $[-1, 1]$ aralığında değer üretir (Şekil 3.8). Kaybolan gradyan problemi Tahn fonksiyonunun kullanımında da görülmektedir ancak gradyanın Sigmoid fonksiyonuna göre daha güçlü olması tercih edilmesini sağlamaktadır.



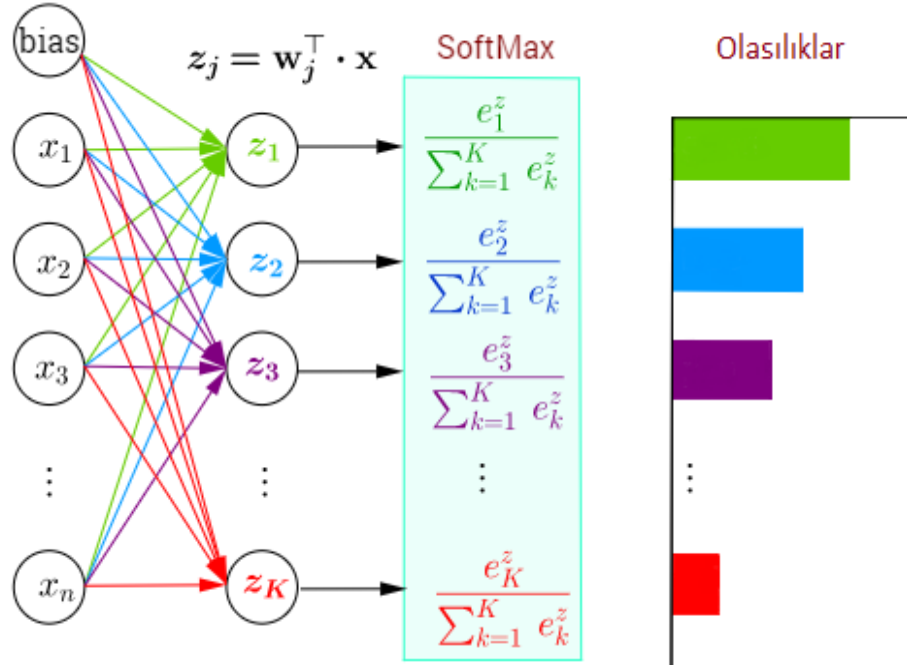
Şekil 3.8 Tahn (Anonymous 2018b)

ReLU kaybolan gradyan problemini önleyen, doğrusal olmayan bir aktivasyon fonksiyonudur (Şekil 3.9). 0'dan küçük değerler için çıktı değeri 0, diğer değerler için çıktı değeri girdiye eşittir. Hızlı sonuç veren ve uygulanabilirliği kolay bir fonksiyondur.



Şekil 3.9 ReLU (Anonymous 2018b)

SoftMax son yıllarda sıkça kullanılan özel bir aktivasyon fonksiyonudur (Şekil 3.10). Sigmoid fonksiyonu gibi $[0, 1]$ aralığında değer üreten bu fonksiyonun farkı üretilen sonuçların toplamının daima 1'e eşit olmasıdır. SoftMax fonksiyonu çoklu sınıflandırma problemleri için kullanılır.



Şekil 3.10 SoftMax (Anonymous 2017a)

3.4.2 Kayıp fonksiyonları

Kayıp fonksiyonu modelin performansını değerlendirme yöntemlerinden biridir ve yapılan tahmin ile gerçek sonuçlar arasındaki farkı gösterir. Eğer tahmin yanlışsa yüksek bir değer, değilse sifıra yakın bir değer verir. Öğrenme yönteminin türüne göre regresyon ve sınıflandırma kayıp fonksiyonları olarak iki kategoriye ayrılırlar.

Regresyon kayıp fonksiyonlarından en bilinenleri Mean Squared Error (MSE), Mean Absolute Error (MAE) ve Huber Loss fonksiyonlarıdır.

Mean Squared Error (MSE) gerçek değerler ve tahmin değerleri arasındaki farkın karesinin ortalaması alınarak hesaplanır. y_i gerçek değeri, \hat{y}_i ise tahmin değerini temsil eder.

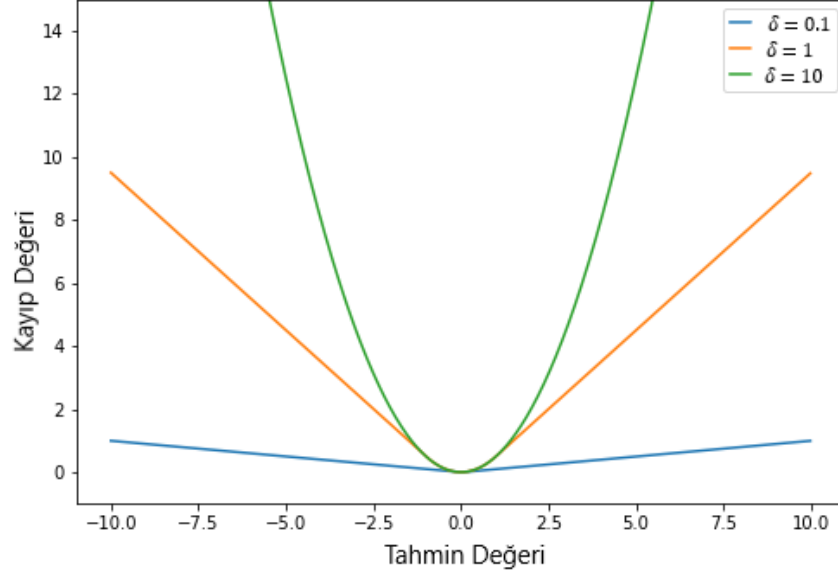
$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3.1)$$

Mean Absolute Error (MAE) gerçek değerler ve tahmin değerleri arasındaki farkların ortalaması alınarak hesaplanır. y_j gerçek değeri, \hat{y}_j ise tahmin değerini temsil eder.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (3.2)$$

Huber Loss, MSE ve MAE fonksiyonlarının avantajlı yönlerinin birleşimidir. MAE fonksiyonunda kayıp değerleri için eğimin aynıdır, bu durum küçük kayıp değerleri için de büyük sonuçların elde edilmesine neden olur. MSE fonksiyonu ise büyük değerler için yüksek bir sonuç verirken değerler küçüldükçe elde edilen sonuç sifıra yaklaşır. Bu problemlerin çözümü amacıyla Huber fonksiyonu kullanılır. Huber fonksiyonu eğer kayıp değeri delta değerine küçük veya eşitse MSE, kayıp değeri delta değerinden büyükse MAE fonksiyonu gibi davranır (Şekil 3.11).

$$L_{\delta}(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & \text{for } |y - f(x)| \leq \delta, \\ \delta |y - f(x)| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases} \quad (3.3)$$



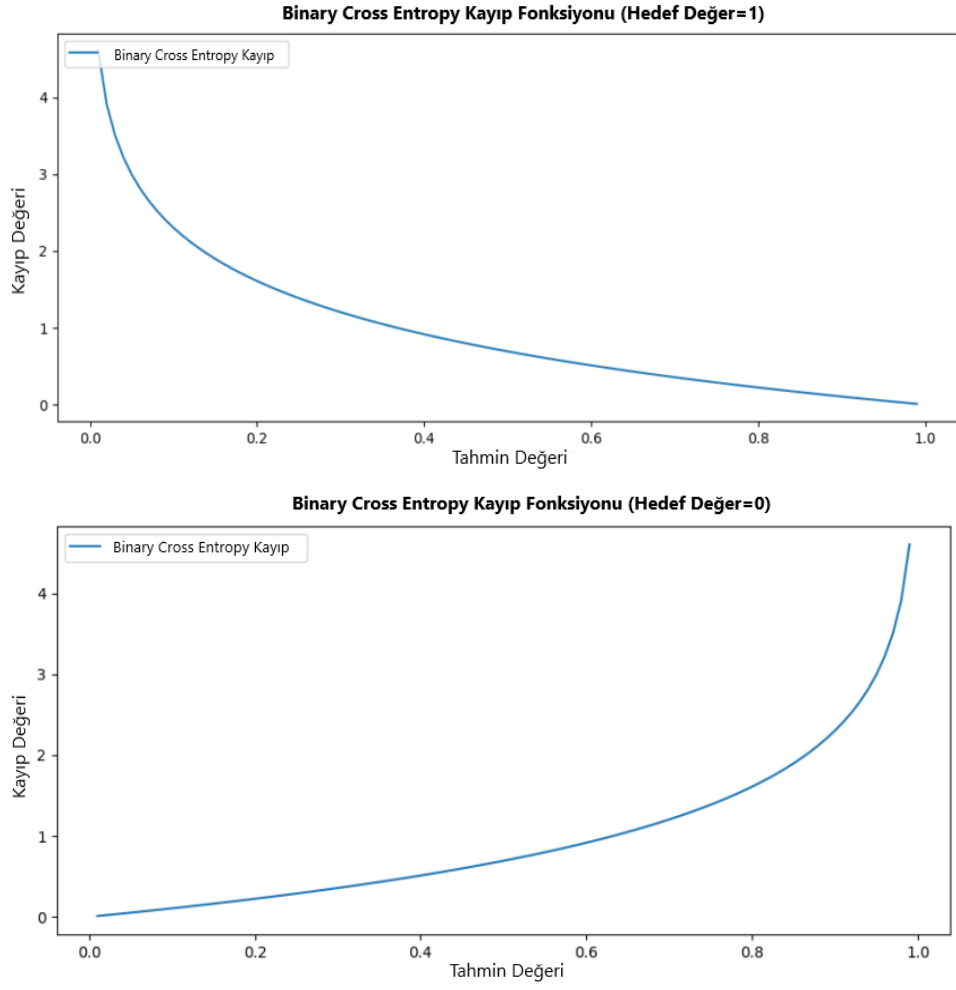
Şekil 3.11 Huber kayıp fonksiyonu (Anonymous 2018a)

Sınıflandırma algoritmalarında Binary Cross Entropy Loss, Hinge Loss, Categorical Cross Entropy Loss ve Kullback Leibler Divergence Loss kayıp fonksiyonları kullanılır.

Binary Cross Entropy Loss ikili sınıflandırma problemlerinde sık kullanılan kayıp fonksiyonlarından biridir.

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (3.4)$$

Formülde y hedef değerdir, 1 veya 0 olabilir. y=1 olduğu durum için modelin tahmin ettiği olasılık p(y), y=0 olduğu durum için ise 1-p(y) olur. Olasılık değerleri kullanılarak hedef değerlere ait kayıp değerleri bulunur. Hedef değer 1 ve 0 için Binary Cross Entropy kayıp fonksiyonu Şekil 3.12'deki gibi görselleştirilebilir.



Şekil 3.12 Binary Cross Entropy kayıp fonksiyonu (Anonymous 2019f)

Hinge Loss ikili sınıflandırma yöntemlerinde kullanılan bir diğer fonksiyondur, genellikle SVM algoritması ile kullanılır. Hedef değeri -1 veya 1 olarak etiketli sınıfları kullanır.

$$L = \max(0, 1 - y * f(x)) \quad (3.5)$$

$y * f(x) \geq 1$ olduğu durumda kayıp değeri 0 olur, $y * f(x) < 1$ olduğunda ise büyük ölçüde artar.

Categorical Cross-Entropy Loss çok sınıflı sınıflandırma işlemlerinde kullanılan bir kayıp fonksiyonudur. One-hot encoded ile temsil edilen hedef vektörleri kullanır. SoftMax aktivasyon fonksiyonu ile kullanılması önerilir. Bunun nedeni SoftMax fonksiyonunun her bir sınıfa ait benzerliği gösteren olasılık değeri oluşturmasıdır.

$$L(y, \hat{y}) = - \sum_{j=0}^M \sum_{i=0}^N (y_{ij} * \log(\hat{y}_{ij})) \quad (3.6)$$

Kullback Leibler Divergence Loss bir dağılımın gerçek dağılımdan ne kadar farklı olduğunu hesaplayan fonksiyondur. $p(x)$ verilerin gerçek dağılımını temsil eder. $q(x)$ ise $p(x)$ 'e yaklaşıklığı temsil eder.

$$D_{KL}(p||q) = \sum_{i=1}^N p(x_i) \cdot (\log p(x_i) - \log q(x_i)) \quad (3.7)$$

3.4.3 Optimizasyon fonksiyonları

Optimizasyon fonksiyonları kayıp değerini en aza indirmek için kullanılan yöntemlerdir. Bu yöntemlerden en sık kullanılanları Stochastic Gradient Descent (SGD), Momentum, Adagrad, Adadelat, RMSprop ve Adam yöntemleridir.

SGD tüm eğitim verisi yerine rastgele seçilen örnekler üzerinden işlem yapar. Bu şekilde daha kısa sürede işlem yapılır ve kayıp fonksiyonlarının değerinin tutulmasına gerek olmadığından daha az bellek gerektirir (Anonymous 2019e).

Momentum geçmiş gradyanların bilgisinin sonraki yinelemeyi etkilediği bir yöntemdir, yüksek eğrilik durumunda hedefe yakınsama hızı artar (Sun vd. 2019). Adagrad öğrenme katsayısını değiştirerek optimizasyon sağlayan bir yöntemdir. Adadelat ve RMSprop, Adagrad yönteminin geliştirilmiş birer versiyonudur. Adagrad yöntemindeki azalan öğrenme katsayısı problemine çözüm getirmeye odaklanmışlardır. Adam,

AdaGrad ve RMSprop'un avantajlı yönlerinin birleşimi olan bir yöntemdir. AdaGrad'ın seyrek gradyanlarla başa çıkabilme ve RMSProp'un ise sabit olmayan hedeflerle başa çıkabilme özelliğini almıştır (Kingma ve Ba 2014).

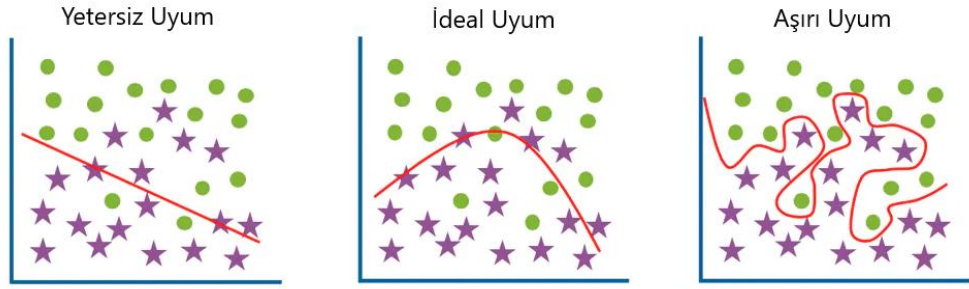
3.4.4 Aşırı uyum ve yetersiz uyum

Modelin eğitim verisi üzerinde gereğinden fazla çalışıp her detayı öğrenmesi durumu aşırı uyum (overfitting) olarak adlandırılmaktadır. Bu durumda model, eğitim verisini ezberlemeye başlar ve yeni bir veri geldiğinde genelleme yapamadığı için kötü tahminlerde bulunur (Anonymous 2019b). Aşırı uyumdan kaçınmanın yolları:

- Eğitim Verisinin Arttırılması: Eğitim setindeki çeşitliliğin artması modelin daha iyi genelleme yapmasına olanak tanır.
- Öznitelik Seçimi: Önemli özniteliklerin belirlenerek seçilmesi ve gereksiz olanların veri setinden çıkarılması özniteliklerin hedefe yakınlığının belirlenmesinde kolaylık sağlar.
- Düzenleştirme (Regularization): Düzenleştirme modelin daha iyi genelleştirme yapabilmesi için öğrenme algoritmasında değişiklikler yapan bir tekniktir (Anonymous 2018c). Düzenleştirme için L1 ve L2 düzenleştirme yöntemleri veya seyreltme (dropout) yöntemi kullanılabilir.
- Erken Durdurma (Early Stopping): Modelin performansında düşüş görüldüğünde eğitimi durdurmaya olanak sağlayan bir yöntemdir.

Yetersiz uyum (underfitting) modelin eğitim verisine uygun olmadığı ve yeni veri için de genelleme yapılamadığı durumdur (Anonymous 2019b). Yetersiz uyum durumunda eğitim ve test veri setinin her ikisi ile elde edilen doğruluk skoru düşük, hata skoru ise yüksek olacaktır. Yetersiz uyumun meydana gelmesinde veri çeşitliğinin azlığı, veri boyutunun küçük olması ve verinin gürültülü olması etken olabilir (Anonim 2020a).

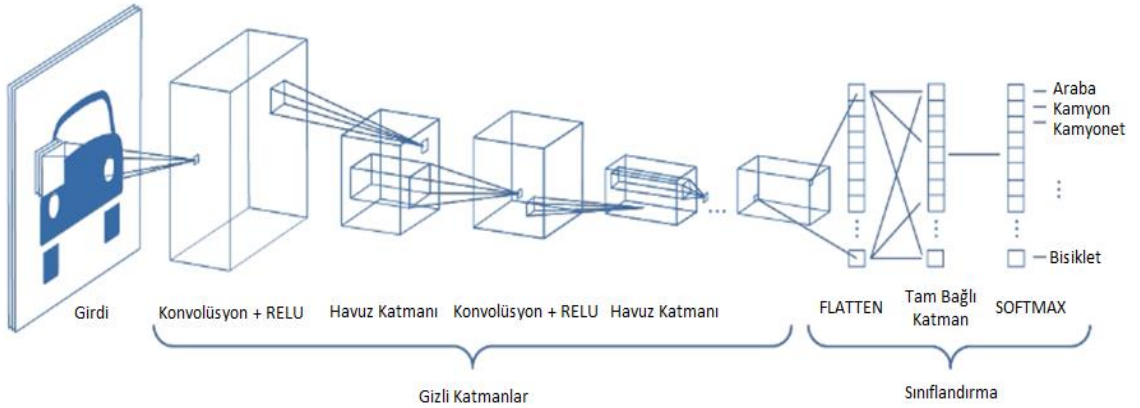
Sınıflandırma problemlerinde yetersiz uyum, ideal uyum ve aşırı uyum durumlarının karşılaştırılması Şekil 3.13'te gösterilmiştir.



Şekil 3.13 Yetersiz uyum, ideal uyum ve aşırı uyum (Anonymous 2021a)

3.4.5 Konvolüsyonel sinir ağları (CNN)

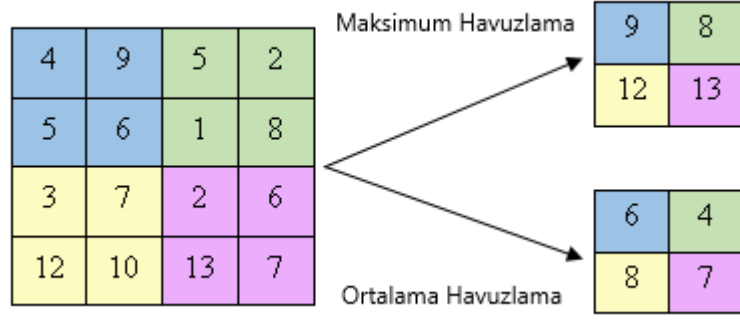
Konvolüsyonel sinir ağı (CNN) ileri beslemeli bir yapay sinir ağı modelidir. CNN görüntü işleme uygulamaları için idealdir, bunun dışında ses ve doğal dil işleme uygulamalarında da kullanılabilir. Bir CNN modeli temelde konvolüsyonel katman, havuz katmanı ve tam bağlı katmandan oluşur (Şekil 3.14).



Şekil 3.14 Konvolüsyonel sinir ağı modeli (Anonymous 2018d)

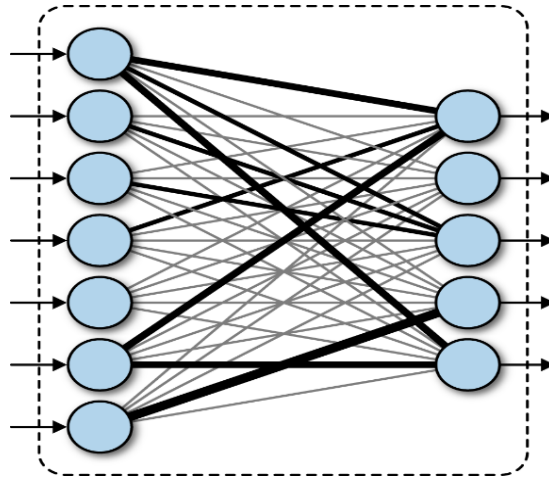
Konvolüsyonel katman ve havuz katmanı öznelik çıkarımının yapıldığı, tam bağlı katman ise sınıflandırma işleminin yapıldığı bölümün bir parçasıdır. Konvolüsyonel katmanda filtreler aracılığı ile öznelik haritası (feature map) çıkarımı yapılır. Havuz katmanı konvolüsyonel katmandan sonra yer alır ve girdi parametrelerinin boyutsal olarak azaltılması işlemini gerçekleştirir. Bu işlem için genellikle maksimum havuzlama (max pooling) ve ortalama havuzlama (average pooling) yöntemlerinden birisi tercih

edilir (Şekil 3.15). Maksimum havuzlama yönteminde filtre girdi boyunca hareket ederek çıktığı oluşturacak maksimum değerleri seçer. Ortalama havuzlama yönteminde, hareket eden filtre bulunduğu alandaki girdi değerlerinin ortalamasını alarak çıktı değerlerini oluşturur.



Şekil 3.15 Maksimum havuzlama ve ortalama havuzlama yöntemi

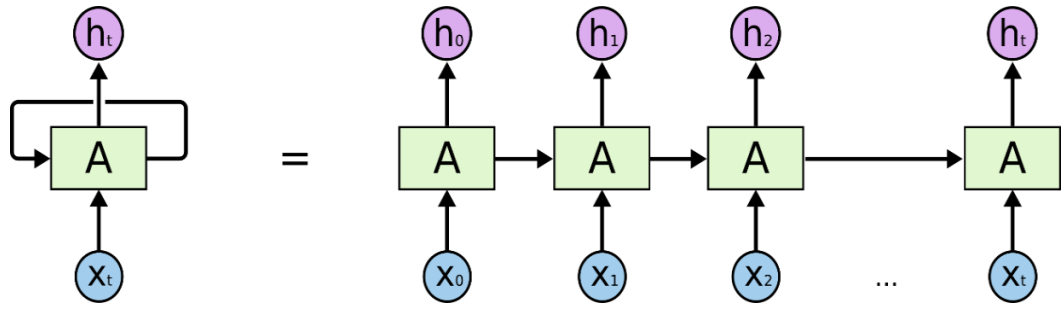
Flatten katmanı gelen girdiyi tek boyutlu diziye çevirerek tam bağlı katman için hazırlar. Tam bağlı katman özneliklerin sınıflandırılmasının yapıldığı katmandır. Tam bağlı katmanda, bir katmandaki her nöron sonraki katmanda yer alan nöronların tamamı ile bağlantılı durumdadır (Şekil 3.16).



Şekil 3.16 Bir tam bağlı katman modeli (Anonymous 2018f)

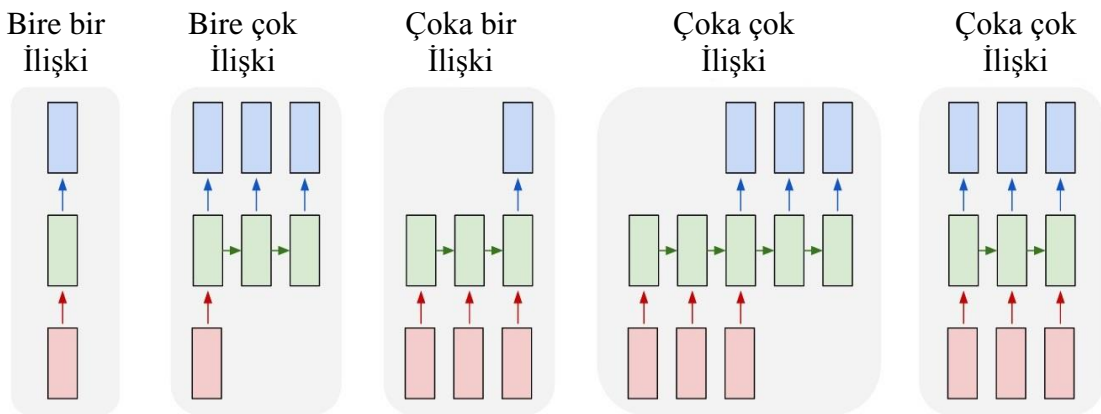
3.4.6 Tekrarlayan sinir ağıları (RNN)

Tekrarlayan Sinir Ağı (RNN) birimler arasındaki bağlantıların yönlendirilmiş bir döngü oluşturduğu sıralı veri setleri için ideal bir yapay sinir ağı modelidir (Şeker vd. 2017). RNN'lerde çıktı sadece mevcut girdiye değil önceden hesaplanan çıktıya da bağlıdır (Şekil 3.17). Örnek olarak bir sonraki kelimeyi öncekilere göre tahmin etmeye çalışan bir dil modeli düşünürsek RNN cümledeki her kelimeyi ve aralarındaki ilişkiyi depolayacak ve ihtiyaç duyduğunda geçmiş bilgileri kullanacaktır (Anonymous 2015a).



Şekil 3.17 Bir tekrarlayan sinir ağı yapısı (Anonymous 2015a)

RNN yapıları kullanım amaçlarına göre farklı ilişki modelleriyle oluşturulabilir (Şekil 3.18).



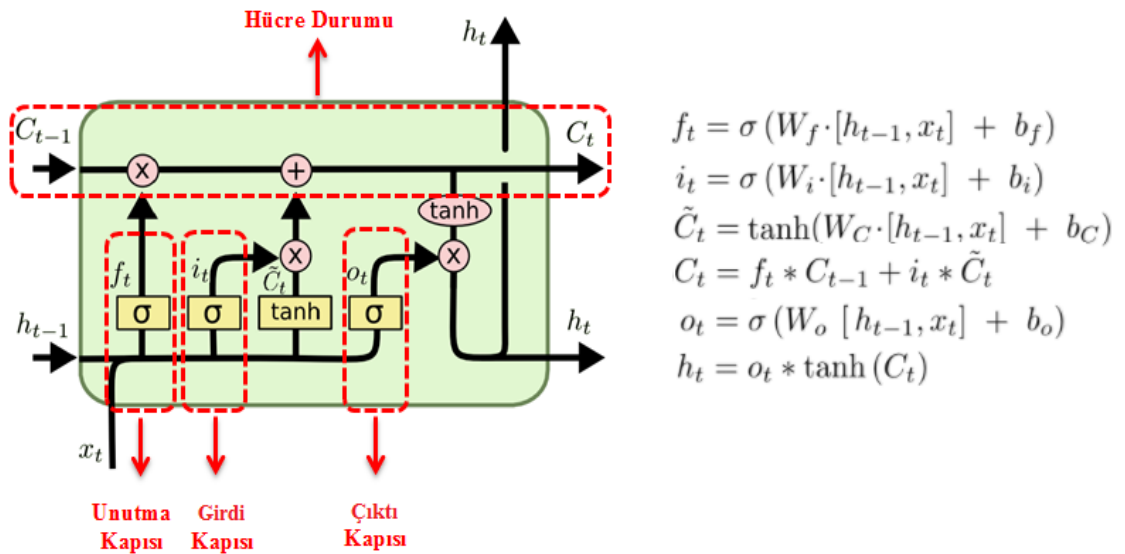
Şekil 3.18 RNN modelleri (Anonymous 2015b)

Şekil 3.18'de kırmızı ile işaretlenmiş birimler giriş ve mavi ile işaretlenmiş birimler çıkış vektörlerini temsil eder. Yeşil ile işaretlenmiş birimlerde ise RNN'nin durum

bilgisi tutulur. Bire bir ilişki modeli geleneksel sinir ağı modelidir. Bire çok ilişki modeli görüntü alt yazılama gibi bir görüntü girdisi sonucu metin üreten uygulamalarda kullanılabilir. Çoka bir ilişki duygu analizi için uygundur, bir sıralı kelime dizisi işlenerek duygu durumu çıktısı elde edilir. Çoka çok ilişki sıralı girdi alır ve sıralı çıktı üretir. Makine çevirisi uygulamaları için bu model için örnektir İngilizce bir metin okunur ve Türkçe karşılığı çıktı olarak alınır. Senkronize sıralı girdi ve çıktıların olduğu çoka çok ilişki modeli için ise örnek olarak varlık ismi tanıma (Named-Entity Recognition) uygulamaları gösterilebilir (Anonymous 2019c). Varlık ismi tanıma uygulamaları metinlerde geçen varlık isimlerini önceden belirlenmiş kategorilere göre sınıflandırmayı amaçlar.

LSTM duygu analizinde en sık kullanılan ve başarılı sonuç veren bir RNN yöntemidir. Hochreiter ve Schmidhuber tarafından 1997 yılında ortaya konulan yöntem, uzun vadeli bağımlılık problemlerinden kaçınmak için tasarlanmıştır (Anonymous 2015a).

Bir LSTM hücresi, hücre durumu ismi verilen bilgi akış hattı ve karar mekanizmasını oluşturan üç adet etkileşimli kapıdan oluşur (Şekil 3.19). Bu kapılar unutma kapısı (forget gate), girdi kapısı (input gate) ve çıktı kapısı (output gate) olarak adlandırılır.



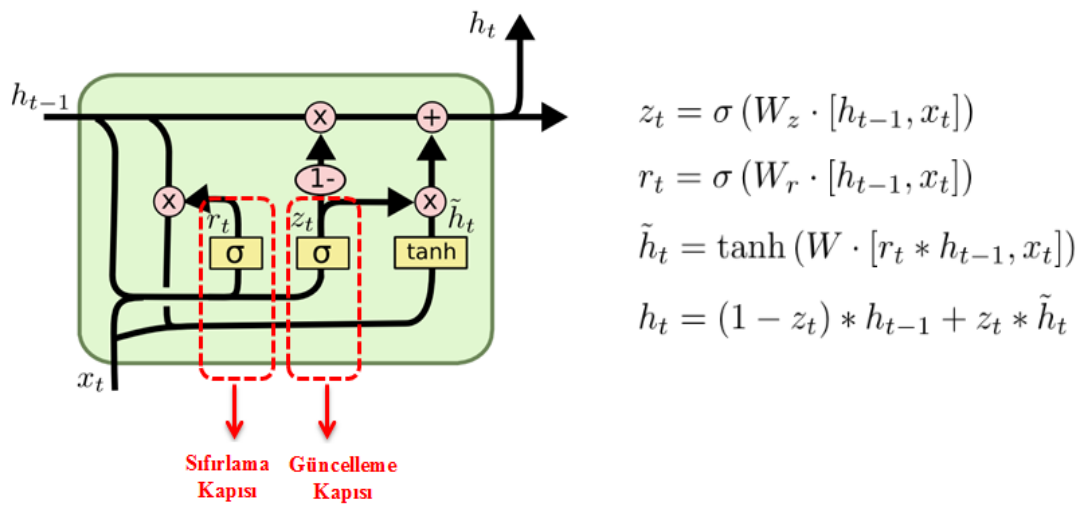
Şekil 3.19 LSTM hücresi (Anonymous 2015a)

LSTM işleyişinde ilk olarak hücre durumunda hangi bilginin tutulup tutulmayacağına karar verilir. Önceki gizli katmandan alınan bilgi ile şu anki girdi bilgisi unutmaya kapısına gönderilir ve sigmoid fonksiyonu yardımıyla sonuç oluşturulur. Sigmoid fonksiyonunun sonucu 0 ile 1 arasında değişmektedir. 0 bilgiyi unut, 1 ise bilgiyi tut anlamına gelmektedir.

İkinci adımda hangi yeni bilginin hücre durumunda depolanacağına karar verilir. Girdi kapısı sigmoid fonksiyonu ile hangi bilgilerin güncelleneceğine karar verir. Sonrasında Tahn fonksiyonu ile değer vektörleri oluşturulur ve sonuçların değerlendirilmesi ile hücre durumu güncellenir.

Son adımda ise hücre durumundaki hangi bilgilerin sonuç olarak gönderileceğinin kararı verilir. Hücre durumundan alınan bilgi Tahn katmanında düzenlenir, çıktı kapısı karar aşamasını gerçekleştirir ve sonuç oluşturulur.

LSTM yapay sinir ağlarının zamanla farklı versiyonları oluşturulmuştur ve Geçitli Tekrarlayan Birim (GRU) bunlardan biridir. Cho vd. (2014) tarafından geliştirilen GRU, LSTM modelinin basitleştirilmiş halidir (Anonymous 2017b). Bir GRU modeli Şekil 3.20’de gösterilmiştir.



Şekil 3.20 GRU (Anonymous 2015a)

GRU iki kapıdan oluşur. Güncelleme kapısı hangi bilginin tutulup hangi bilginin atılacağına karar verir. Sıfırlama kapısı ise geçmiş bilgilerinin ne kadarının tutulacağına karar verir. GRU birimleri güncelleme ve sıfırlama kapıları ile bilgileri eleyerek yalnızca ihtiyacı olan bilgiyi tutar ve böylece kaybolan gradyan problemine çözüm getirir (Anonymous 2017b). GRU, LSTM ye göre daha az hiper parametreye sahip olduğu için eğitimi daha hızlıdır ve eğitim için daha az bilgiye ihtiyaç duyar (Anonymous 2015c).

4. MATERYAL VE YÖNTEM

Bu bölümde yapılan duygu analizi çalışmasına ait işlem adımları sırası ile açıklanacaktır. Uygulama temelde veri düzenlemesi, derin öğrenme modellerinin oluşturulması ve değerlendirilmesi, sonuçların hazırlanması adımlarından oluşur.

Uygulama geliştirme gereksinimleri:

- İşletim Sistemi: Windows 7 64 bit
- İşlemci (CPU): Intel Core i7 2.50 GHz
- Yüklü Bellek (RAM): 8 GB
- Grafik İşlemci Birimi (GPU): NVIDIA GeForce GTX 850M 2GB
- Programlama Dili: Python 3.8
- Uygulama Geliştirme Ortamı: PyCharm 2020.3.5
- Kullanılan Kütüphaneler: Tensorflow, Keras, pandas, NumPy, NLTK, scikit-learn, Matplotlib, Gensim, Mlxtend, Zemberek

4.1 Veri Seti

Çalışma için Hepsiburada (<https://www.hepsiburada.com/>) e-ticaret sitesindeki ürün yorumları ile oluşturulan veri seti (Anonymous 2018e) kullanılmıştır. Veri setinde 5 ölçülü değerlendirme üzerinden 272216 adet yorum bulunmaktadır (Çizelge 4.1).

Çizelge 4.1 Veri setinde yer alan yorumların skor dağılımı

| Skor | Yorum Sayısı |
|----------|--------------|
| 5 yıldız | 167427 |
| 4 yıldız | 62394 |
| 3 yıldız | 28719 |
| 2 yıldız | 6450 |
| 1 yıldız | 7226 |

- Analizi yapılacak metinlerde yazım yanlışları bulunabilmektedir, bu durum uygulamanın yazım yanlışı olan kelimeyi farklı bir kelime olarak anlamasına neden olmaktadır. Yazım yanlışlarının giderilmesi işlemi normalizasyon olarak adlandırılır. Yapılan çalışmada ilk olarak tüm ürün yorumlarındaki yazım yanlışları Zemberek (Anonim 2019a) doğal dil işleme kütüphanesinin metin normalizasyonu metodundan faydalanılarak düzeltilmiştir. Örnek;

“ses çikisi çooook kötü bu telefonla müzik dinlemek istemezsiniz”

“ses çıkışı çok kötü bu telefonla müzik dinlemek istemezsiniz”

- Metin içerisindeki büyük harfler küçük harfe dönüştürülmüştür.
- Metin içerisinden e-posta adresleri, URL, ip adresleri ve numaralar çıkartılmış, tüm noktalama işaretleri temizlenmiştir.
- Metin içerisinden cümlenin duygusuna etkisi olmayan “ve”, “ama”, “belki” gibi Türkçe durak kelimeler (stop words) çıkartılmıştır. Durak kelimelerin çıkartılmasında Python NLTK (Loper ve Bird 2002) kütüphanesinden faydalanılmıştır.

4.3 Kelime Gömme (Word Embedding)

Yapılan çalışmada performansı arttırmak amacı ile FastText tarafından CBOW ve Skip-gram yaklaşımıyla oluşturulan önceden eğitilmiş kelime gömme modelleri kullanılmıştır. Metinde geçen her kelimenin vektörel karşılığı ile oluşturulan embedding matrisi, derin öğrenme modellerinde yer alan embedding katmanına girdi olarak gönderilerek işlem yapılmıştır. Katmanda boyut FastText’in vektör modelinin boyutuna uygun olarak belirlenmiştir.

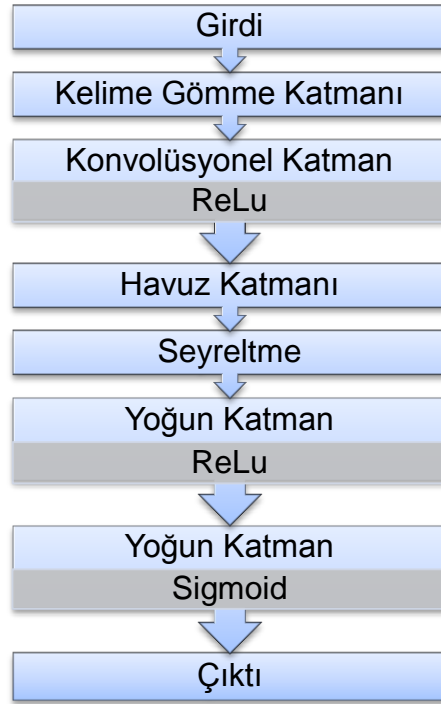
4.4 Uygulama Modelleri

Çalışmada yapılan deneylerde CNN, LSTM ve GRU yapay sinir ağı modelleri kullanılmıştır. Modellerin oluşturulmasında Tensorflow (Abadi vd. 2016) ve Keras

kütüphanelerinden faydalanılmıştır. TensorFlow makine öğrenmesi modellerini oluşturmak için kullanılan açık kaynaklı bir platformdur. Keras, Tensorflow üzerinde çalışan bir derin öğrenme kütüphanesidir.

4.4.1 CNN modeli

CNN modeli kelime gömme katmanı, konvolüsyonel katman, havuz katmanı ve tam bağlı katman (yoğun katman) ile oluşturulmuştur (Şekil 4.2). Kelime gömme katmanında FastText'in önceden eğitilmiş kelime vektörlerinden faydalanılmıştır. Modelde seyreltme (dropout) işlemi yapılmış ve eşik değeri 0.5 olarak belirlenmiştir. Aktivasyon fonksiyonu olarak ReLu ve Sigmoid fonksiyonları kullanılmıştır.



Şekil 4.2 CNN model mimarisi

4.4.2 LSTM ve GRU modeli

LSTM ve GRU yöntemleri için benzer modeller oluşturulmuştur. Her model kelime gömme katmanı, tekrarlayan katman ve tam bağlı katmandan oluşmaktadır (Şekil 4.3).

Modellerin kelime gömme katmanının oluşturulmasında önceden eğitilmiş FastText kelime vektörlerinden faydalanılmıştır. Tekrarlayan katman yorumlardaki duygu durumunun bulunması işlemi yapan temel katmandır. Bu katmanda seyreltme (dropout) işlemi yapılmış ve eşik değeri 0.2 olarak belirlenmiştir. Modellerde aktivasyon fonksiyonu olarak Sigmoid kullanılmıştır.



Şekil 4.3 LSTM ve GRU model mimarisi

4.5 Performans Metrikleri

Modellere ait tahminsel sonuçların değerlendirilmesi için metrik fonksiyonlarından faydalanılmıştır. Metrik fonksiyonlarının temelini karar matrisi (confusion matrix) oluşturmaktadır (Şekil 4.4).

| | | Gerçek Değerler | |
|------------------|-------------|-----------------|-------------|
| | | Pozitif (1) | Negatif (0) |
| Tahmin Değerleri | Pozitif (1) | DP | YP |
| | Negatif (0) | YN | DN |

Şekil 4.4 Karar matrisi

Karar matrisindeki parametreler aşağıda açıklanmıştır:

DP (Doğru Pozitifler - True Positives): Pozitif olarak işaretlenmiş pozitif örnekler

YP (Yanlış Pozitifler - False Positives): Pozitif olarak işaretlenmiş negatif örnekler

DN (Doğru Negatifler - True Negatives): Negatif olarak işaretlenmiş negatif örnekler

YN (Yanlış Negatifler - False Negatives): Negatif olarak işaretlenmiş pozitif örnekler

Kullanılan metrik fonksiyonlarının tanımı ve formülleri aşağıda listelenmiştir.

Doğruluk (Accuracy) sınıflandırıcının ulaştığı doğru sonuçların tüm sonuçlara oranıdır.

$$\text{Doğruluk} = \frac{DP+DN}{DP+YN+YP+DN} \quad (4.1)$$

Kesinlik (Precision) sınıflandırıcının doğru olarak tahmin ettiği sonuçlardan ne kadarının pozitif olduğunu belirler.

$$\text{Kesinlik} = \frac{DP}{DP+YP} \quad (4.2)$$

Duyarlılık (Recall) sınıflandırıcının gerçek pozitiflerden ne kadarını pozitif olarak etiketlediğini belirler.

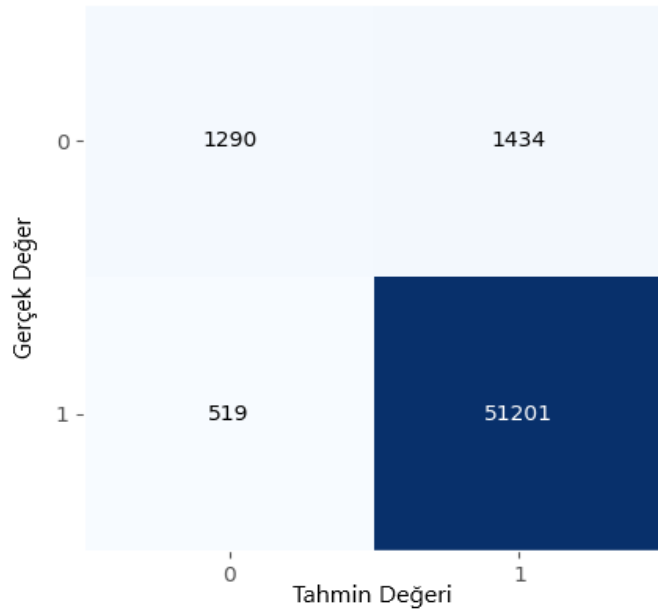
$$\text{Duyarlılık} = \frac{DP}{DP+YN} \quad (4.3)$$

F1-Skoru (F1-Score) kesinlik ve duyarlılığın ağırlıklı ortalamasıdır. F1-Skoru formülü Şekil 5.12' de gösterilmiştir.

$$\text{F1-Skoru} = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4.4)$$

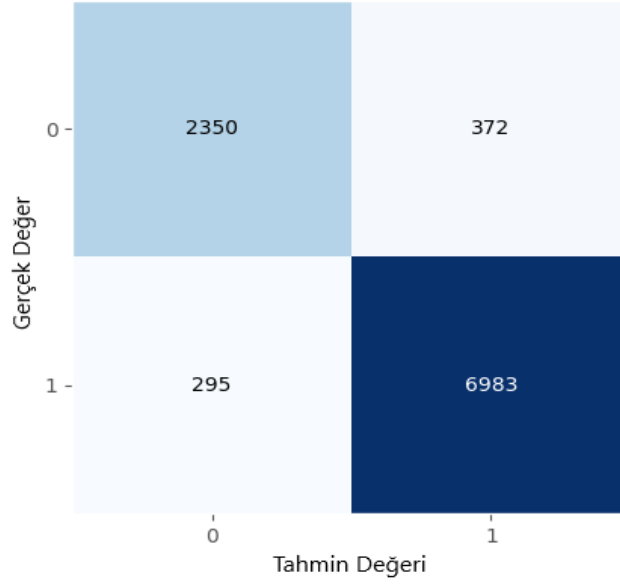
5. BULGULAR VE TARTIŞMA

İlk aşamada tüm kayıtları içeren veri setine ön işlemler uygulanmıştır. Bu şekilde duygu analizine etkisi olmayan kelimeler, rakamlar ve noktalama işaretleri temizlenmiştir. Sonrasında veri seti FastText (Skip-gram) önceden eğitilmiş kelime vektörleri kullanılarak LSTM modeli ile 10 tur eğitilmiştir. Elde edilen sonuçlar karar matrisi ile analiz edilmiştir (Şekil 5.1).



Şekil 5.1 Duygu sınıflandırması karar matrisi (dengesiz dağılımlı veri)

Şekil 5.1 modelde bir öğrenme probleminin olduğunu göstermektedir. Eğitim verisi çok fazla pozitif yorum içermekte ve negatif yorumların sayısı yetersiz kalmaktadır. Model pozitif yorumları başarılı bir şekilde yorumlarken negatif yorumların %50'sinden fazlası doğru şekilde yorumlanamamıştır. Bu nedenle pozitif ve negatif yorum sayısı daha dengeli olan bir veri seti hazırlanmıştır. 36500 pozitif ve 13500 negatif yorum içeren bu veri seti aynı LSTM modeli ile eğitilerek Şekil 5.2'de yer alan sonuç elde edilmiştir.



Şekil 5.2 Duygu sınıflandırması karar matrisi (dengeli dağılımlı veri)

Dengeli dağılıma sahip veri setini kullanmanın eğitim başarısını arttırdığı gözlemlenmiştir. Negatif yorumlar için yapılan yanlış tahmin oranı %15 civarına düşmüştür. Bu nedenle yapılacak deneylerde dengeli veri seti kullanılacaktır.

Eğitimde başarı oranını arttıran bir diğer faktör hiper-parametre seçimidir. Hiper-parametreler nöron sayısı, öğrenme kat sayısı, aktivasyon fonksiyonu ve optimizasyon fonksiyonu gibi modelin eğitim sürecinin kontrol edilmesini sağlayan ayarlanabilir parametrelerdir. CNN, LSTM ve GRU modelleri için belirlenen hiper-parametreler sırası ile açıklanacaktır.

Modeller için optimizasyon fonksiyonu olarak Adam kullanılmış ve öğrenme oranı 0.001 olarak belirlenmiştir. Kayıp fonksiyonu olarak ise Binary Cross-Entropy kullanılmıştır. Bu parametrelerde değişiklik yapılmamıştır.

Paket boyutu model ağırlıkları güncellenmeden üzerinde çalışılacak örnek sayısını ifade etmektedir. İdeal paket boyunun bulunması için CNN, LSTM ve GRU modeli FastText (Skip-gram) kelime vektörleri kullanılarak 32, 64 ve 128 paket boyutu ile eğitilmiş ve aşağıdaki sonuçlar elde edilmiştir (Çizelge 5.1).

Çizelge 5.1 Paket boyutunun CNN, LSTM ve GRU modellerinin performansına etkisi

| | Paket Boyutu | Doğruluk (%) | Kayıp (%) | F1-Skoru (%) | Kesinlik (%) | Duyarlılık (%) |
|------------------------------------|---------------------|---------------------|------------------|---------------------|---------------------|-----------------------|
| CNN + FastText (Skip-gram) | 32 | 91.55 | 22.22 | 94.20 | 92.00 | 96.73 |
| | 64 | 91.76 | 21.96 | 94.25 | 93.41 | 95.33 |
| | 128 | 91.61 | 21.92 | 94.17 | 92.77 | 95.82 |
| LSTM + FastText (Skip-gram) | 32 | 93.34 | 20.89 | 95.35 | 94.85 | 96.03 |
| | 64 | 93.23 | 18.93 | 95.35 | 93.59 | 97.36 |
| | 128 | 93.38 | 18.36 | 95.45 | 93.44 | 97.73 |
| GRU + FastText (Skip-gram) | 32 | 93.60 | 19.62 | 95.56 | 94.31 | 97.00 |
| | 64 | 93.53 | 19.01 | 95.54 | 94.01 | 97.31 |
| | 128 | 93.73 | 17.03 | 95.65 | 94.61 | 96.88 |

CNN modelinde 64, LSTM ve GRU modelinde 128 paket boyutu ile eğitilen modellerin başarı oranının yüksek, kayıp değerinin ise düşük olduğu gözlemlenmiştir. Bu nedenle sonraki deneylerde CNN modellerinin paket boyutu 64, LSTM ve GRU modellerinin paket boyutu 128 olarak kullanılacaktır.

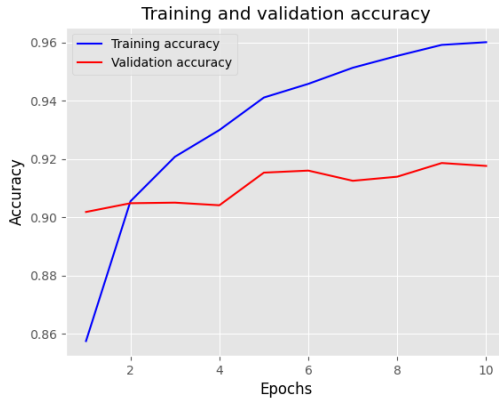
Modelin eğitim başarısını etkileyen bir diğer önemli parametre eğitim turu (epoch) sayısıdır. Eğitim turu öğrenme algoritmasının eğitim verisi üzerinde kaç kez çalışacağını belirler. FastText (Skip-gram) kelime vektörleri kullanan CNN, LSTM ve GRU modelleri için 5, 10, 15 eğitim turu ile elde edilen sonuçlar paylaşılmıştır (Çizelge 5.2).

Çizelge 5.2 Tur sayısının FastText (Skip-gram) kelime vektörleri kullanan modellerin performansına etkisi

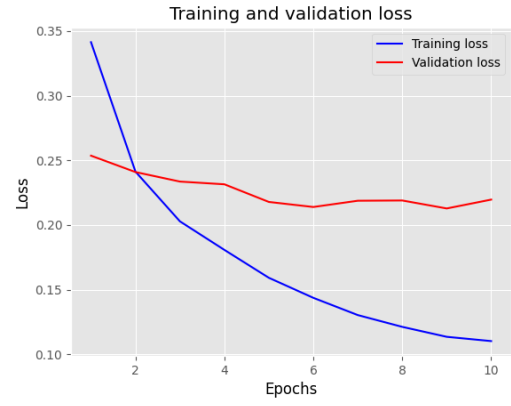
| | Tur Sayısı | Doğruluk (%) | Kayıp (%) | F1-Skoru (%) | Kesinlik (%) | Duyarlılık (%) |
|-----------------------------|------------|--------------|-----------|--------------|--------------|----------------|
| CNN + | 5 | 91.44 | 21.70 | 94.15 | 91.73 | 96.96 |
| FastText (Skip-gram) | 10 | 91.76 | 21.96 | 94.25 | 93.41 | 95.33 |
| | 15 | 91.67 | 22.47 | 94.31 | 91.92 | 97.06 |
| LSTM + | 5 | 92.44 | 18.99 | 94.73 | 93.75 | 95.94 |
| FastText (Skip-gram) | 10 | 93.38 | 18.36 | 95.45 | 93.44 | 97.73 |
| | 15 | 92.83 | 21.80 | 95.07 | 92.90 | 97.54 |
| GRU + | 5 | 92.90 | 18.24 | 95.07 | 93.60 | 96.78 |
| FastText (Skip-gram) | 10 | 93.73 | 17.03 | 95.65 | 94.61 | 96.88 |
| | 15 | 93.67 | 19.43 | 95.63 | 94.02 | 97.46 |

Çizelge 5.2 analiz edildiğinde başarı oranının belirli bir tur sayısına kadar arttığı sonrasında ise azalmaya başladığı ve modeller için optimum tur sayısının 10 olduğu gözlemlenmiştir.

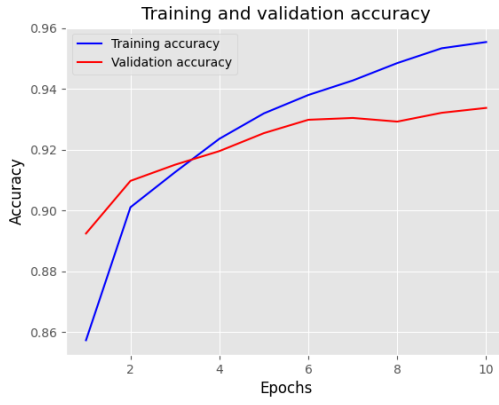
CNN, LSTM ve GRU modellerinin her biri için 10 tur sayısı ile elde edilen doğruluk ve kayıp grafiği Şekil 5.3’de gösterilmiştir.



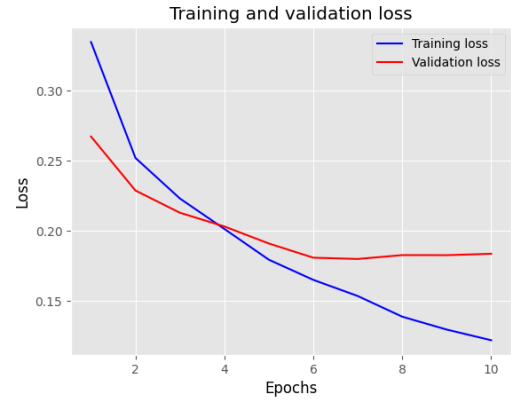
(A)



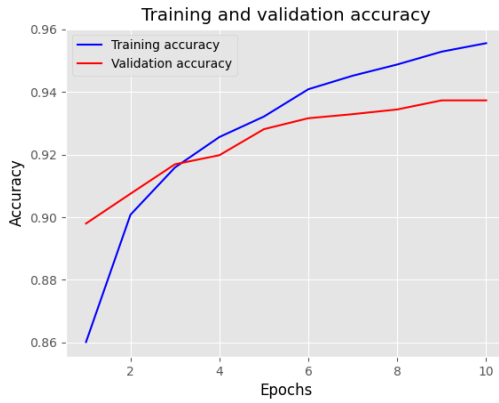
(B)



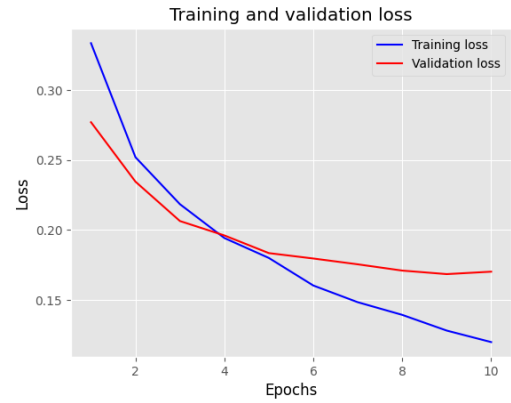
(C)



(D)



(E)



(F)

Şekil 5.3 (A) CNN+FastText (Skip-gram) doğruluk grafiği, (B) CNN+FastText (Skip-gram) kayıp grafiği, (C) LSTM+FastText (Skip-gram) doğruluk grafiği, (D) LSTM+FastText (Skip-gram) kayıp grafiği, (E) GRU+FastText (Skip-gram) doğruluk grafiği, (F) GRU+FastText (Skip-gram) kayıp grafiği

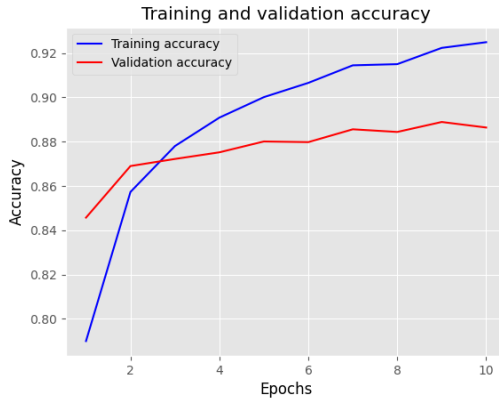
FastText (CBOW) kelime vektörleri kullanan modeller için de 5, 10, 15 tur sayıları ile eğitim yapılmış ve sonuçlar paylaşılmıştır (Çizelge 5.3).

Çizelge 5.3 Tur sayısının FastText (CBOW) kelime vektörleri kullanan modellerin performansına etkisi

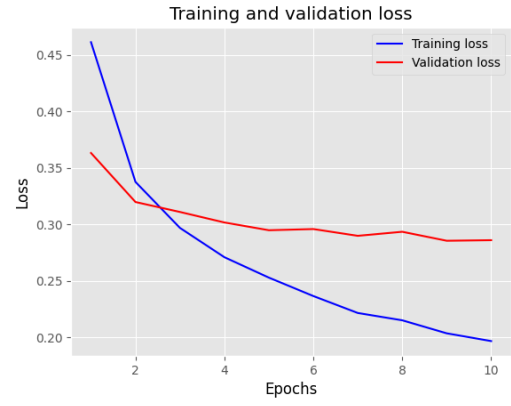
| | Tur Sayısı | Doğruluk (%) | Kayıp (%) | F1-Skoru (%) | Kesinlik (%) | Duyarlılık (%) |
|-------------------------------|-------------------|---------------------|------------------|---------------------|---------------------|-----------------------|
| CNN + FastText (CBOW) | 5 | 87.81 | 29.82 | 91.90 | 87.65 | 96.91 |
| | 10 | 88.64 | 28.60 | 92.39 | 88.56 | 96.87 |
| | 15 | 88.43 | 29.19 | 92.22 | 88.83 | 96.21 |
| LSTM + FastText (CBOW) | 5 | 89.85 | 25.71 | 92.99 | 91.66 | 94.66 |
| | 10 | 90.84 | 22.76 | 93.62 | 92.84 | 94.67 |
| | 15 | 91.56 | 22.23 | 94.13 | 93.81 | 94.68 |
| GRU + FastText (CBOW) | 5 | 89.98 | 25.64 | 92.97 | 92.47 | 93.75 |
| | 10 | 91.12 | 22.88 | 93.80 | 92.89 | 94.98 |
| | 15 | 91.80 | 21.74 | 94.26 | 94.04 | 94.71 |

FastText'in CBOW ile önceden eğitilmiş kelime vektörlerini kullanan CNN modeli, Skip-gram ile önceden eğitilen kelime vektörlerini kullanan modeller ile benzer bir davranış sergilemiştir. Doğruluk değeri 10. tura kadar artmış sonrasında düşüş göstermiştir. LSTM ve GRU modellerinde eğitim sayısı arttıkça doğruluk değerinin arttığı ve kayıp değerinin azaldığı gözlemlenmiştir.

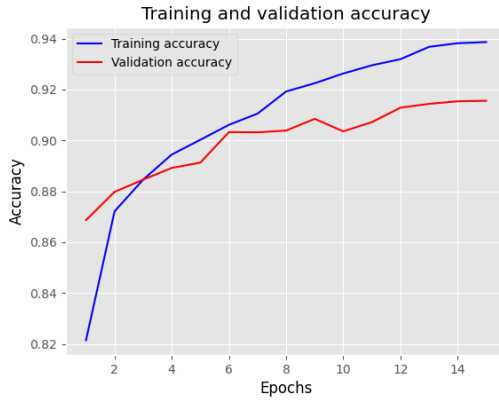
CNN modeli için 10, LSTM ve GRU modelleri için 15 tur sayısı ile elde edilen doğruluk ve kayıp grafiği Şekil 5.4'de gösterilmiştir.



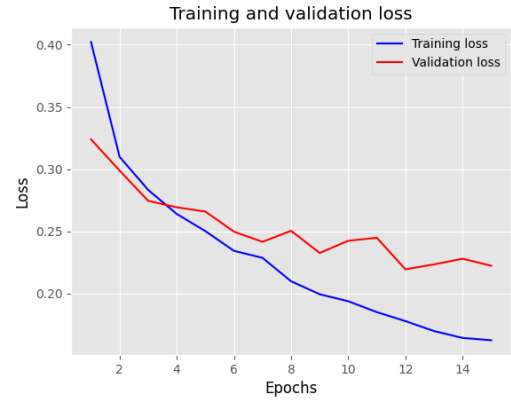
(A)



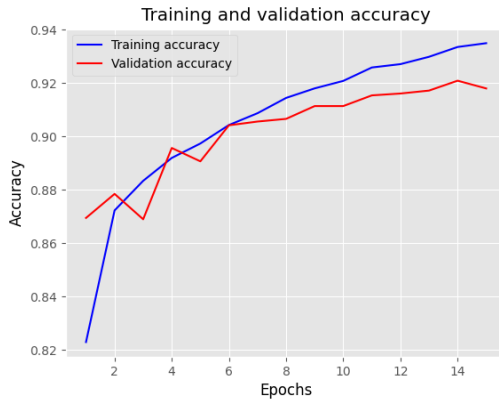
(B)



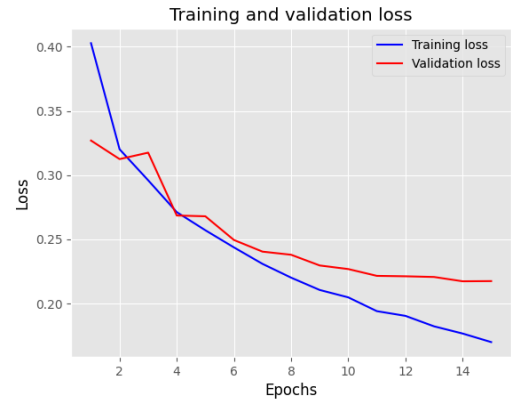
(C)



(D)



(E)



(F)

Şekil 5.4 (A) CNN+FastText (CBOW) doğruluk grafiği, (B) CNN+FastText (CBOW) kayıp grafiği, (C) LSTM+FastText (CBOW) doğruluk grafiği, (D) LSTM+FastText (CBOW) kayıp grafiği, (E) GRU+FastText (CBOW) doğruluk grafiği, (F) GRU+FastText (CBOW) kayıp grafiği

Çizelge 5.2 ve Çizelge 5.3'ün karşılaştırması yapıldığında, duygu analizinde önceden eğitilmiş FastText'in Skip-gram ile eğitilen kelime vektörlerini kullanan modellerin CBOW ile eğitilen kelime vektörlerini kullanan modellerden daha iyi sonuç verdiği gözlemlenmiştir.

Modellere ait performans ölçütlerinden biri de eğitim süresidir. Tüm modellerin eğitim turu 10 için çalışma süreleri Çizelge 5.4'de gösterilmiştir.

Çizelge 5.4 CNN, LSTM ve GRU eğitim süreleri

| | Eğitim Süresi (sn.) |
|------------------------------------|----------------------------|
| CNN + FastText (Skip-gram) | 255.33 |
| CNN + FastText (CBOW) | 254.22 |
| LSTM + FastText (Skip-gram) | 816.07 |
| LSTM + FastText (CBOW) | 836.98 |
| GRU + FastText (Skip-gram) | 698.43 |
| GRU + FastText (CBOW) | 675.9 |

CNN modeline ait çalışma süresi diğer modellerden az olmasına rağmen başarı oranı en düşük olan model olmuştur. GRU modeli LSTM modelinin basitleştirilmiş bir versiyonudur ve LSTM ile karşılaştırıldığında daha kısa sürede daha etkili sonuçlar vermiştir.

Duygu analizinde FastText (Skip-gram) önceden eğitilmiş kelime vektörlerini kullanan GRU modeli uygun parametreler kullanıldığında CNN ve LSTM modellerine göre daha başarılı performans göstermiş ve 93.73% doğruluk değeri elde edilmiştir.

Duygu analizinde GRU modeli kullanan diğer çalışmalar incelenmiştir. Santur (2019) Türkçe ürün yorumları verisini kullanarak GRU modeli ile 95% doğruluk elde etmiştir. Onan (2020) İngilizce Twitter ürün yorumları veri seti ile CNN, RNN, LSTM, GRU ve CNN-LSTM modellerinin bir karşılaştırmasını sunmuştur. CNN-LSTM modelinin 93.85% doğruluk değeri ile diğer modellerden daha başarılı olduğunu göstermiştir.

6. SONUÇ

Son yıllarda e-ticaret platformlarının hızla gelişmesiyle birlikte satılan ürünler hakkında yapılan inceleme ve deneyimlerin paylaşımında artış görülmektedir. Hizmet kalitesinin arttırılması açısından geri bildirimlerin analizi önemli bir rol oynamaktadır. Bu çalışmada bir e-ticaret platformuna ait ürün yorumları verisi üzerinde duygu analizi yöntemleri uygulanmış ve ürün yorumlarının duygu durumunun olumlu veya olumsuz olarak tespit edilmesi hedeflenmiştir. Bu hedefi gerçekleştirirken CNN, LSTM ve GRU derin öğrenme modelleri kullanılmıştır. Başarı oranını arttırmak amacıyla FastText tarafından Skip-gram ve CBOW yöntemiyle oluşturulan önceden eğitilmiş kelime vektörlerinden faydalanılmıştır. Modellerin eğitimi sonucu elde edilen sonuçlar karşılaştırılmış ve sınıflandırma başarısı en yüksek olan yöntemin GRU olduğu gözlemlenmiştir.

İlerleyen çalışmalarda Google AI Language ekibi tarafından 2018 yılında yayınlanan Bidirectional Encoder Representations from Transformers (BERT) (Devlin vd. 2018) kelime temsil modeli kullanılarak duygu analizi yapılması planlanmaktadır. Diğer kelime temsil modellerinden farklı olarak BERT her kelimenin sağ ve sol tarafındaki bağlamı dikkate alacak şekilde tasarlanmıştır. Soru cevaplama ve varlık ismi tanıma gibi doğal dil işleme problemlerinde yüksek başarı sağlamıştır.

KAYNAKLAR

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin M., Ghemawat S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y. and Zheng X. 2016. TensorFlow: A System for Large-Scale Machine Learning. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), 265-283.
- Akba, F., Uçan, A., Sezer, E. A., Sever, H. 2014. Assessment of Feature Selection Metrics for Sentiment Analyses: Turkish Movie Reviews. 8th European Conference on Data Mining, 191(2002); 180-184.
- Anonim. 2019a. Web Sitesi: <https://github.com/ahmetaa/zemberek-nlp>. Erişim Tarihi: 2019.
- Anonim. 2019b. Web Sitesi: http://www.tuik.gov.tr/PreTablo.do?alt_id=1028. Erişim Tarihi: 2019.
- Anonim. 2020a. Web Sitesi: <https://medium.com/kodluyoruz/makine-%C3%B6%C4%9Frenmesinde-overfitting-underfitting-best-fitting-kavramlar%C4%B1-9f80a5d42719>. Erişim Tarihi: 2021.
- Anonymous. 2015a. Web Sitesi: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Erişim Tarihi: 2019
- Anonymous. 2015b. Web Sitesi: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>. Erişim Tarihi: 2019
- Anonymous. 2015c. Web Sitesi: <http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-grulstm-rnn-with-python-and-theano/>. Erişim Tarihi: 2019.
- Anonymous. 2017a. Web Sitesi: https://isaacchanghau.github.io/post/activation_functions/. Erişim Tarihi: 2019.
- Anonymous. 2017b. Web Sitesi: <https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be>. Erişim Tarihi: 2019
- Anonymous. 2017c. Web Sitesi: https://www.researchgate.net/figure/a-The-building-block-of-deep-neural-networks-artificial-neuron-or-node-Each-input-x_fig1_312205163. Erişim Tarihi: 2019
- Anonymous. 2018a. Web Sitesi: <https://heartbeat.fritz.ai/5-regression-loss-functions-all-machine-learners-should-know-4fb140e9d4b0>. Erişim Tarihi: 2021
- Anonymous. 2018b. Web Sitesi: <https://medium.com/@shrutijadon10104776/survey-on-activation-functions-for-deep-learning-9689331ba092>. Erişim Tarihi: 2019

- Anonymous. 2018c. Web Sitesi: <https://www.analyticsvidhya.com/blog/2018/04/fundamentals-deep-learning-regularization-techniques/>. Erişim Tarihi: 2021.
- Anonymous. 2018d. Web Sitesi: <https://www.freecodecamp.org/news/an-intuitive-guide-to-convolutional-neural-networks-260c2de0a050/>. Erişim Tarihi: 2019.
- Anonymous. 2018e. Web Sitesi: <https://www.kaggle.com/cebeci/turkishreviews>. Erişim Tarihi: 2019.
- Anonymous. 2018f. Web Sitesi: <https://www.oreilly.com/library/view/tensorflow-for-deep/9781491980446/ch04.html>. Erişim Tarihi: 2019
- Anonymous. 2019a. Web Sitesi: <https://developer.nvidia.com/deep-learning>. Erişim Tarihi: 2019.
- Anonymous. 2019b. Web Sitesi: <https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/>. Erişim Tarihi: 2021.
- Anonymous. 2019c. Web Sitesi: <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks>. Erişim tarihi: 2019.
- Anonymous. 2019d. Web Sitesi: <https://towardsdatascience.com/nlp-101-word2vec-skip-gram-and-cbow-93512ee24314>. Erişim Tarihi: 2021.
- Anonymous. 2019e. Web Sitesi: <https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6>. Erişim Tarihi: 2021.
- Anonymous. 2019f. Web Sitesi: <https://www.machinecurve.com/index.php/2019/10/22/how-to-use-binary-categorical-crossentropy-with-keras/>. Erişim Tarihi: 2021.
- Anonymous. 2019g. Web Sitesi: <https://www.mathworks.com/discovery/deep-learning.html>. Erişim Tarihi: 2019.
- Anonymous. 2021a. Web Sitesi: <https://www.ibm.com/cloud/learn/overfitting#toc-overfittin-7bQGsqfX>. Erişim Tarihi: 2021.
- Asur, S. and Huberman, B. A. 2010. Predicting the Future With Social Media. Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, IEEE, 492-499.
- Ayata, D., Saraçlar, M. ve Özgür, A. 2017. Makine Öğrenmesi ve Kelime Vektör Temsili ile Türkçe Tweet Sentiment Analizi. 2017 International Conference on Computer Science and Engineering (UBMK), IEEE, 911-916.
- Cho, K., Merriënboer, V. B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv preprint arXiv:1406.1078.

- Çoban, Ö., Özyer, B. and Özyer, G. T. 2015. Sentiment analysis for Turkish Twitter feeds. 23rd Signal Processing and Communications Applications Conference (SIU), IEEE.
- Dang, N. C., Moreno-García, M. N. and De la Prieta, F. 2020. Sentiment Analysis Based on Deep Learning: A Comparative Study. *Electronics*, 9(3); 483.
- Dehkharghani, R., Yanikoglu, B., Saygin, Y. and Oflazer, K. 2015. Sentiment Analysis in Turkish: Towards a Complete Framework. *Natural Language Engineering*, 1.
- Devlin, J., Chang, M. W., Lee, K. and Toutanova, K. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
- Dos Santos, C. and Gatti, M. 2014. Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts. *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 69-78.
- Go, A., Bhayani, R. and Huang, L. 2009. Twitter Sentiment Classification using Distant Supervision. *CS224N Project Report*, 1(12); 2009.
- Hochreiter, S. and Schmidhuber, J. 1997. Long Short-Term Memory. *Neural Computation*, 9(8); 1735-1780.
- Joulin, A., Grave, E., Bojanowski, P. and Mikolov, T. 2016. Bag of Tricks for Efficient Text Classification. arXiv preprint arXiv:1607.01759.
- Karakuş, B. A., Talo, M., Hallaç, I. and Aydin, G. 2018. Evaluating Deep Learning Models For Sentiment Classification. *Concurrency and Computation: Practice and Experience*, 30(21); e4783.
- Kingma, D. P. and Ba, J. 2014. Adam: A Method For Stochastic Optimization.» arXiv preprint arXiv:1412.6980.
- Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.
- Loper, E. and Bird, S. 2002. NLTK: The Natural Language Toolkit. arXiv preprint cs/0205028.
- Medhat, W., Hassan, A., and Korashy, H. 2014. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113.
- Meehan, K., Lunney, T., Curran, K. and McCaughey, A. 2013. Context-Aware Intelligent Recommendation System for Tourism. 2013 IEEE international conference on pervasive computing and communications workshops (PERCOM workshops), IEEE, 328-331.
- Mikolov, T., Chen, K., Corrado, G. and Dean, J. 2013. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781.

- Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y. and Ngo, D. C. 2015. Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment. *Expert Systems with Applications*, 42(1); 306-324.
- Oğul, B. B. ve Ercan, G. 2016. Türkçe Otel Yorumlarından Duygu Analizi Sentiment Classification on Turkish Hotel Reviews. 24th Signal Processing and Communication Application Conference (SIU), IEEE, 497-500.
- Onan, A. 2020. Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. *Concurrency and Computation: Practice and Experience*, e5909.
- Pang, B., Lee, L. and Vaithyanathan, S. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. *Empirical Methods in Natural Language Processing*.
- Pang, B. and Lee, L. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 271.
- Pang, B. and Lee, L. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2); 1-135.
- Pennington, J., Socher, R. and Manning, C. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532-1543.
- Ravi, K. and Ravi, V. 2015. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 89; 14-46.
- Salur, M. U. and Aydin, İ. 2018. Sentiment Classification Based On Deep Learning. 2018 26th Signal Processing and Communications Applications Conference (SIU), IEEE, 1-4.
- Santur, Y. 2019. Sentiment Analysis Based on Gated Recurrent Unit. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), IEEE, 1-5.
- Sohangir, S., Wang, D., Pomeranets, A. and Khoshgoftaar, T. M. 2018. Big Data: Deep Learning for financial sentiment analysis. *Journal of Big Data*, 5(1); 3.
- Somasundaran, S., Wilson, T., Wiebe, J. and Stoyanov, V. 2007. QA with Attitude: Exploiting Opinion Type Analysis for Improving Question Answering in On-line Discussions and the News. *ICWSM*.
- Sun, S., Cao, Z., Zhu, H. and Zhao, J. 2019. A Survey of Optimization Methods From a Machine Learning Perspective. *IEEE Transactions on Cybernetics*, 50(8); 3668-3681.

- Şeker, A., Diri, B. ve Balık, H. H. 2017. Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme. Gazi Mühendislik Bilimleri Dergisi (GMBD), 3(3); 47-64.
- Tang, D., Qin, B., Feng, X. and Liu, T. 2015. Effective LSTMs for Target-Dependent Sentiment Classification. arXiv preprint arXiv:1512.01100.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. 2011. Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, 37(2); 267-307.
- Turney, P. D. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, 417-424.
- Türkmenoğlu, C. T. and Tantığ, A. C. 2014. Sentiment Analysis in Turkish Media. International Conference on Machine Learning (ICML).
- Wang, H., Can, D., Kazemzadeh, A., Bar, F. and Narayanan, 2012. A System for Real-time Twitter Sentiment Analysis of 2012 U.S.Presidential Election Cycle. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, 115-120.
- Yadav, A., Vishwakarma, D. K. 2020. Sentiment analysis using deep learning architectures: a review. Artificial Intelligence Review, 53(6); 4335-4385.
- Yang, L., Li, Y., Wang, J. and Sherratt, R. S. 2020. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access, 8; 23522-23530.
- Zhang, L., Wang, S. and Liu, B. 2018. Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4); e1253.
- Zhuang, L., Jing, F. and Zhu, X.-Y. 2006. Movie Review Mining and Summarization. Proceedings of the 15th ACM International Conference on Information and Knowledge Management, 43-50.