

Türkçe Metinlerde Duygu Analizi

Sentiment Analysis in Turkish Texts

Sinem TOKCAER, İzmir Ekonomi Üniversitesi, Türkiye, sinem.tokcaer@ieu.edu.tr

Orcid No: 0000-0001-8842-3574

Öz: Günümüzde yaygınlaşan internet ve sosyal medya kullanımının artmasıyla ortaya çıkan büyük verinin, analiz edilerek anlamlı bilgiye dönüştürülmesi oldukça büyük bir öneme sahiptir. Duygu analizinde ise, bir metin içerisinde fikir içeren verinin sistematik olarak incelenmesi ve metne dair duygu kategorisinin ve duygu polaritesinin belirlenmesi sürecidir. Sadece dil biliminde değil, finansal piyasalar, pazarlama ve sosyal medya analizi gibi birçok farklı alanda sıklıkla duygu analizi yaklaşımları kullanılmaktadır. İngilizcenin dünyada konuşulan ortak dil olması sebebiyle, literatürde İngilizce metinler üzerine yapılmış birçok duygu analizi çalışması bulunmaktadır. Ancak, Türkçe metinlerde duygu analizi hâlen geliştirilmeye açık bir araştırma alanıdır. Bu çalışmada, Türkçe metinlerde duygu analizi literatürü incelenerek, literatürde sıklıkla kullanılan yöntemler, açık kaynaklı kütüphaneler ve veri tabanları ortaya konulmuş ve araştırmaya açık konular irdelenmiştir

Anahtar Kelimeler: Duygu Analizi, Doğal Dil İşleme, Duygu Sınıflandırması, Duygu Polaritesi

JEL Sınıflandırması: C55, C80, Y20

Abstract: Recently, the need for analyzing and understanding the story behind big data coming from increasing use of internet and social media becomes a trend and has gained a huge impact on businesses. Sentiment analysis is a method to analyze the opinion in a text in a systematic way that it defines the sentiment category and polarity in that text. The respected areas where sentiment analysis is applied is not only limited to linguistics, but also several applications in financial markets, marketing and social media analysis are observed in the literature. Although sentiment analysis can be applied in any language, English has a dominance in the literature because it is a globally spoken language. Therefore, sentiment analysis in Turkish texts still requires further attention as an open research area. In this research, we examined the relevant literature on sentiment analysis in Turkish texts in terms of frequently applied methodologies, open source libraries and databases. Consequently, we define research gaps and further research topics in application of sentiment analysis in Turkish texts.

Keywords: Sentiment Analysis, Natural Language Processing, Sentiment Categorization, Sentiment Polarity

JEL Classification: C55, C80, Y20

1. Giriş

Günümüzde internet ağları ve sosyal medya kullanımı, insanların başka kişi ya da kurumlar hakkındaki görüşlerini kolay ve açık şekilde ifade edebilmesine ve yorumlamasına imkân sağlamıştır. Bu imkânın giderek yaygınlaşması ile büyüyen zengin bir içerik ortaya çıkmaktadır. Bunun sonucunda internetten elde edilen büyük verinin analiz edilerek anlamlı bilgiye dönüştürülmesi ve kullanılması son yıllarda yoğun olarak çalışılan bir konudur. Bu bağlamda, son on yılda, sosyal medya içeriğini analiz etmek için yapılan çalışmalar hızlı bir artış göstermiş ve netnografi çalışmalarından (Kozinets, 2009; Hine, 2005) fikir madenciliğine (Dave vd., 2003) ve doğal dil işlemeye (Nasukawa ve Yi, 2003) kadar çok sayıda yöntem önerilmiştir. Duygu analizi ise büyük veri içindeki öznel bilginin sistematik olarak analiz edilmesini sağlayan bir doğal dil

Makale Geçmişi / Article History

Başvuru Tarihi / Date of Application : 27 Nisan / April 2021

Kabul Tarihi / Acceptance Date : 30 Mayıs / May 2021

işleme yöntemidir. Metin madenciliği ve hesaplamalı dilbilim araçlarını kullanarak, araştırmacıların metin içindeki duyguları belirlemelerine, duygu kategorilerine göre sınıflandırmalarına ve kategorilerin duygusal polaritelerini bulmalarına yardımcı olur (Arazy ve Woo, 2007; Jacobson, 2009).

Duygu analiz modelleri, satış tahminleri (Fan vd.,2017), ürün derecelendirme (Çalı ve Balaman, 2019), kamuoyu yoklamaları (O'Connor vd., 2010), oylama sonucu tahminleri (Wang vd., 2012), menkul değerler piyasasında yatırımcı davranışları inceleme (Baker ve Wurgler, 2007) gibi birçok alanda kullanılmıştır. Pazarlama alanında ise müşterinin sesi (voice of customer) çalışmalarında (Aguwa vd., 2017) ve sosyal medya analizlerinde (Ceron vd., 2014) sıklıkla kullanılır. O'Leary (2011) ve Thelwall vd (2011) tarafından belirtildiği gibi, birinci elden tüketici perspektifinden ürünler / hizmetler hakkında fikir içeren bloglar, tweetler, çevrimiçi anketler gibi sosyal medya araçları, müşteri verilerini kaydetmek, işlemek ve altında yatan anlamı analiz etmek için büyük öneme sahiptir. Bunun yanı sıra, müşteri deneyimlerinin hangi yönlerinin belirli yanıtları ve duyguları ifade ettiğini ortaya koymak duygu analizi ile mümkün olabilmektedir (Wright, 2009). Wright (2009) ve Deshpande ve Sarkar (2010) 'a göre, duygu analizinin temel amacı, sosyal medya aracılığıyla paylaşılan verileri işleyerek elde edilen bilgiyi performans değerlendirme ve ürün / hizmet tasarımları gibi stratejik kararlar için bir rehber olarak kullanmaktır. Bunun yanı sıra, duygu analizinin bir diğer özelliği müşteri deneyimleri hakkındaki örüntüleri, zaman içerisindeki değişimi ve olaylar arasındaki ilişkiyi ortaya koyabilme gücüdür (Ku vd., 2019). Bu kapsamdan yola çıkarak Misopoulos vd. (2014) tarafından yapılan çalışmada, müşteri memnuniyetini arttıran ve iyileştirme gerektiren müşteri hizmetleri unsurları belirlenmiştir. Benzer şekilde Park ve Ha (2018) ise yaptıkları çalışma ile son kullanıcı yorumlarından yola çıkarak, müşteri şikayetlerini kategorize etmiş ve şikayetlerin bu kategoriler altında zaman içerisindeki değişimini ortaya koymuştur. Sıklıkla kullanılan yöntemler ile ilgili literatür, Medhat vd. (2014)'nin çalışmasında özetlemiştir. Duygu analizinin güncel kullanım alanları için Kumar ve Jaiswal'in (2020) yapmış olduğu literatür çalışmasına başvurulabilir.

Duygu analizi çalışmaları her ne kadar popüler olsa da geliştirilen algoritmalar ve kütüphaneler ele alındığında, İngilizce üzerine yapılan çalışmalar ön plandadır ve farklı dillerin analizine imkân sağlayan yöntemlerin ve kütüphanelerin daha fazla gelişmeye ihtiyacı vardır. Bu motivasyondan yola çıkılarak, bu çalışma ile Türkçe duygu analizi literatürü incelenecektir. Çalışmanın temel amacı, Türkçe metinlerde duygu analizi üzerine son dönemlerde yapılan çalışmaların kapsamını

ve genişliğini göstermek, bulgularını özetlemek ve mevcut literatürdeki boşlukları belirlemektir. Bu bağlamda, Türkçe duygu analizi alanında yapılan güncel çalışmaları ve bu çalışmalarda kullanılan yöntemler, özellikle sosyal bilimler alanındaki duygu analizi uygulamaları, kullanılan açık kaynaklı kütüphaneler ve veri tabanlarını ortaya konulmuştur. Buna bağlı olarak, çalışma ile hem bu alanda çalışma yapacak araştırmacılara yön göstermek, hem de ileri araştırma alanlarını ortaya koymak hedeflenmiştir.

Çalışma şu şekilde düzenlenmiştir; birinci kısımda çalışma metodolojisi ve incelenen literatüre ait özet verilmiş, ikinci kısımda Türkçe duygu analizi literatüründe sıklıkla kullanılan öznelik seçimi, makine öğrenme ve sözlük tabanlı yöntemler irdelenmiştir. Dördüncü kısımda Türkçe duygu analizinde kullanılan açık kaynaklı kütüphaneler ve veri tabanları hakkında bilgi verilmiş ve son kısımda literatürde gözlemlenen eksiklikler ve üzerinde araştırma yapılabilecek açık konular özetlenmiştir.

2. Metodoloji

Bu çalışmada, Türkçe metinlerde duygu analizi yazını incelenirken kapsam derlemesi yöntemi izlenmiştir. Kapsam derlemesi yöntemi, belirli bir konu etrafındaki literatürün çerçevesini çizmek için tasarlanmış tanımlayıcı bir yaklaşımdır ve yöntem ile elde edilecek bulgular yazındaki boşlukları belirleyerek gelecekteki araştırmalara ışık tutmaktadır. Bu bağlamda, son 10 yıl içerisinde hakemli dergilerde yayınlanmış ve ilgili enstitülerce kabul almış her türlü Türkçe duygu analizi çalışmaları taranmış ve 43 adet yayın ve 12 adet tez (yüksek lisans ve doktora) incelenmiştir. İncelenen yayınların içeriği farklı başlıklar altında değerlendirilmiş ve çalışmalarda kullanılan yöntemler ve bulguları özetlenerek, yazının genel çerçevesi elde edilmiştir. Tablo 1’de incelenen yayınların içeriği özet olarak verilmiştir. Üçüncü kolonda yayının kullandığı yaklaşımı (Sözlük Tabanlı – LB veya Makine Öğrenimi – ML) gösterilirken, dördüncü kolonda ilgili çalışmada kullanılan yöntem belirtilmiştir. Beşinci kolon ilgili çalışmada duygu kutbu incelenip incelenmediği, altıncı kolonda ise çalışmada analiz edilen verinin içeriği ve sekizinci kolonda veri setinin kaynağı gösterilmiştir. Son iki kolonda ise, sırasıyla çalışmada, eğer bildirilmişse, kullanılan programlama dili ve ön işleme kütüphanelerine yer verilmiştir.

Tablo 1. İncelenen Literatür Özet Tablosu

<i>Referans</i>	<i>M L/ LB</i>	<i>Algoritma</i>	<i>İçerik</i>	<i>Veri Kaynağı</i>	<i>Progra mlama Dili</i>	<i>Ön İşleme Kütüphane</i>
<i>Kaya, Fidan ve Torosoğlu (2012)</i>	M L	NB, ME, SVM, N-Gram	Politik olaylar ile ilgili haberler	Haber siteleri	Java	ZEMBEREK
<i>Vural vd. (2013)</i>	LB	Dictionary lexicon	Film yorumları	beyazperde	Java	ZEMBEREK
<i>Türkmenoğlu ve Tantuğ (2014)</i>	LB - M L	DB Lexicon and SVM, NB and Decision Trees	ML ve LB yöntemleri karşılaştırması	twitter, beyazperde	N/A	Manual
<i>Nizam ve Akın (2014)</i>	M L	NB, RF, SMO, DT, NN1	Ürün yorumları	twitter	Java (WEKA)	Manual
<i>Akba vd. (2014)</i>	M L	SVM, NB	Film yorumları	beyazperde	R	ZEMBEREK
<i>İskender ve Batı (2015)</i>	M L	Multinomial Naïve Bayes (MNB), SVM	Girişimcilik	twitter	N/A	N/A
<i>Çoban, Özyer ve Özyer (2015)</i>	M L	SVM, NB, MNB, KNN	Performans değerlendirme	twitter	Java	ZEMBEREK
<i>Nalçakan, Bayramoğlu ve Tuna (2015)</i>	M L	NB, RF	Ürün yorumları	twitter	Python	Manual
<i>Gözükkara ve Özel (2016)</i>	M L	SVM	Ürün yorumları	-	C#	Manual
<i>Akgül, Ertano ve Diri (2016)</i>	LB	Dictionary lexicon	Genel veri seti	twitter	N/A	N/A
<i>Kaynar vd. (2016)</i>	M L	NB, CBC, MLP, SVM	Film yorumları	beyazperde	MatLab, Python	NLTK
<i>Parlar ve Özel (2016)</i>	M L	NB	Ürün yorumları	-	Python	NLTK
<i>Çatal ve Nangir (2017)</i>	M L	Multiple Classifiers	Performans değerlendirme	Antoloji, beyazperde, hepsiburada	N/A	N/A
<i>Kama vd. (2017)</i>	LB	aspect - sentiment word mapping algorithm	Ürün yorumları	donanım haber	N/A	ZEMBEREK
<i>Bilgin ve Şentürk (2017)</i>	M L	Semi-supervised vector algorithm	Genel veri seti	twitter	Python	Manual
<i>Seyfioğlu ve Demirezen (2017)</i>	M L	Semi-supervised vector algorithm	Havayolu müşteri değerlendirmeleri	-	Python (Gensim)	ZEMBEREK
<i>Dehkharghani vd. (2017)</i>	LB	Dictionary and NLP	Film yorumları	beyazperde	N/A	ITU NLP
<i>Pervan ve Keleş (2017)</i>	M L	Word2Vec RF	Ürün yorumları	-	Python	Manual
<i>Kamışlı, Erçiçek ve Ergül (2017)</i>	M L	NB	Öğrenci yorumları	twitter	Python	NLTK

<i>Öztürk ve Ayvaz (2018)</i>	LB	NLP	Suriye krizi ve mülteciler	twitter	R	Rsentiment (English), Manual (Turkish)
<i>Çetin ve Eryiğit (2018)</i>	LB	sequence tagging, CRF, Linear classifier	Restoran yorumları	SemEval ABSA 2016	Python	ITU NLP Web servisi
<i>Karaöz ve Gürsoy (2018)</i>	LB	Adaptive learning	TV programı yorumları	twitter	R, excel	N/A
<i>Akın ve Gürsoy (2018)</i>	LB	Adaptive learning	TV programı yorumları	twitter	R, excel	N/A
<i>Velioğlu, Yıldız ve Yıldırım (2018)</i>	M L	NB, LR, SVM, DT	Genel veri seti	twitter	Python	FastText
<i>Karamollaoğlu vd. (2018)</i>	LB	Semantic	Genel veri seti	twitter	Java	ZEMBEREK
<i>Çiftçi ve Apaydın (2018)</i>	M L	Deep Learning, RNN	Ürün yorumları, film yorumları	hepsiburada, beyazperde	Python	ITU NLP
<i>Dehkharghani (2018)</i>	LB	Hybrid lexicon and corpus algorithm	Film yorumları, otel yorumları	beyazperde, Semeval-2016	N/A	N/A
<i>Dülger (2018)</i>	M L	NB, LR, DT, RF, MP	Genel veri seti	-	Java	ZEMBEREK
<i>Parlar, Özel ve Song (2018)</i>	M L	BN.SVM,ME,DT	Film yorumları, ürün yorumları	beyazperde, hepsiburada	Python	NLTK
<i>Shehu vd. (2019)</i>	M L	polarity lexicon (PL) and artificial intelligence (AI).	Performans değerlendirme	twitter	R	ZEMBEREK
<i>Alpkocak vd. (2019)</i>	M L	ANN, SVM, RF, KNN	Performans değerlendirme	TREMO	N/A	Manual
<i>Atan ve Çınar (2019)</i>	LB	Dictionary Based	Finansal piyasalar	newspapers	R	ZEMBEREK
<i>Karcioğlu ve Aydın (2019)</i>	M L	Linear SVM, LR	Genel veri seti	twitter	Python	Manuel, Porter Stemmer, TurkishStemmer
<i>Uslu, Tekin ve Aytakin (2019)</i>	LB - M L	Dictionary based Lexicon, NB, LR, DT	Film yorumları	beyazperde	N/A	ZEMBEREK
<i>Santur (2019)</i>	M L	Gated Recurrent Unit	Ürün yorumları	hepsiburada	Python	N/A
<i>Bayraktar, Yavaoğlu ve Özbilen (2019)</i>	M L	LDA, C-Value, WSBFE	Restoran yorumları	SemEval ABSA 2016	Java	ZEMBEREK, Yandex.XML
<i>Demirci, Keskin ve Doğan (2019)</i>	M L	LR,SVM,RF,NN	Politik olaylar ile ilgili haberler	twitter	Python	Turkish NLP
<i>Yurtalan, Koyuncu ve Turhan (2019)</i>	LB	Dictionary lexicon	Genel veri seti	twitter	N/A	ZEMBEREK
<i>Erşahin vd. (2019)</i>	Hy bri d	Dictionary lexicon, NB, SVM	Film yorumları, otel yorumları, twitter	beyazperde, otelpuan, twitter	N/A	ZEMBEREK, ITU NLP
<i>Karagöz vd. (2019)</i>	LB	LSTM, CRF	Ürün yorumları	donanım haber	Python	ZEMBEREK

<i>Ahmetođlu ve Daş (2020)</i>	M L	Recurrent NN	Otel yorumları	-	Python	NLTK
<i>Açıklalın, Bardak ve Kutlu (2020)</i>	M L	BERT (NN)	Film yorumları, otel yorumları	beyazperde, otelpuan	N/A	FastText, ITU NLP
<i>Shehu ve Tokat (2020)</i>	M L	SVM, RF	Film yorumları, twitter	twitter, beyazperde	N/A	ZEMBEREK

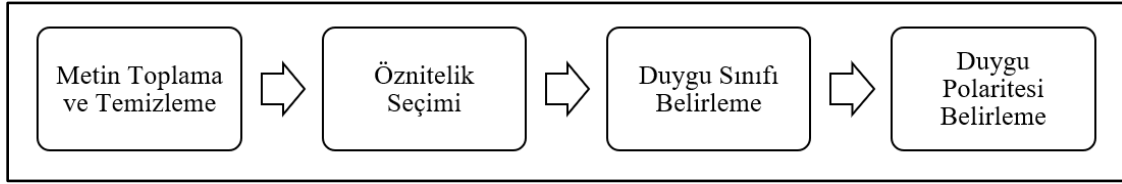
Taranan yazın iki ana başlık altında incelenmiş olup, ilk kısımda Türkçe metinlerde duygu analizinde kullanılan yöntemler ele alınmış, ikinci kısımda ise Türkçe metinlerde duygu analizinde kullanılan açık kaynaklı kütüphaneler ve veri tabanları özetlenmiştir.

3. Türkçe Metinlerde Duygu Analizi

Duygu analizi çalışmalarında, cümle, metin (doküman) ve hedef kategori seviyesi olmak üzere üç farklı seviyede analiz yapmak mümkündür. Cümlelerin de birer metin olması varsayımından yola çıkarak cümle ve metin seviyesi analizlerin kapsamının aynı olduğunu söylemek mümkündür. Her ne kadar uzun metinler duyguyu betimlemede daha elverişli olsa da kısa metinlerin kirli bilgiden arınmış olması sebebiyle duygu analizini kolaylaştırdığını gösteren çalışmalar da bulunmaktadır (Birmingham, 2010). Bu sebeple, sosyal medya analizlerinde twitter platformunun sıklıkla kullanıldığı görülmektedir (Thelwall vd., 2011). Metin ve cümle seviyesi analizlerde, bir metnin öznel fikir veya duygu içerip içermediği ve öznel fikir var ise duygu polaritesi irdelenmektedir. Bu anlamda, duygu analizinde iki veya üç duygu kategorisi ya iki (pozitif ve negatif) ya da 3 (pozitif, negatif ve nötr) kategori olduğu gözlenmektedir. Metnin ve cümle seviyesinde analizler ile metinde incelenen konuya dair tüm özellikleri elde etmek güçtür. Örneğin bir ürün için fiyat ve kalite iki farklı performans ölçütü olabilir ve bir metin içerisinde her iki ölçüt için farklı duygular belirtilmiş olabilir. Bu sebeple bu ürün ile ilgili duygu farklı kategoriler altında değerlendirilmelidir. Bu tür ayrıntıları elde etmek hedef tabanlı duygu analizi (*Aspect Based Sentiment Analysis*) ile mümkündür. Hedef tabanlı duygu analizinde temel amaç incelenen olguya dair duyguyu belirli kategoriler altında incelemektir. Türkçe metinlerde duygu analizi literatürü incelendiğinde, metin seviyesinde yapılan çalışmaların daha yoğun olduğu hedef tabanlı analizlerin daha az çalışılan bir alan olduğu görülmüştür.

Duygu analizi süreci Şekil 1’de gösterildiği üzere veri toplama ve temizleme işlemi ile başlamaktadır. Temizlenen ve kirli veriden ayıklanan veri seti ile duygu analizi yapmak, analiz performansını arttırmaktadır. Bir diğer analiz performansını arttıran işlem ise ikinci aşamada yer alan öznitelik seçimidir. Üçüncü aşamada ise, analiz yöntemi belirlenerek amaca ve metne uygun

model belirlenir ve sınıflandırma işlemi gerçekleştirilir. Son olarak duygu sınıflarının polaritesi belirlenir.



Şekil 1. Duygu Analizi Süreci

Duygu analizi süreçleri ele alındığında, analizin sonucunu ve kalitesini etkileyen en önemli iki nokta, öznitelik seçimi ve duygu kategorisi belirlenirken kullanılacak yöntemdir. Bu bölümün devamında öncelikle Türkçe metinlere uygulanan duygu analizi çalışmalarında kullanılan öznitelik seçimi yöntemleri, sonrasında duygu kategorisi belirleme yöntemleri incelenecektir.

3.1. Türkçe Metinlerde Duygu Analizinde Kullanılan Öznitelik Seçimi Yöntemleri

Duygu analizinde öznitelik seçimi, metin içerisinde sıklıkla tekrarlanan niteliklerin belirlenmesi, az rastlanan niteliklerin elenerek metnin ana karakterini belirleme çalışmasıdır ve üç temel amacı vardır. Bu amaçlar; (i) karmaşıklığı azaltması sebebiyle hesaplama zamanında fayda, (ii) veri ile fazlaca uyumlu analiz yapmayı önleme ve (iii) sınıflandırma doğruluğunun geliştirilmesidir. Bu bağlamda öznitelik seçimi duygu analizinin ve sınıflandırmanın performansında oldukça belirleyicidir.

Öznitelik olarak bir takım farklı özellikler bir arada kullanılabilir. Bu amaca yönelik olarak, kelime, kelime, cümle parçası, n-gram, sözcük türü etiketleme (*Part of Speech tags – POS tags*), bağımlılık ilişkisi (*dependency information*), ifade ikonlar (*emoticon*), noktalama işaretleri ise büyük harf kullanımı gibi metin içerisindeki duyguyu anlamlandıracak öznitelikler sıklıkla kullanılmaktadır (İskender ve Batı, 2015). Aynı zamanda, kelime torbası (*bag of words*) kullanımının oldukça yaygın olduğu görülmektedir (Çatal ve Nangir, 2017).

Öznitelik seçiminde öncelikli işlem metin temizleme işlemidir. Bu aşamada, “ve”, “bir” gibi duygu anlamı içermeyen gereksiz kelimelerin metinden çıkartılmakta ve kelime köklerinin tanımlanması gibi metin sadeleştirme işlemleri yapılmaktadır. Türkçe duygu analizi çalışmalarında metin sadeleştirme aşamasında kullanılan açık kaynaklı kütüphaneler bulunmaktadır ve sıklıkla kullanılan kütüphanelerin Zemberek-NLP (Akın ve Akın, 2007) ve ITU-NLP (Eryiğit, 2014) olduğu görülmüştür.

Ön işleme ve özniteliklerin belirlenmesinden sonraki aşama ise, özniteliklerin seçimidir ve

uygulamanın kolay olması sebebiyle Ki-Kare (*Chi-Square*), Bilgi Kazanımı (*Information Gain*), ve Gömme Algoritmaları (*Embedding*) gibi yöntemler sıklıkla kullanılmaktadır. Ki-kare, belirlenen kategorik değişkenler ile öznitelik arasında anlamlı bir ilişki olup olmadığını belirleyen istatistiksel bir yöntemdir. Bilgi Kazanımında ise bir özneliğin önemi, diğer alternatif özniteliklerin önemine oranına bakılarak belirlenir ve belirli bir alt sınır üzerinde öneme sahip olan öznitelikler seçilir. Gömme Algoritmalarında ise öznitelik seçimi ve sınıflandırma süreçleri eş zamanlı olarak yürütülmekte ve makine öğrenimi aşamasında elde edilen bilgiler kullanılarak belirgin öznitelikler seçilmektedir.

Türkçe metinler üzerinde yapılan bazı çalışmalarda farklı öznitelik seçimlerinin analiz performansı üzerindeki etkisi incelenmiştir. Bu doğrultuda, Parlar (2016) çalışmasında sıklıkla kullanılan Ki-Kare ve Bilgi Kazanımı yöntemleri ile Metin Fark Frekansı (*Document Frequency Difference*) ve Optimal Dikey Ağırlık Merkezi (*Optimal Orthogonal Centroid*) yöntemlerinin duygu analizi performansı üzerindeki etkisini araştırmıştır. Çalışmada, anlamlı şekilde üstün olan bir öznitelik seçimi belirlenmemekle birlikte, noktalama işaretlerinin ve etkisiz kelimelerin nitelik olarak kalmasının analiz performansı üzerinde olumlu etkiye sahip olduğu belirlenmiştir. Bu çalışmadan yola çıkarak, Parlar ve Özel (2016) yeni bir öznitelik seçimi yöntemi (*Query Expansion Ranking*) geliştirmişler ve önerilen yöntemi kullanarak Ki-Kare ve Metin Fark Frekansı yöntemlerine göre daha iyi analiz performansı elde ettiklerini ortaya koymuşlardır. Gözükara ve Özel (2016) ise yapmış oldukları çalışmada, Ki-kare ve Bilgi Kazanımı yöntemlerinin analiz sonuçlarında anlamlı farklılık olmadığını ancak ikili kelime gruplarının analiz performansını iyileştirdiğini göstermişlerdir.

Kelime grupları oluşturma'nın analiz performansı üzerindeki etkileri açısından inceleme yapan çalışmalarda ise genellikle izlenen stratejinin farklı n-gram gruplamanın analiz performansı üzerindeki etkisini inceleme olduğu gözlemlenmiştir. Türkmenoğlu ve Tantuğ (2014) aynı veri seti üzerinde uyguladıkları farklı makine öğrenme yöntemlerinin tamamında iki kelimelik gruplamanın analiz performansını iyileştirdiğini ortaya koymuşlardır. Çoban, Özyer ve Özyer (2015) kelime torbası ve farklı seviyede kelime grupları ile yapılan analizlerin sonuçlarını karşılaştırmış ve 3-gram gruplamanın en iyi performansı gösterdiği belirtmiştir. Kaya, Fidan ve Toroslu (2012) makine öğrenimi kullanarak yaptıkları duygu analizi çalışmasında farklı seviyelere ve özelliklere sahip kelime gruplarının analiz performansı üzerindeki etkisini incelemişler ve sıfat ekine sahip tek kelime gruplarının diğer seviyelere göre daha iyi performans gösterdiğini ortaya

koymuşlardır. Ahmetođlu ve Daş (2020) ise farklı kapsamlara sahip kelime torbaları kullanarak Word2Vec algoritması ile eğitilmiş modellerin performansını değerlendirmiş ve veri setinin kelime modeli olarak kullanıldığı durumun daha anlamlı ilişki kurguladığı belirtmişlerdir.

Duygu analizinde ironi içeren metinlerin belirlenmesi karşılaşılan bir diğer zorluktur ve öznitelik seçiminde dikkate alınması gereken bir husustur. Diğer dillerde, özellikle İngilizce metinlerde ironinin tespit edilmesi üzerine birçok çalışma olmasına rağmen (Farias, Benedi ve Ross, 2015; Farias, Patti ve Rosso, 2016), Türkçe çalışmalarda bu konu üzerine yapılan çalışma sayısının az olduğu gözlenmiştir. Bu kapsamda, Taşlıođlu (2014) tez çalışmasında noktalama işaretlerinin ironi ile olan ilişkisi üzerinde durmuş, belirli noktalama işareti kullanımlarını öznitelik olarak duygu analizinde kullanmış ve parantezli ünlem işareti, ünlem işareti ve soru işaretinin ironi içeren ifadeleri tespit edebilecek öznitelikler olduğunu ifade etmiştir. Dölger (2018) ise, 3 farklı kategoride (noktalama işareti, duygusal zıtlıklar ve zamansal zıtlıklar) belirlediği kurallar çerçevesinde, ironiye işaret eden öznitelikleri belirlemiş ve farklı makine öğrenme yöntemleri ile yaptığı analizler sonucunda doğruluk performansı yüksek sonuçlar elde etmiştir.

3.2. Türkçe Metinlerde Duygu Analizinde Kullanılan Duygu Sınıflaması Yöntemleri

Duygu analizinde duygu sınıflandırması yapılırken makine öğrenimi (*machine learning*) ve sözlük-tabanlı (*lexicon-based*) olmak üzere iki temel yaklaşım söz konusudur. Makine öğrenimi yaklaşımları, denetimli (*supervised*) ve denetimsiz (*unsupervised*) öğrenme olmak üzere iki ana kategoride toplanmıştır (Medhat vd., 2014) ve her ikisi de duygu analizinde sıklıkla kullanılmaktadır (Thelwall vd., 2010).

İki yöntemin kapsamı farklılık göstermekle birlikte, alan odaklı çalışmalarda makine öğrenme yöntemleri daha sık kullanılmaktayken, ölçeklenebilir olması bakımından genel analizlerde sözlük temelli yaklaşımlar daha çok tercih edilmektedir (Sađlam, 2019). Literatür incelendiğinde, makine öğrenme yaklaşımlarının daha çok kullanıldığı görülmektedir. Sözlük tabanlı yaklaşımlar, kelimelerin sözlüksel anlamını ve eşanlamlılarını incelemesi sebebiyle, metnin içerdiği gerçek anlam ve duyguyu tam olarak yakalamamaktadır ve bu durum yaklaşımların duygu kategorisini belirleme performansını azaltmaktadır (Thelwall, 2011). Ayrıca, makine öğrenme yöntemlerinin kolaylıkla uygulanmasına imkan sağlayan açık kaynaklı kütüphanelerin bulunması da bu yaklaşımlara olan eğilimi arttırmaktadır. Benzer bir eğilim Türkçe metinlerde duygu analizinde de görülmektedir. İncelenen literatürde 43 adet makaleden 29 (%60) adedinde makine öğrenme yöntemleri kullanıldığı görülmüştür.

Bu bölümün devamında, bu yöntemler ve ilgili çalışmalardaki uygulamalar, denetimli ve denetimsiz makine öğrenimi yaklaşımları ve sözlük tabanlı yaklaşımlar başlıkları altında incelenmiştir. Devamında, makine öğrenme ve sözlük tabanlı yaklaşımların beraber kullanıldığı melez yaklaşımlar ve iki yöntemin performanslarının karşılaştırıldığı çalışmaların sonuçları özetlenmiştir.

3.2.1. Denetimli Makine Öğrenme Yaklaşımları

Denetimli makine öğrenme yaklaşımlarında temel amaç, etiketli eğitim verisi kullanılarak duygu kategorisini belirleyecek modeli tahmin etmektir. Eğitilen algoritma sonucu ortaya konulan model doğrulama veri seti üzerinde uygulanarak algoritmanın duygu kategorisi belirleme performansı ölçülür. Yukarıda da belirtildiği üzere, öznelik seçiminin, uygulanacak modelin performansı üzerinde büyük etkisi bulunmaktadır.

Türkçe metinlerde duygu analizi çalışmaları incelendiğinde, sıklıkla Naive Bayes (NB), Maksimum Entropi (ME), Destek Vektör Makinesi (SVM: *Support Vector Machine*) ve Rastgele Orman (*Random Forest*) sınıflandırıcılarının kullanıldığı görülmüştür. Naive Bayes sınıflandırıcıda bir özelliğin görülme olasılığı hesaplanarak, olasılık değeri yüksek özelliklere göre sınıflandırma yapılmaktadır. Uygulaması kolay ve hızlı çalışan bir yöntem olması sebebiyle sıklıkla kullanılmaktadır. Destek Vektör Makinesi ise, iki sınıfın maksimum uzaklıkta olmasını amaçlar ve NB sınıflandırıcıya benzer şekilde uygulanması kolay olduğu için çalışmalarda yoğun olarak kullanılmıştır. Literatürdeki farklı çalışmalarda, iki yöntem üzerine kurgulanan modellerin performansının yüksek olduğu raporlanmıştır. Örneğin Parlar vd. (2018) tarafından yapılan öznelik seçiminin performans üzerindeki etkisini ölçmeye yönelik çalışmada ise, çokterimli NB, SVM, Lojistik Regresyon ve Karar Ağacı sınıflandırma algoritmaları denenmiş ve çokterimli NB yönteminin diğer yöntemlere göre daha iyi performans gösterdiği raporlanmıştır. NB ve SVM dışında yüksek performans gösteren bir diğer yöntem ise Maksimum Entropi'dir. Kaya, Fidan ve Torosoğlu (2012) çalışmasında, n-gram öznelik seçimi ile kurguladıkları ve Türkçe haber metinleri üzerinde uyguladıkları NB, ME ve SVM modelleri arasında ME modelinin en iyi performansı gösterdiğini raporlamışlardır. Çatal ve Nangir (2017) çalışmasında ise, NB, SVM ve *Bagging* yöntemlerinin bir arada kullanıldığı *çoklu sınıflandırma algoritması* önermiş ve klasik sınıflandırma yöntemlerinin bir arada kullanımının duygu analizi performansını arttırdığını ortaya koymuşlardır.

3.2.2. Denetimsiz Makine Öğrenme Yaklaşımları

Denetimsiz makine öğrenmede etiketsiz veri kullanan modelin, verideki bilinmeyen ilişkileri ortaya koyması ve sınıflandırma için faydalı olan özellikleri veya kuralları ortaya çıkarması temel amaçtır. Bu noktada sıklıkla kullanılan yöntemler kural tabanlı yaklaşımlardır. Noktasal Karşılıklı Bilgi Katsayısı (*PMI – Pointwise Mutual Information*) kelime türlerinin bir arada bulunma olasılıkları üzerinden yüksek karşılıklı bilgi katsayısına sahip özelliklerin belirlenmesi işlemidir. Modelde bu bilgi kullanarak belirlenen kurallar çerçevesinde duygu polaritesi belirlenmektedir (Turney, 2002). Başka bir istatistiksel yaklaşım olan Örtülü Anlam Dizinleme (*LSI - Latent Semantic Indexing*), bir kelimenin farklı anlamlarından yola çıkarak, cümle içerisinde hangi anlamda kullanıldığını tespit etmeye ve cümlenin doğru anlaşılmasına yönelik analiz yöntemidir. LSI ile metni eksen düzeninde ifade edilerek, kelimelerin olası anlamları hakkında daha bilgi elde edilir ve en yüksek olasılıklı anlam seçilir. LSI yaklaşımının bir uzantısı olan Örtülü Dirichlet Dağılımı (*LDA - Latent Dirichlet Allocation*) ise LSI'nin olasılık yaklaşımını kullanarak metni anlam ve konu ekseninde düzenler ve metinde açıkça yer almayan konuların tespit edilmesine imkan sağlar.

Türkçe duygu analizinde denetimsiz makine öğrenme yaklaşımları kullanılarak yapılan çalışmaların sayısı oldukça azdır. Metin ve Karaoğlan (2016) tarafından doğal dil işleme üzerine yapılan çalışmada, Türkçede eşdizimlerin belirlenmesinde gözlenme sıklığı (*Occurrence Frequency*) ve PMI yöntemleri kullanılmış ve PMI'nin daha iyi performans gösterdiği raporlanmıştır.

3.2.3. Sözlük Tabanlı Yaklaşımlar

Sözlük tabanlı makine öğrenme yaklaşımları, pozitif veya negatif duygu içeren kelimelerin belirli durumlarda ortaya çıktığı varsayımından yola çıkar ve metin içinde duygu içeren kelimelerin tespit edilmesi yoluyla metni pozitif veya negatif olarak etiketler. Sözlük temelli duygu analizinde genel duygu sözlüğü kullanılmaktadır. Sözlük kullanımının bir dezavantajı, sözlükteki terimlerin bağlamdan bağımsız olarak değerlendirilmesi ve bu yüzden performans kaybına sebep olmasıdır (Sağlam, 2019). Sözlük temelli yaklaşımlarda izlenen genel uygulama, diğer kelime sözcük türlerine göre sıfatların duygu betimlemede daha büyük bir öneme sahip olduğu varsayımdır. Ayrıca, “çok”, “gerçekten” gibi kuvvetlendiricilerin kullanımıyla terime ait duygu polaritesinin belirlenmesi bir diğer uygulamadır. Bu bağlamda, Dehkharghani (2015) çalışmasında kuvvetlendirici ifadelerin sınıflandırma performansına katkısı olduğunu gösterilmiştir. Bu

çalışmadan yola çıkarak, Dehkharghani (2018) duygu polaritesi sözlüğü oluşturmak için farklı sınıflandırma algoritmalarını kullanan melez bir yaklaşım önermiş ve önerilen algoritmanın sınıflandırma performansını %71-%91 aralığında olduğunu raporlamıştır. Yurtalan, Koyuncu ve Turhan (2019) ise sözlük tabanlı sınıflandırma modelinde Türkçe dilinin yapısal özelliklerini kullanarak, kelime, ikili ve üçlü kelime grupları, deyim ve atasözlerinin sınıflandırma performansını arttırdığını göstermişlerdir.

Sözlük tabanlı makine öğrenme yöntemleri hedef tabanlı duygu analizinde de sıklıkla kullanılmaktadır. Çetin ve Eryiğit (2018) kelime vektörleri ve sözcük ve cümle analizi bilgilerini kullanarak CFR (*Conditional Random Fields*) uygulamış ve kullanılan yöntemin hedef kategori belirlemede daha başarılı olduğunu raporlamışlardır.

Türkçe duygu analizi literatüründe makine öğrenme ve sözlük tabanlı yöntemleri bir arada kullanan melez modeller de kullanılmaktadır. Erşahin vd. (2019) öznitelik belirlemek için sözlük tabanlı bir öznitelik üretme algoritması geliştirmişlerdir. Bu algoritmanın çıktıları, farklı makine öğrenme sınıflandırıcılarına girdi olarak verilmiş ve öznitelik üretme algoritmasının, özellikle NB sınıflandırıcısının performansını oldukça arttırdığı raporlanmıştır.

Türkçe metinlerde sözlük tabanlı yaklaşımlar ve makine öğrenme yaklaşımlarının karşılaştırmasını yapan çalışmalar incelendiğinde, genel olarak makine öğrenme yaklaşımlarının daha yüksek performans gösterdiği raporlanmıştır. Uslu, Tekin ve Aytekin (2019) çalışması kapsamında ANEW kütüphanesinden çeviri yaparak oluşturdukları Türkçe duygu sözlüğü oluşturmuşlardır. Aynı veri seti üzerinde bu kütüphanenin ve farklı makine öğrenme yaklaşımlarının (NB, LR ve Karar Ağacı) performansını karşılaştırmış ve makine öğrenme modelleri ile daha yüksek doğruluk oranına sahip sonuçlar elde etmişlerdir. Ayrıca, bu çalışma ile veri kümesinin boyutu ile makine öğrenme algoritmalarının performansı arasında doğrusal bir ilişki olduğunu ve özellikle büyük veri setlerinde makine öğrenme yaklaşımlarının performansının sözlük tabanlı yaklaşımlardan daha yüksek olduğunu vurgulamışlardır. Bunun yanı sıra, Türkmenoğlu ve Tantuğ (2014) yapmış olduğu çalışma ile çeşitli makine öğrenme ve sözlük tabanlı duygu analizi yöntemlerini karşılaştırmış ve makine öğrenme yöntemlerinin, özellikle tekli ve ikili kelime grupları ile çalıştırılan NB ve SVM modellerinin daha yüksek sınıflandırma performansına sahip olduğunu tespit etmiştir.

Sözlük tabanlı yaklaşımlar ile Doğal Dil İşleme yaklaşımları birlikte kullanılarak sözdizimsel ilişkilerin ortaya konulması mümkündür (Methat, Hassan ve Korashy, 2014). Çalı ve Balaman

(2019) müşteri yorumlarına uyguladıkları duygu analizi sonuçlarını bir Çok Amaçlı Karar Verme algoritması olan VIKOR ile sıralamış ve bir ürün önerme sistemi geliştirmişlerdir. Kullandıkları duygu analizi algoritmasında sözlük yaklaşımı ile kelime ve terimlere ait duygu skorları belirlenmiş, bu skorlar ve dil işleme kuralları birlikte kullanılarak sınıflandırma yapılmıştır.

İncelenen literatürde, sözlük tabanlı yaklaşımların uygulandığı çalışmalarda gözlemlenen bir diğer unsur ise, değişik konular üzerindeki toplumsal fikri veya olguların birbirleri ile olan ilişkilerini ortaya koyan çalışmalarda sıklıkla kullanılmasıdır. Makine öğrenme metotlarını kullanan çalışmalarda genellikle sınıflandırma performansı vurgulanmaktayken, sözlük tabanlı çalışmalarında ele alınan konuların farklılık gösterdiği görülmektedir. Öztürk ve Ayvaz (2018) Suriyeli mülteciler hakkında Türkçe ve İngilizce twitter iletileri üzerinde uyguladıkları sözlük tabanlı duygu analizi ile Türkçe ve İngilizce tweet'lerin duygu polaritesini belirlemişler ve iki dildeki polaritenin farklılıklarını ortaya koymuşlardır. Atan ve Yetkin (2019) ise finansal piyasalarda yayınlanan haberlerin duygu polaritesini Türkçeye çevrilmiş duygu sözlüğü kullanarak belirlemişlerdir. Buna bağlı olarak, haberlerin duygu polaritesi ile BIST30 şirketlerinin finansal değerleri arasında anlamlı bir ilişki olduğunu ortaya koymuşlardır. Karaöz ve Şimşek (2018 (a)) yarı denetimli makine öğrenme ve sözlük tabanlı yöntemleri birlikte kullandıkları model önermişlerdir. Akın ve Şimşek (2018(b)) ise bu model üzerinde kurguladıkları çalışmada, televizyon programları hakkında paylaşılan tweet'leri yapısal hale getirmişler ve programa ait duygu ile programın reytingi arasındaki ilişkiyi ortaya koymuşlardır.

4. Türkçe Metinlerde Duygu Analizinde Kullanılan Açık Kaynaklı Kütüphaneler ve Veri Tabanları

Çalışmanın bu kısmında Türkçe metinlerde duygu analizinde kullanılan açık kaynaklı kütüphaneler ve sıklıkla kullanılan veri tabanları üç ana başlık altında incelenmiştir. İlk kısımda genel amaçlı dil işleme kütüphaneleri, ikinci kısımda özellikle sözlük tabanlı yöntemler için geliştirilmiş yaygın olarak kullanılan genel amaçlı duygu kütüphaneleri, farklı programlama dillerinde sıklıkla kullanılan makine öğrenme paketleri ve son olarak literatürde sıklıkla kullanılan açık kaynaklı veri tabanları irdelenmiştir.

4.1. Genel Amaçlı Dil İşleme Kütüphaneleri

Duygu analizinin öncül aşamalarından olan, toplanan işlenmemiş verinin ön işleme sürecinde, genel amaçlı dil işleme kütüphanelerinin kullanımı oldukça yaygındır. Türkçe metinlerde,

Türkçe'nin morfolojik yapısını dikkate alarak geliştirilmiş oldukça etkili doğal dil işleme kütüphaneleri bulunmaktadır. Bu kütüphaneler arasında en sık kullanılan ve performansı en yüksek kütüphane Zemberek'dir. Zemberek Java dilinde yazılmış açık kaynaklı bir Doğal Dil İşleme (DDİ) kütüphanesi olup, özellikle metin temizleme ve ön işleme aşamalarında sıklıkla kullanılmaktadır. Yazım denetimi, hatalı kelimeler için öneri, heceleme, hatalı kodlama temizleme işlevlerinin yanı sıra, belirteçlere ayırma, cümle bölümlenme ve sözcük ayırma (*tokenizations, sentence segmentation, lemmatizations*) gibi işlevlere sahiptir. Güncel sürümü Zemberek-NLP olup, bu sürüm ile metin sınıflandırma özelliği de kütüphaneye dahil edilmiştir. Ancak, literatürde ön işleme sürecinde sıklıkla Zemberek kütüphanesinin kullanılmasına rağmen bu kütüphane kullanılarak sınıflandırma yapılan çalışmaya rastlanmamıştır. Zemberek kütüphanesine GitHub üzerinden erişmek mümkündür¹.

Bir diğer DDİ kütüphanesi olan ITU NLP Turkish Parser (Eryigit 2014) İTÜ Doğal Dil İşleme Grubu tarafından geliştirilmiş olup, hizmet olarak yazılım (*Software as a service – SaaS*) modelinde çalışan bir doğal dil işleme kütüphanesidir. İnternet tarayıcı aracılığı ile kullanım imkanının yanı sıra Uygulama Programlama Arayüzü (*Application Programming Interface -API*) erişimi yoluyla da kullanmak mümkündür. Kütüphane temelde hata ayıklama, cümle bölümlenme gibi özelliklerin yanı sıra yeniden işleme, morfoloji, sözdizimi ve varlık tanıma (*reprocessing, morphology, syntax and entity recognition*) işlemlerini çözümlenmektedir. İTÜ-NLP SaaS uygulamasına herhangi bir tarayıcı üzerinden erişmek mümkündür².

4.2. Genel Amaçlı Duygu Kütüphaneleri

Duygu sözlüğü, duygu analizi için tasarlanmış sözlük tabanlı bir kaynaktır ve duygu analizinde, özellikle sözlük tabanlı yöntemlerde kullanılmaktadır. Dünyada yaygın olarak kullanılan WordNet bir dildeki kelimelerin tanımlarını, eş anlamlılarını ve kelimeler arasındaki farklı seviyedeki anlambilimsel ilişkileri tanımlayan oldukça kapsamlı bir veritabanıdır ve günümüzde 200'den fazla dil için tasarlanmış WordNet bulunmaktadır. Türkçe WordNet ise 2001-2004 yılları arasında geliştirilen ve 14K anlamdaş küme ve 20K terimden barındıran ilk ve tek Türkçe sözlüksel veri tabanıdır. Ancak Türkçe WordNet'in diğer dillere göre bazı dezavantajları bulunmaktadır. Örneğin, İngilizce WordNet kapsamında 117K anlamdaş terim olduğu düşünüldüğünde Türkçe WordNet oldukça düşük kapasiteye sahiptir. Aynı zamanda, Türkçe WordNet açık erişime

¹ Ahmet Afsin Akın, Zemberek-NLP; Erişim adresi; <https://github.com/ahmetaa/zemberek-nlp>

² İTÜ Türkçe Doğal Dil İşleme Yazılım Zinciri; Erişim adresi; <http://tools.nlp.itu.edu.tr/>

kapalıdır ve veritabanının güncellenmesine dair literatürde herhangi bir çalışmaya rastlanılmamıştır (Sağlam, 2019).

Duygu sözlükleri incelendiğinde dünyada SentiWordNet kullanımının yaygındır. Özellikle duygu analizi için tasarlanmış olup SentiWordNet, WordNet üzerine kurgulanmış duygu sözlüğüdür. WordNet'te var olan anlamsal ilişkilerin yanı sıra, kelimelere ait 3 kategoride (pozitif/negatif/nötr) duygu puanını içermektedir. Türkçe metinlerin analizi için düzenlenmiş farklı açık kaynaklı duygu sözlükleri bulunmaktadır. Bu bağlamda, incelenen literatürde sıklıkla başvurulan genel duygu sözlüğünün SentiTurkNet (Dehkharghani vd., 2016) olduğu görülmüştür. 3 farklı İngilizce kütüphane kullanılarak hazırlanan SentiTurkNet'te, SenticNet sözlüğü Türkçeye çevrilmiş, Türkçe ve İngilizce WordNet kütüphaneleri ile eşleştirilmiş ve 15K terim için İngilizce SentiWordNet'ten duygu puanları elde edilmiştir. Erşahin vd. (2019) çalışmasında ise SentiTurkNet kütüphanesini Türk Dil Kurumuna ait ASDICT kullanarak genişletilmiş ve sözlüğün kapsamını %78 oranında arttırılmıştır. Uçan (2014) ise İngilizce SentiWordNet veritabanını Türkçeye çevirerek 27K boyutlu bir duygu sözlüğü oluşturmuş, Sağlam (2019) ise bu sözlüğü temel alarak 37K boyutunda daha kapsamlı bir duygu sözlüğü geliştirmiştir. Her ne kadar başarılı Türkçe genel duygu sözlüğü oluşturma çalışmaları olsa da, mevcut sözlüklerin boyutu dünya örnekleri ile karşılaştırıldığında, bu alanda gelişime ihtiyaç duyulmaktadır.

4.3. Makine Öğrenme Paketleri

İncelenen literatürde, özellikle makine öğrenme yaklaşımını kullanan çalışmalarda çoğunlukla değinilmese de farklı programlama dillerinde kullanılan makine öğrenme paketleri olduğu göze çarpmaktadır. Bu anlamda yazında sıklıkla Python programlama dili kullanıldığı ve makine öğrenme yaklaşımlarında Python ile geliştirilmiş olan açık kaynaklı Scikit-Learn³ modülünden faydalandığı görülmüştür. Scikit-Learn, 2007 yılında başlayan bir proje olmasına rağmen günümüzde giderek önem kazanan bir makine öğrenme aracı olarak karşımıza çıkmaktadır. Duygu analizi çalışmaları açısından ise, makine öğrenme ile sınıflandırma çalışmalarında sıklıkla kullanılan Naive Bayes, destek vektör makinesi ve maksimum entropi gibi yöntemleri Scikit-Learn paketi kullanarak uygulamak mümkündür.

Diğer sıklıkla kullanılan programlama dili olan R ile Python'a benzer şekilde makine öğrenme yaklaşımlarını uygulamaya imkan sağlayan paketler bulunmaktadır. Bu kapsamda, kullanımı en

³ Scikit Learn: Machine Learning in Python: <https://github.com/scikit-learn/scikit-learn> veya <https://scikit-learn.org>

yaygın paketin MLR3⁴ olduğu görülmüştür. MLR paketinin güncellenmiş versiyonu olarak kullanımda olan MLR3 ile duygu sınıflandırmasında sıklıkla kullanılan yöntemleri uygulamak mümkündür.

4.4. Açık Kaynaklı Veri Tabanları

İncelenen literatürde kullanılan veri setleri incelendiğinde sıklıkla başvuru alan birkaç farklı kaynak göze çarpmaktadır. Twitter'ın kullanıcılarına geliştirici API hizmeti aracılığıyla hızlı ve kolayca metin tarama imkanı sağlaması sebebiyle, farklı konulardaki tweet gönderilerinin duygu analizi sıklıkla başvuru alan bir yöntemdir. Ancak bu şekilde derlenen verinin etiketsiz olması, duygu analizinde, özellikle denetimli makine öğrenimi yaklaşımlarda, bazı zorlukları beraberinde getirmektedir.

Özellikle denetimli makine öğrenme yöntemleri uygulanan çalışmalar için hazırlanmış olan, açık kaynaklı ve etiketli bir veri seti Hacettepe Üniversitesi MIRL (*Multimedia Information Laboratory*) tarafından sağlanmaktadır. Bu veri seti iki farklı kaynaktan toplanmış negatif ve pozitif etikete sahip 350K adet film ve otel yorumlarını içermektedir. Veri setinde yer alan yorumları beyazperde.com sitesinden, otel yorumları ise otelpuan.com sitesinden elde edilmiştir.

Türkçe duygu analizi literatüründe kullanılan, kullanıcılara ait ürün yorumlarının yer aldığı etiketli veri setleri de bulunmaktadır. Literatürde kullanılan bir diğer veri seti ise SEMEVAL 2016 veri setidir. SEMEVAL her yıl doğal dil işleme ve anlam analizinin farklı alanları üzerine düzenlenen atölyelerdir. 2016 yılında düzenlenen SEMEVAL atölyesinin ana konusu duygu analizidir ve bu senede düzenlenen atölyede kullanılan veri setinde Türkçe dilini de içermekte ve açık kaynaklı olarak paylaşılmaktadır.

5. Sonuç

Bu çalışmanın ile, Türkçe duygu analizi üzerine yapılan güncel çalışmalar ve bu çalışmalarda kullanılan yöntemler, özellikle sosyal bilimler alanındaki duygu analizi uygulamaları, kullanılan açık kaynaklı kütüphaneler ve veri tabanlarını ortaya konulmuştur. Araştırmaya konu olan yazın taraması için 43 adet yayın ve 12 adet tez (yüksek lisans ve doktora) incelenmiş ve duygu analizinin farklı aşamaları açısından ele alınarak irdelenmiştir.

Metinlerin ön işlenmesi sürecinde genellikle açık kaynaklı kütüphanelerin kullanımı oldukça yaygın olup, bu kütüphaneler arasında en sık kullanılan ve performansı en yüksek kütüphaneler

⁴ MLR3: Machine Learning in R: <https://mlr3.ml-org.com/>

Zemberek-NLP ve ITU NLP Turkish Parser'dir. Bir sonraki aşama olan özniteliklerin belirlenmesi üzerine yapılan çalışmalarda, Türkçe metinlerde ikili veya üçlü kelime gruplarının, başka bir deyişle n-gram yaklaşımlarının analiz performansını iyileştirdiği gösterilmiştir. Başka bir dikkate alınması gereken durum ise, noktalama işaretlerinin ve etkisiz kelimelerin nitelik olarak kalmasının analiz performansı üzerinde olumlu etkiye sahip olmasıdır. Öznitelik seçiminde ise, Ki-Kare (*Chi-Square*), Bilgi Kazanımı (*Information Gain*) gibi istatistiksel modellerin oldukça yaygın olduğu görülmüştür. Bu modellerin anlaşılabilirliğinin ve uygulanabilirliğinin kolay olması sebebiyle daha çok tercih edildiği düşünülmektedir. Bunun yanı sıra, özellikle sosyal medyada kinayeli ve ironi içeren paylaşımların oldukça fazla olması, bu alanda yapılacak analizlerde ironi tespiti analiz performansını etkileyecektir ve ironi tespitine dair kuralların belirlenerek öznitelik seçimine dahil edilmesi, analiz performansını arttıracak bir unsurdur. İncelenen literatürde ironi tespitine dair çalışmaların az olduğu ve kapsamı daha geniş çalışmalara ihtiyaç duyulduğu gözlenmiştir. Bunun yanı sıra, farklı dillerde yapılmış olan ironi tespiti yöntemleri ve bu yöntemlerin analiz performansına etkisi odağında yapılacak bir sistematik literatür taraması, Türkçe metinlerde ironi tespitine dair geliştirilmesi gereken yöntemlere ışık tutacaktır.

İncelenen literatür kullanılan duygu kategorileştirme yöntemleri açısından ele alındığında, hem analiz performansının daha yüksek olması hem de kolay uygulanabilir kütüphanelere sahip olması sebebiyle, özellikle makine öğrenme yaklaşımlarının uygulandığı çalışmaların daha yaygın olduğu görülmüştür. Denetimsiz makine öğrenme ve sözlük tabanlı yaklaşımlarının Türkçe analizlerdeki performansının incelenmesi başka bir çalışma alanı olarak karşımıza çıkmaktadır. Sözlük tabanlı yaklaşımlarda ise kullanılan genel ve duygu sözlüklerinin Dünyada yaygın olarak kullanılan WordNet ve SentiWordnet kadar geniş kapsamlı olmadığı ve Türkçe sözlük veri tabanlarının boyutunun dünya örneklerine kıyasla oldukça küçük olduğu gözlenmiştir. Bu bağlamda, Türkçe duygu analizinde kullanılan duygu sözlükleri, geliştirilmesi gereken bir alandır.

Diğer dillerde, özellikle İngilizce metinlerde, makine öğrenme ve sözlük tabanlı yöntemleri bir arada kullanan melez yaklaşımların analiz performansının oldukça yüksek olduğu bilinmektedir. İncelenen literatürde, Türkçe metinlerde melez yaklaşımların kullanıldığı çalışmaların çok fazla olmadığı gözlemlenmiştir ve bu yaklaşımlar ileri çalışma alanı olarak görülebilir.

KAYNAKÇA

- Açıkalin, Utku Umur, Benan Bardak, Mucahid Kutlu. 2020. Turkish Sentiment Analysis Using BERT. *28th Signal Processing and Communications Applications Conference (SIU)*. IEEE: 1-4.
- Aguwa, C., M.H. Olya ve L. Monplaisir. 2017. Modeling of fuzzy-based voice of customer for business decision analytics. *Knowledge-Based Systems*: 125, 136-145.
- Ahmetoğlu, Hüseyin, and Resul Daş. 2020. "Türkçe Otel Yorumlarıyla Eğitilen Kelime Vektörü Modellerinin Duygu Analizi İle İncelenmesi." *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, August, 455–63.
- Akba, Firat, Alaettin Uçan, Ebru Akcapinar Sezer, and Hayri Sever. 2014. "Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews". In 8th European Conference on Data Mining. 191(2002): 180-184.
- Akba, F. Assesment of Feature Selection Metrics for Sentiment Analysis: Turkish Movie Reviews, Published Master's Thesis, Ankara, (2014).
- Akgül, Eyüp Sercan, Caner Ertano, and Banu Diri. 2016. "Sentiment analysis with Twitter." *Pamukkale University Journal of Engineering Sciences* 22 (2): 106–10.
- Akın, Ahmet Afsin, ve Mehmet Dündar Akın. 2007. "Zemberek, an open source nlp framework for turkic languages." *Structure* 10: 1-5.
- Alpkoçak, Adil, Mansur Alp Tocoglu, Azer Çelikten, and İrfan Aygün. 2019. "Türkçe Metinlerde Duygu Analizi İçin Farklı Makine Öğrenmesi Yöntemlerinin Karşılaştırılması." *Deu Muhendislik Fakültesi Fen ve Muhendislik* 21 (63): 719–25.
- Arazy, O. ve Woo, C. 2007. Enhancing information retrieval through statistical natural language processing: a study of collocation indexing, *MIS Quarterly*, 31(3): 525-546.
- Atan, Suat, ve Yetkin Çınar. 2019. Borsa İstanbul'da finansal haberler ile piyasa değeri ilişkisinin metin madenciliği ve duygu (sentiment) analizi ile incelenmesi. *Ankara Üniversitesi SBF Dergisi*, 74(1): 1-34.
- Baker, M., ve J. Wurgler. 2007. Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2), 129-152.
- Bayraktar, Kivanc, Uraz Yavanoglu, and Alper Ozbilen. 2019. "A Rule-Based Holistic Approach for Turkish Aspect-Based Sentiment Analysis." *2019 IEEE International Conference on Big Data (Big Data)*. <https://doi.org/10.1109/bigdata47090.2019.9005473>.
- Birmingham, A., ve Alan F. S. 2010. Classifying sentiment in microblogs: is brevity an advantage?. In Proceedings of the 19th ACM international conference on Information and knowledge management, 1833-1836.
- Bilgin, Metin ve İzzet Fatih Şentürk. 2017. Sentiment analysis on Twitter data with semi-supervised Doc2Vec. In 2017 international conference on computer science and engineering (UBMK).IEEE: 661-666.
- Burcu, Akın, ve Şimşek, Umma, Tuğba, Gürsoy. 2018. Sosyal Medya Analitiği İle Değer Yaratma: Duygu Analizi İle Geleceğe Yönelim. *Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 5(3): 797-811.
- Catal, Cagatay, and Mehmet Nangir. 2017. "A Sentiment Classification Model Based on Multiple Classifiers." *Applied Soft Computing* 50 (January): 135–41.
- Ceron, A., L. Curini, S. Iacus ve G. Porro. 2014. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New media & society*, 16(2): 340-358.
- Ciftci, Basri, and Mehmet Serkan Apaydin. 2018. "A Deep Learning Approach to Sentiment Analysis in Turkish." *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*. <https://doi.org/10.1109/idap.2018.8620751>.
- Çalı, Sedef, and Şebnem Yılmaz Balaman. 2019. "Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment." *Computers & Industrial Engineering* 129: 315-332.
- Catal, Cagatay, and Mehmet Nangir. 2017. "A sentiment classification model based on multiple classifiers." *Applied Soft Computing* 50: 135-141.
- Çetin, Fatih Samet, and Gülşen Eryiğit. 2018. "Türkçe Hedef Tabanlı Duygu Analizi İçin Alt Görevlerin İncelenmesi – Hedef Terim, Hedef Kategori Ve Duygu Sınıfı Belirleme." *Bilişim Teknolojileri Dergisi*, January. <https://doi.org/10.17671/gazibtd.325865>.
- Çoban, Önder, Baris Özyer, and Gülsah Tümöklü Özyer. 2015. "Türkçe Twitter Mesajlarının Duygu Analizi Sentiment Analysis for Turkish Twitter Feeds." In 2015 23rd Signal Processing and Communications Applications Conference (SIU). IEEE. 2388-2391.

- Dave, K., Lawrence, S. ve Pennock, D.M. 2003, Mining the peanut gallery: opinion extraction and semantic classification of product reviews, Proceedings of the 12th International World Wide Web, ACM, Budapeşte.
- Dehkharghani, Rahim, Berrin Yanikoglu, Yucel Saygin, and Kemal Oflazer. 2017. "Sentiment Analysis in Turkish at Different Granularity Levels." *Natural Language Engineering* 23 (4): 535–59.
- Dehkharghani, Rahim, Yucel Saygin, Berrin Yanikoglu, and Kemal Oflazer. 2016. "SentiTurkNet: a Turkish polarity lexicon for sentiment analysis." *Language Resources and Evaluation* 50, no. 3: 667-685.
- Dehkharghani, Rahim. 2018. "A Hybrid Approach to Generating Adjective Polarity Lexicon and Its Application to Turkish Sentiment Analysis." *International Journal of Modern Education & Computer Science*, 10(11). <https://doi.org/10.5815/ijmecs.201.11.0>.
- Dehkharghani, Rahim, Sentiment Analysis in Turkish: Resources and Techniques, Doktora Tezi, Sabancı University, İstanbul (2015).
- Demirci, Gozde Merve, Seref Recep Keskin, and Gulustan Dogan. 2019. "Sentiment Analysis in Turkish with Deep Learning." *2019 IEEE International Conference on Big Data (Big Data)*. <https://doi.org/10.1109/bigdata47090.2019.9006066>.
- Dülger, Oğuzhan. 2018. "Türkçe Metinlerde İroni Tespiti." Proceedings of the 12th Turkish National Software Engineering Symposium (UYMS).
- Erşahin, Buket, Özlem Aktaş, Deniz Kiliç, and Mustafa Erşahin. 2019. "A Hybrid Sentiment Analysis Method for Turkish." *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES* 27 (3): 1780–93.
- Fan, Z. P., Che, Y. J., ve Chen, Z. Y. 2017. Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research* 74: 90-100.
- Fariás, Delia Irazú Hernández, Viviana Patti, and Paolo Rosso. "Irony detection in twitter: The role of affective content." *ACM Transactions on Internet Technology (TOIT)* 16, no. 3 (2016): 1-24.
- Fariás, Irazú Hernández, José-Miguel Benedí, and Paolo Rosso. "Applying basic features from sentiment analysis for automatic irony detection." In *Iberian Conference on Pattern Recognition and Image Analysis*, pp. 337-344. Springer, Cham, 2015.
- Gözükara, Furkan, and Selma Ayşe Özel. 2016. "An Experimental Investigation of Document Vector Computation Methods for Sentiment Analysis of Turkish and English Reviews."
- Hine, C. 2005, Virtual Methods: Issues in Social Research on the Internet, Berg, Oxford.
- İskender, Elyase, and Gülgönül Bozoğlu Batı. 2015. "Comparing Turkish Universities Entrepreneurship and Innovativeness Index's Rankings with Sentiment Analysis Results on Social Media." *Procedia - Social and Behavioral Sciences* 195 (July): 1543–52.
- Kama, Batuhan, Murat Ozturk, Pinar Karagoz, I. Hakki Toroslu, and Murat Kalender. 2017. "Analyzing Implicit Aspects and Aspect Dependent Sentiment Polarity for Aspect-Based Sentiment Analysis on Informal Turkish Texts." In *Proceedings of the 9th International Conference on Management of Digital EcoSystems*. New York, NY, USA: ACM. <https://doi.org/10.1145/3167020.3167041>.
- Kamisli Ozturk, Z., Z. İ. Erzurum Cicek, and Z. Ergul. 2017. "Sentiment Analysis: An Application to Anadolu University." *Acta Physica Polonica Series a* 132 (3): 753–55.
- Karagoz, Pinar, Batuhan Kama, Murat Ozturk, I. Hakki Toroslu, and Deniz Canturk. 2019. "A Framework for Aspect Based Sentiment Analysis on Turkish Informal Texts." *Journal of Intelligent Information Systems* 53 (3): 431–51.
- Karamollaoğlu, Hamdullah, İbrahim Alper Doğru, Murat Dörterler, Anıl Utku, Oktay Yıldız. 2018. Sentiment analysis on Turkish social media shares through lexicon based approach. In *2018 3rd International Conference on Computer Science and Engineering*. IEEE. 45-49
- Karaöz, Burcu, and U. Tuğba Gürsoy. 2018. "Adaptif Öğrenme Sözlüğü Temelli Duygu Analiz Algoritması Önerisi." *Bilişim Teknolojileri Dergisi*, August. <https://doi.org/10.17671/gazibtd.342419>.
- Karcioğlu, Abdullah Ammar ve Tolga Aydın. 2019. Sentiment analysis of Turkish and english twitter feeds using Word2Vec model. In *2019 27th Signal Processing and Communications Applications Conference (SIU)*. IEEE: 1-4
- Kaya, M., G. Fidan, and I. H. Toroslu. 2012. "Sentiment Analysis of Turkish Political News." In *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, 1:174–80.
- Kaynar, Oguz, Yasin Görmez, Mustafa Yıldız, and Ayşegül Albayrak. 2016. "Makine Öğrenmesi Yöntemleri İle Duygu Analizi." In *International Artificial Intelligence and Data Processing Symposium (IDAP'16)*: 17-18.
- Keleş, Pervan Nergis; 2017. "Sentiment Analysis Using a Random Forest Classifier on Turkish Web Comments." *Communications Faculty Of Science University of Ankara*. https://doi.org/10.1501/commua1-2_0000000105.

- Kozinets, R. 2009, Netnography: Doing Ethnographic Research Online, Sage Publications, Londra.
- Ku, L.W., Ho, H.W. and Chen, H.H. 2009. Opinion mining and relationship discovery using CopeOpi opinion analysis system, *Journal of the American Society for Information Science and Technology*, 60 (7): 1486-1503.
- Kumar, A., ve A. Jaiswal. 2020. Systematic literature review of sentiment analysis on Twitter using soft computing techniques. *Concurrency and Computation: Practice and Experience*, 32(1): e5107.
- Liu, Bing. 2012. "Sentiment analysis and opinion mining." *Synthesis lectures on human language technologies* 5, no. 1: 1-167.
- Medhat, Walaah, Ahmed Hassan, and Hoda Korashy. "Sentiment analysis algorithms and applications: A survey." *Ain Shams engineering journal* 5, no. 4 (2014): 1093-1113.
- Metin, Senem Kumova, and Bahar Karaođlan. "Türkiye Türkçesinde Eşdizimlerin İstatistiksel Yöntemlerle Belirlenmesi." *Bilig* 78 (2016): 253-286.
- Misopoulos, F., Mitic, M., Kapoulas, A. and Karapiperis, C. 2014. "Uncovering customer service experiences with Twitter: the case of airline industry", *Management Decision*, 52 (4): 705-723.
- Nalçakan, Yağiz, Şan Sitki Bayramođlu, and Samed Tuna. 2015. "Sosyal Medya Verileri Üzerinde Yapay Öğrenme İle Duygu Analizi Çalışması." Technical Report. Trakya Üniversitesi
- Nielsen, Finn Årup. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs." arXiv preprint arXiv:1103.2903 (2011).
- Nizam, Hatice, and Saliha Sila Akin. 2014. "Sosyal Medyada Makine Öğrenmesi İle Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması." IX. Türkiye'de İnternet Konferansı, 1-6.
- Nasukawa, T. ve Yi, J. 2003. Sentiment analysis: capturing favorability using natural language processing, K-CAP 2003 Proceedings of the 2nd International Conference on Knowledge Capture, Sanibel Island, FL, 70-77.
- O'Connor, B., R. Balasubramanyan, B. R. Routledge ve N. A. Smith. 2010. From tweets to polls: Linking text sentiment to public opinion time series. *Tepper School of Business*, 559.
- O'leary, D. E. 2011. The use of social media in the supply chain: Survey and extensions. *Intelligent Systems in Accounting, Finance and Management*, 18(2-3): 121-144.
- Özsert, Cüneyd Murad, and Arzucan Özgür. "Word polarity detection using a multilingual approach." In *International Conference on Intelligent Text Processing and Computational Linguistics*, pp. 75-82. Springer, Berlin, Heidelberg, 2013.
- Öztürk, Nazan, and Serkan Ayvaz. 2018. "Sentiment Analysis on Twitter: A Text Mining Approach to the Syrian Refugee Crisis." *Telematics and Informatics* 35 (1): 136-47.
- Park, K., ve Ha, S. H. 2018. Mining user-generated opinions online with LDA model to discover service complaints. *International Information Institute (Tokyo). Information*, 21(3): 875-884.
- Parlar, Tuba, and Selma Ayse Ozel. 2016. "A New Feature Selection Method for Sentiment Analysis of Turkish Reviews." *2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*. <https://doi.org/10.1109/inista.2016.7571833>.
- Parlar, Tuba, Selma Ayşe Özel, and Fei Song. 2018. "QER: A New Feature Selection Method for Sentiment Analysis." *Human-Centric Computing and Information Sciences* 8 (1). <https://doi.org/10.1186/s13673-018-0135-8>.
- Sağlam, F. Otomatik Duygu Sözlüğü Geliştirilmesi ve Haberlerin Duygu Analizi. Doktora Tezi, Hacettepe Üniversitesi, Ankara (2019).
- Santur, Yunus. 2019. "Sentiment Analysis Based on Gated Recurrent Unit." *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*. <https://doi.org/10.1109/idap.2019.8875985>.
- Seyfiođlu, Mehmet, and Mustafa Demirezen. 2017. "A Hierarchical Approach for Sentiment Analysis and Categorization of Turkish Written Customer Relationship Management Data." In *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems*. IEEE. <https://doi.org/10.15439/2017f204>.
- Shehu, H. A., and S. Tokat. 2020. "A Hybrid Approach for the Sentiment Analysis of Turkish Twitter Data." *Artificial Intelligence and Applied Mathematics in Engineering Problems*. https://doi.org/10.1007/978-3-030-36178-5_15.
- Shehu, Harisu Abdullahi, Sezai Tokat, Md Haidar Sharif, and Sahin Uyaver. 2019. "Sentiment Analysis of Turkish Twitter Data." *AIP Conference Proceedings* 2183 (1): 080004.
- Taşhođlu, H.: Irony Detection On Turkish Microblog Texts, Yüksek Lisans Tezi, Orta-dođu Teknik Üniversitesi, Bilgisayar Mühendisliđi, Ankara (2014).
- Thelwall, Mike, Kevan Buckley, and Georgios Paltoglou. 2011. "Sentiment in Twitter events." *Journal of the American Society for Information Science and Technology* 62(2): 406-418.
- Thelwall, M., Kevan Buckley, Georgios Paltoglou, D. Cai ve A. Kappas. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558.

- Tricco, Andrea C., Erin Lillie, Wasifa Zarin, Kelly O'Brien, Heather Colquhoun, Monika Kastner, Danielle Levac et al. "A scoping review on the conduct and reporting of scoping reviews." *BMC medical research methodology* 16, no. 1 (2016): 1-10.
- Turkish WordNet: Ehsani R, Solak E, Yildiz OT. Constructing a WordNet for Turkish using manual and automatic annotation. *ACM Transactions on Asian Language Information Processing* 2018; 17 (3): 1-15. doi: 10.1145/3185664
- Turney P. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In: *Proceedings of annual meeting of the Association for Computational Linguistics (ACL'02)*; 2002
- Türkmenoğlu, C. Türkçe Metinlerde Duygu Analizi, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul (2015).
- Türkmenoğlu, Cumali, and Ahmet Cüneyd Tantug. "Sentiment analysis in Turkish media." In *International Conference on Machine Learning (ICML)*. 2014.
- Ucan A, Naderalvojud B, Sezer EA, Sever H. SentiWordNet for new language: automatic translation approach. In: *12th International Conference on Signal-Image Technology & Internet-Based Systems*; Naples, Italy; 2016. pp. 308-315.
- Ucan, A. "Automatic sentiment dictionary translation and using in sentiment analysis." *Yükseklisans Tezi, Hacettepe Üniversitesi, Ankara, Turkey* (2014).
- Uslu, Abdullah, Sefa Tekin, and Tefvik Aytakin. 2019. Sentiment analysis in Turkish film comments. In *2019 27th Signal Processing and Communications Applications Conference (SIU)*. IEEE:1-4.
- Wang, H., D. Can, A. Kazemzadeh, F. Bar ve S. Narayanan. 2012. A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 system demonstrations*, 115-120.
- Wright, A. 2009. Our sentiments, exactly, *Communications of the ACM*, 52 (4): 14-15.
- Velioglu, Riza, Tugba Yildiz, and Savas Yildirim. 2018. "Sentiment Analysis Using Learning Approaches Over Emojis for Turkish Tweets." *2018 3rd International Conference on Computer Science and Engineering (UBMK)*. <https://doi.org/10.1109/ubmk.2018.8566260>.
- Vural, A. Gural, B. Barla Cambazoglu, Pinar Senkul, and Z. Ozge Tokgoz. "A framework for sentiment analysis in turkish: Application to polarity detection of movie reviews in turkish." In *Computer and Information Sciences III*, pp. 437-445. Springer, London, 2013.
- Yurtalan, Gökhan, Murat Koyuncu, and Çiğdem Turhan. 2019. "A Polarity Calculation Approach for Lexicon-Based Turkish Sentiment Analysis." *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*, March, 1325-39.