İSTANBUL UNIVERSITY PRESS

Chapter 16

Sentiment Analysis Using Data Obtained from Twitter

Sedanur Yeşilkaya Koç [1], Derya Çakıcı [2] & Murat Gezer [3]

[1] İstanbul University, Institute of Graduate Studies in Sciences, Informatics Program, İstanbul, Türkiye
[2] Garanti BBVA, İstanbul, Türkiye
[3] İstanbul University, Faculty of Science Department of Computer Science, İstanbul, Türkiye

Abstract

The widespread use of internet technology has caused social networks to have a great place in human life. Individuals who shared about their lives and thoughts provided researchers with a great deal of data for their studies on various subjects. It has been observed that Twitter is widespread, especially since it is text-based in sharing thoughts. The increasing number of shared tweets has caused researchers to encounter a huge data crowd. Since it is not possible to manually analyze or classify this large-scale data, it has become a necessity to analyze and classify it by a system. For this reason, it has been inevitable for researchers to turn to the fields of text mining and sentiment analysis. In this article, the subject of text mining and sentiment analysis is examined together with its details and approaches and presented as a literature review. In the application part, a sentiment analysis study was conducted with the data taken from Twitter on a determined subject.

Keywords Sentiment analysis · twitter · machine learning

Citation: Yeşilkaya Koç, S., Çakıcı, D. & Gezer, M. (2025). Sentiment analysis using data obtained from Twitter. in S. Çelik, S. Gülseçen & M. Eryılmaz (Eds.), *Creating A Human-Focused Future* (p. 172-183). İstanbul University Press. https://doi.org/10.26650/B/T3T18.2025.013.016

This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License.

© 2025. Yeşilkaya Koç, S., Çakıcı, D. & Gezer, M.

Corresponding author: Sedanur Yeşilkaya Koç sedayesilkaya@gmail.com

Introduction

The rise of social media, driven by the dominance of Web 2.0 technology in the internet world, continues to increase in popularity. Individuals who dedicate substantial time to social media in their daily lives have become open to both the positive and negative aspects of the evolving digital world (Buzzi et al., 2011; Pak & Paroubek, 2010). Social networks, particularly Twitter, have transformed societal structures and communication methods, influencing the way individuals think and behave in both personal and social matters. In the examination of social networks, it is observed that these platforms offer significant opportunities for researchers to observe, identify, and interpret the thinking patterns of the public (Fermanoğlu, 2020).

One of the most widely used social networks today is Twitter, which stands out in these examinations. Twitter's unique position stems from its ability to allow users to freely express their thoughts on various topics, providing an opportunity for both social researchers and producing companies to infer future preferences (Bermingham & Smeaton, 2011). With over 330 million users worldwide, Twitter has become a platform where over 8,000 tweets are shared per second, according to recent figures (Karabulut & Küçüksille, 2018; Murthy, 2018). Given the current number of tweets shared per second, the magnitude of the obtained data is particularly noteworthy.

Since its inception in 2006, Twitter has increasingly become an integral part of individuals' lives worldwide. Especially by offering publicly accessible data, Twitter contributes significantly to text mining researchers. Researchers can easily access content that has not been restricted by users, allowing them to complete their studies based on a defined problem and reach various results.

In research, Twitter has been recognized for having its own language, causing challenges in understanding the existing data due to reasons such as misspelled words. As a result, various studies related to natural language processing and data mining have been initiated, and industry products have been introduced using these methods (Ayan et al., 2019).

Two prominent methods in the process of processing Twitter data and making it understandable and analyzable are text mining and sentiment analysis. Text mining is utilized in making the obtained data understandable, while sentiment analysis is employed in the analysis of data and obtaining a result.

Twitter, by offering users the opportunity to express themselves with words rather than visual posts, has become a platform where ideas are shared in various fields, from politics to psychology, from movies to music albums, from customer experiences to business life, and even daily life problems. At this point, sentiment analysis plays a crucial role in determining subjective information such as emotions, opinions, and attitudes, enabling strategic decisions in marketing, predicting financial ratios, customer churn analysis, determining industry opportunities and threats, and pre-determining competitors' activities based on the data compiled with text mining (Onan, 2017).

Matrix of Literature Review for the Findings of the Study

A total of 20 articles within the scope of the study were examined by two researchers (graduate students). Studies on Twitter sentiment analysis were analyzed based on the approaches they employed. The examinations were organized in terms of authors, year of publication, dataset used, approach, algorithm, and results. The results are summarized in Table 1.

Table 1

Findings Reached Through Literature Review on Approaches to Twitter Sentiment Analysis

Yazar	*Yıl*	*Veri Seti*	*Yaklaşım*	*Algoritma*	*Sonuç*
Sarah E. Shukri, Rawan I. Yaghi, Ibrahim Aljarah, Hamad Alsawalgah	2015	Data from 1000 Twitter posts	Dictionary-based	Naive Bayes (NB)	The study's findings indicate that the percentage of joy is higher for BMW compared to Mercedes and Audi, while the percentage of sadness is higher for Mercedes and Audi when compared to BMW.
Prerna Mishra, Dr. Ranjana Rajnish, Dr.Pankaj Kumar	2016	Twitter data expressing opinions on the subject	Dictionary-based	SentiWord	The analysis results indicate that opinions and thoughts are observed to be 50% positive, 20% negative, and 30% neutral.
Laila M. Qaisi, Ibrahim Aljarah	2016	Twitter and cloud systems data	Dictionary-based	Naive bayes (NB)	At the conclusion of the study, it was observed that in the positive category of joy, Microsoft Azure is in a better condition compared to Amazon, and Amazon has a higher percentage of negativity.
Entesar M. Milod Eljali	2019	50,000 Twitter data obtained from 5 different Twitter accounts	Dictionary-based		This thesis study has provided a significant contribution to the literature by offering important information flow. Particularly, the development and introduction of a useful tool, based on new empirical methods, for identifying potential emotional reasons in the emotional flow of a specific user has been a key aspect of this contribution
Hatice Kübra Küçükkartal	2020	10,000 English Twitter data	Dictionary-based		As a result of sentiment analysis, the posts made by individuals on the subject were analyzed to determine whether they were positive, negative, or neutral. The findings reveal that approximately 42% of the tweets are positive, 14% are negative, and 43% are neutral.
Marco Pennacchiotti, Ana-Maria Popescu	2011	Twitter data expressing opinions on the subject	Machine learning-based	Gradient Boosted Decision Trees (GBDT)	The analysis results indicate that the study has largely achieved its purpose and obtained positive outcomes.

Hao Wang, Dogan Can, Abe Kazemzadeh, François Bar, Shrikanth Narayanan	2012	Simultaneous Twitter data expressing opinions on the subject	Machine learning-based	Naive bayes (NB)	As a result, a real-time system has been developed for the ongoing elections, where Twitter data is analyzed in real-time.
Geetika Gautam, Divakar yadav	2014	Twitter data.	Machine learning-based	Naive bayes(NB), maximum entropy(MaxEnt), support vector machines(SVM)	The study has contributed to the literature by conducting a comprehensive data analysis.
Hatice Nizam, Saliha Sıla Akın	2014	Two separate Twitter datasets created with a total of 2824 tweets	Machine learning-based	Naive Bayes (NB), Sequential Minimal Optimization (SMO), Decision Tree (J48), Random Forest (RF) and 1-Nearest Neighbors (IB1)	The experimental results revealed that the balanced dataset, exhibiting an appropriate distribution among the results, outperformed the imbalanced dataset. According to the analysis, the classification algorithm that demonstrated the best performance was SMO with an average accuracy rate of 72.33%.
Önder Çoban, Barış Özyer, Gülşah Tümüklü Özyer	2015	20,000 Twitter data, consisting of 10,000 positive and 10,000 negative entries	Machine learning-based	SVM, Naive Bayes, Multinom Naive Bayes and KNN	This study, based on the findings obtained, confirms the thesis that experimental results show Twitter messages can be classified using machine learning methods.
Yağız Nalçakan, Şan Sıtkı Bayramoğlu, Samed Tuna	2015	Three separate Twitter datasets comprising 1500 tweets each about Samsung, Apple, and LG	Machine learning-based	Naive Baye(NB), Random Forest(RF), LibSVM, J48 and KStar	When all the results are examined, it has been decided to work with a larger dataset to find a meaningful result. Nevertheless, evaluating the results obtained, it can be seen that in both studies, the Naive Bayes (NB) algorithm yielded the best results among the three companies.
Varsha Sahayak, Vijaya Shete, Apashabi Pathan	2015	Twitter data expressing opinions on the subject	Machine learning-based	Naive bayes(NB), maximum entropy(MaxEnt), support vector machines(SVM)	The result of the study is positive feedback from survey results. In other words, participants have concluded that the data from Twitter would be effective in sentiment analysis.

Tuba Parlar, Esra Saraç, Selma Ayşe Özel	2017	A total of 3000 Twitter data belonging to three different classes (positive, negative, neutral)	Machine learning-based	Maximum Entropy Modeling (MEM)	According to the results of the study, it has been observed that feature selection has significantly positive effects on classification performance. Reducing the number of features not only enhances classification performance but also sharply decreases the classification time. This indicates that feature selection is an available method for improving the overall performance of the system.
Aytuğ Onan	2017	A total of 10,600 Twitter data, comprising 5,300 positive and 5,300 negative entries	Machine learning-based	Naive Bayes classifier (NB), Support Vector Machines (SVM), and Logistic Regression (LR)	As a result of experimental studies, the highest performance (%77.78) is achieved when the feature set is represented by combining 1-gram and 2-gram feature sets, and the Naive Bayes algorithm is used as the classification algorithm.
Buğra Ayan, Birol Kuyumcu, Bünyamin Ciylan	2019	Twitter data obtained in the English language	Machine learning-based	Linear Ridge Regression and Naive Bayes	According to the analysis results, for positive tweets, better outcomes were achieved in the Ridge model compared to the Naive Bayes classifier. In Ridge Regression, a accuracy rate of 96.3% was reached, while in the Naive Bayes Classifier, the accuracy rate was 95.3%.
Cihan Çılgın, Ceyda Ünal, Serkan Alıcı, Ekin Akkol, Yılmaz Gökşen	2020	Twitter data expressing opinions on the subject	Machine learning-based	Artificial Neural Networks, Logistic Regression, Support Vector Machines, and Naive Bayes	The analysis revealed a correlation coefficient of 0.681, indicating a moderately strong positive relationship.
Haruna Isah, Paul Trundle, Daniel Neagu	2014	Twitter and Facebook data	Hybrid	Naive bayes (NB)	The analyses resulted in positive outcomes.
Eyüp Sercan Akgül, Caner Ertano, Banu Diri	2015	Three separate Twitter datasets with 500 tweets (VS1), 1200 tweets (VS2), and 5100 tweets (VS3)	Hybrid		The dictionary-based and character-based n-gram methods achieved success rates of approximately 70% and 69%, respectively. It was observed that the dictionary-

					based approach yielded more successful results.
Thakare Ketan Lalji, Sachin N. Deshmukh	2016	Twitter data expressing opinions on the subject	Hybrid	Support Vector Machine (SVM)	According to the results, the dictionary-based approach demonstrated high precision when used alone, but performance issues were observed. The researchers enhanced performance efficiency by combining it with a machine learning-based approach.
Harisu Abdullahi Shehu	2019	13,000 Twitter data	Hybrid	Random Forest and Support Vector Machines	The analysis results indicate that Support Vector Machines have the shortest execution time, the Random Forest method performs better on raw data, and the performance of the method using a polarity dictionary improves in a way not observed in other methods.

In reviewing the table, it is observed that the first study on sentiment analysis on Twitter in the literature was conducted by Pennacchiotti and Popescu (2011). Subsequently, one study continued in 2012, and between 2014-2015, 8 studies were conducted, followed by 5 studies between 2016-2017, and 4 studies between 2018-2020.

When looking at the datasets used in the examined studies, it is seen that in 18 studies only Twitter data was used, and in the remaining 2 studies, a dataset created by combining Twitter and Facebook or Twitter and cloud systems data was preferred. The tweets contain individuals' opinions and emotions on relevant topics in their entirety.

When the examined studies are analyzed based on the approaches they rely on, it is observed that 5 dictionary-based, 11 machine learning-based, and 4 hybrid approaches were preferred. From these results, it can be seen that machine learning-based approaches were predominantly used.

When evaluated according to the classification algorithms used, Naive Bayes (NB) was used in 12 studies, Support Vector Machine (SVM) in 8 studies, Maximum Entropy (MaxEnt) in 3 studies, Random Forest (RF) in 3 studies, Logistic Regression (LR) in 2 studies, Decision Trees (J48) in 2 studies, Gradient Boosted Decision Trees (GBDT), knn, kstar, Sequential Minimal Optimization (SMO), Linear Ridge Regression, 1-Nearest Neighbors (IB1), artificial neural networks, and Multinom Naive Bayes classifiers were used in 1 study each. As seen, Naive Bayes (NB) and Support Vector Machine (SVM) classification algorithms were predominantly preferred.

In general, the overall result of the studies indicates that sentiment analyses conducted on Twitter provide quite positive results in determining what individuals think about various topics, being effectively utilized.

Materials and Methods

The common feature of patriarchal societies in many regions of the world, including Turkey, is that the place of women in society shows similarities. Women, who have been kept in the background in patriarchal societies from the past to the present, have started to actively participate in all areas of life and live freely in every sense with the integration of civilization into life and modernization. During this transition process, the issue of violence against women, a common feature in patriarchal societies around the world, has become more audible, especially with the development of technology and social media, and the reactions given at this point have become visible to everyone.

The main idea of this study is to evaluate tweets about "violence against women" during a specified period on Twitter, one of the largest social media platforms where people express their thoughts verbally, and to classify these ideas as positive/negative using sentiment analysis method.

Methodology

Under this heading, the structure and quantity of the data used in the study, along with the data collection and cleaning stages, have been discussed, with a final emphasis on the sentiment analysis section. A dictionary-based approach was employed for the analysis.

Dataset

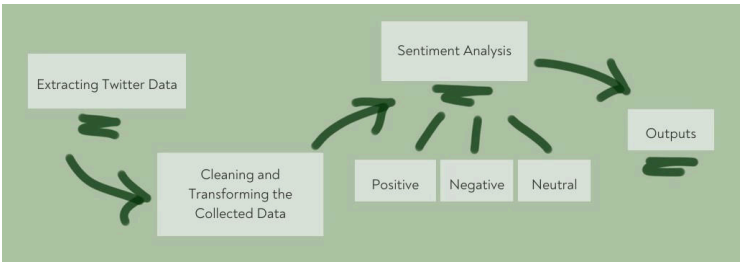
Python programming language was chosen to collect data from Twitter. Tweets containing the keyword "kadinasiddet" (violence against women) were collected. For these operations, the Selenium module was used, which allows the automation of browser actions, such as opening a browser like Chrome or Firefox with driver support installed on the computer, to simulate human-like interactions through programming language. In the context of the application, 819 tweets posted between 2010 and 2019 were accessed using the Firefox browser.

Data Preprocessing Steps

Twitter denotes user names with the "@" character and hashtags with the "#" character. In the process of cleaning the data, these characters are initially removed from the tweets. It is essential for the tweets to undergo this preprocessing step before the classification process. In this study, this process was carried out using the NLTK library in Python. Starting with these characters, the top 30 stop words determined for Turkish were removed to facilitate the analysis stage. In summary, initially, all uppercase letters in the collected tweets were converted to lowercase. Then, non-alphabetic characters were removed from the tweets. Subsequently, user names, hashtags, and URL addresses in the tweets were cleaned, and finally, stop words were removed, leading to the analysis stage.

Figure 1

Sentiment Analysis



In sentiment analysis, dictionaries such as HowNet (Fu et al., 2017), SentiWordNet (Montejo-Ráez et al., 2014), and MPQA (Saif et al., 2016) are used in various regions of the world. In this study, SentiTurkNet was employed as the dataset for Turkish sentiment analysis. SentiTurkNet contains 15,000 words and word sets, each with three separate polarity scores for positive, negative, and neutral sentiments. In the study, the words in each tweet collected from Twitter were compared with the words in the SentiTurkNet dataset to determine the polarity and polarity score of each word. Following this stage, as shown in Figure 1, each tweet was analyzed based on the collected polarity score and the negative/positive/neutral status. The score values were determined according to the number of positive and negative words in the tweets.

Conclusion

In this study, research on text mining and sentiment analysis was examined, and the results were evaluated based on the approaches used. The concepts of text mining and sentiment analysis were explained in detail in the study, presented in a language that can be understood by anyone, including readers and researchers who are new to the subject. Additionally, three methods used in sentiment analysis were explained.

When examining research on machine learning-based approaches, it was observed that generally two or three datasets were used. Typically, these datasets are separated into training and testing datasets. When machine learning-based approaches were used in the research, it was observed that more reliable results were obtained, and the approach allowed predictions for the future, assisting in making strategic decisions related to the subject. In research comparing classification algorithms used in this approach, Naive Bayes algorithms were found to provide more successful results.

In the dictionary-based approach, researchers have two options for the database. When a ready-made database is selected, the sentiment scores valid for that database are taken into account. Another option for researchers is to create their own corpus. In corpora, words and sentiment scores are determined by researchers. In the articles examined on the dictionary-based approach, it is mostly observed that a ready-made database is preferred. This approach is generally used to uncover information about the sentiments expressed by words used in the study related to the examined subject, whether they are positive, negative, or neutral. In articles, the dictionary-based approach is often seen to prefer weighting with TF-IDF (Term Frequency - Inverse Document Frequency). Research results provide positive, negative, and neutral ratios.

In the hybrid approach, both the dictionary-based and machine learning-based approaches are used. The machine learning-based approach is included when the dictionary-based approach is used and is insufficient. In research using this approach, it was observed that the dictionary-based approach yielded more successful results when compared to the machine learning-based approach. In another study, it was observed that the dictionary-based approach showed high accuracy when used alone but faced performance issues, and when used in conjunction with the machine learning-based approach, it improved performance. These two cases show that the successful approach in the hybrid approach depends on the research's goal.

When all the articles examined in this study are evaluated, it is seen that each of the three approaches can be used depending on the research's goal. The desired results were achieved in sentiment analyses of data collected from Twitter. It was clearly observed that researchers could use sentiment analysis in solving problems related to their areas of interest. Additionally, it was revealed that sentiment analysis would shed light on various research topics in the future, such as social events, customer analyses for companies, and the examination of individual personality traits.

In the final section of the study, tweets containing 'kadinasiddet' (violence against women) posted on Twitter within a specific date range were examined using the data collected from Twitter. The dataset obtained by extracting data from Twitter was evaluated through sentiment analysis, and scores were determined. According to the sentiment analysis, out of 819 tweets posted between 2010 and 2020, 71 were scored positively, 145 received negative scores, and 341 were neutral. The remaining tweets were not used after preprocessing.

In conclusion, with the data obtained in this study, it is possible to examine when society expresses its views on 'violence against women' more in certain periods through social media, when more negative results are obtained, and when the data follows a normal course. As a more comprehensive study, starting a new study with high processing power and conducting research on visual and video content tweets using image processing libraries can be considered. Finally, by adding the location factor, tweets can be listed, and tweets about 'violence against women' for different locations and countries can be examined.

Author Details

Sedanur Yeşilkaya Koç

[1] İstanbul University, Institute of Graduate Studies in Sciences, Informatics Program, İstanbul, Türkiye

0009-0004-3260-5597 sedayesilkaya@gmail.com

Derya Çakıcı

[2] Garanti BBVA, İstanbul, Türkiye

0009-0003-1056-2277 deryacakici16@gmail.com

Murat Gezer

[3] İstanbul University, Faculty of Science Department of Computer Science, İstanbul, Türkiye

0000-0002-7286-3943 murat.gezer@istanbul.edu.tr

References

Amolik, A., Jivane, N., Bhandari, M., & Venkatesan, M. (2015). Twitter Sentiment Analysis Of Movie Reviews Using Machine Learning Techniques. *International Journal of Engineering and Technology*. *7*(6), 2038–2044.

Akgül, E. S., Ertano, C., & Diri, B. (2016). Twitter Verileri İle Duygu Analizi. *Pamukkale University Journal of Engineering Sciences*, *22*(2).

Ayan, B., Can, M., & Gürsoy, U. T. (2019). Sosyal Medya Madenciliği İle Firmaların Twıtter Verilerinin İncelenmesi. *International Journal of Economic & Administrative Studies*, (23).

Aydın, G. (2017). *No-Sql Eri Tabanları Üzerinde Bir Metin Madenciliği Uygulaması* [Doctoral Dissertation, İstanbul Aydın Üniversitesi Fen Bilimleri Enstitüsü].

Aytuğ, Onan. (2017). Twitter Mesajları Üzerınde Makine Öğrenmesi Yöntemlerine Dayalı Duygu Analizi. *Yönetim Bilişim Sistemleri Dergisi*, *3*(2), 1-14.

Bermingham, A., & Smeaton, A. (2011, November). On Using Twitter To Monitor Political Sentiment And Predict Election Results. In *Proceedings of the Workshop On Sentiment Analysis Where AI Meets Psychology (Saaıp 2011)* (Pp. 2-10).

Bougie, R., Pieters, R., & Zeelenberg, M. (2003). Angry Customers Don't Come Back, They Get Back: The Experience And Behavioral Implications of Anger and Dissatisfaction in Services. *Journal of the Academy of Marketing Science*, *31*(4), 377-393.

Buzzi, M. C., Buzzi, M., & Leporini, B. (2011, September). Web 2.0: Twitter And The Blind. In *Proceedings of the 9th Acm Sıgchı Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity* (Pp. 151-156).

Çoban, Ö., Özyer, B., & Özyer, G. T. (2015, May). Sentiment Analysis for Turkish Twitter Feeds. In *2015 23nd Signal Processing and Communications Applications Conference (Sıu)* (Pp. 2388-2391). Ieee.

Çılgın, C., Ceyda, Ünal., Alıcı, S., Akkol, E., & Gökşen, Y. (2020). Metin Sınıflandırmada Yapay Sinir Ağları İle Bitcoin Fiyatları Ve Sosyal Medyadaki Beklentilerin Analizi. *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, *4*(1), 106-126.

Das, S., Chen, M. (2001, July). Yahoo! For Amazon: Extracting Market Sentiment From Stock Message Boards. In Proceedings of the Asia Pacific Finance Association Annual Conference (Apfa) (Vol. 35, P. 43).

Diri B. (2014). *Doküman Sınıflandırma Sunumu*. Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü.

Elmas, Ş. Ş. (2019). *Sosyal Medya Mesajlarının Veri Madenciliği Yöntemi İle Duygu Analizi (Sivas İli Örneği)* (Master's Thesis, Sivas Cumhuriyet Üniversitesi-Sosyal Bilimler Enstitüsü).

Eljalı, E. M. M. (2019). *Arapça Dilini Etkileyen Twitter Kullanıcılarının Duygu Etkenleri Arkasındaki Olası Sebepleri Tespit Etmek* [Doctoral Dissertation, Kastamonu Üniversitesi].

Ergün, K. (2012). *Metin Madenciliği Yöntemleri ile Ürün Yorumlarının Otomatik Değerlendirilmesi*. [Doktora Tezi, Sakarya Üniversitesi Fbe].

Erten, F. (2015). *Metin Madenciliği Tabanlı Bir Web Sitesi Sınıflandırma Aracı Tasarımı* [Master's Thesis, Maltepe Üniversitesi, Fen Bilimleri Enstitüsü].

Fermanoğlu, Ö. M. Sosyal Medya Ve Yerel Seçimler: Ak Parti# Gönülbelediyeciliği Etiketi Üzerine Bir Analiz. *Selçuk İletişim*, *13*(1), 44-70.

Fu, X., Liu, W., Xu, Y., & Cui, L. (2017). Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. *Neurocomputing, 241*, 18-27.

Gautam, G., & Yadav, D. (2014, August). Sentiment Analysis of Twitter Data Using Machine Learning Approaches and Semantic Analysis. In *2014 Seventh International Conference on Contemporary Computing (Ic3)* (Pp. 437-442). Ieee.

Güven, A. (2007). *Türkçe Belgelerin Anlam Tabanlı Yöntemlerle Madenciliği* [Doctoral Dissertation, Ytü Fen Bilimleri Enstitüsü].

http://Mis.Sadievrenseker.Com/2014/06/Metin-Madenciligi-Text-Mining/[Ziyaret Tarihi: 27.05.2020]

Isah, H., Trundle, P., & Neagu, D. (2014, September). Social Media Analysis for Product Safety Using Text Mining and Sentiment Analysis. In *2014 14th Uk Workshop on Computational Intelligence (Ukci)* (Pp. 1-7). Ieee.

Karabulut, Y. E., & Küçüksille, E. U. (2018). Twitter Profesyonel İzleme ve Analiz Aracı. *Teknik Bilimler Dergisi*, *8*(2), 17-24.

Kasaba, E., & Yıldıztepe, E. (2016). *Destek Vektör Makinesi Yöntemi İle Bir Duygu Çözümlemesi*. Akademik Bilişim.

Kızılkaya, Y. M. (2018). Duygu Analizi ve Sosyal Medya Alanında Uygulama.

Küçükkartal, H. K. (2020). Twitter'daki Verilere Metin Madenciliği Yöntemlerinin Uygulanması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, *1*(2), 10-13.

Lalji, T., & Deshmukh, S. (2016). Twitter Sentiment Analysis Using Hybrid Approach. *International Research Journal of Engineering and Technology*, *3*(6), 2887-2890.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms And Applications: A Survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113.

Mishra, P., Rajnish, R., & Kumar, P. (2016, October). Sentiment Analysis of Twitter Data: Case Study on Digital India. In *2016 International Conference on Information Technology (Incıte)-The Next Generation it Summit on the Theme-Internet of Things: Connect Your Worlds* (Pp. 148-153). Ieee.

Mohammad, S. M., & Yang, T. W. (2011, June). Tracking Sentiment in Mail: How Genders Differ on Emotional Axes. In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis* (Pp. 70-79). Association For Computational Linguistics.

Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., & Ureña-López, L. A. (2014). Ranked wordnet graph for sentiment polarity classification in twitter. *Computer Speech & Language, 28*(1), 93–107.

Murthy, D. (2018). Twitter. Cambridge, Uk: Polity Press, 24,25. (Murthy, 2018).

Nalçakan, Y., Bayramoğlu, Ş. S., & Tuna, S. (2015). *Sosyal Medya Verileri Üzerinde Yapay Öğrenme İle Duygu Analiziçalışması*. Technical Report.

Nizam, H., & Akın, S. S. (2014). Sosyal Medyada Makine Öğrenmesi İle Duygu Analizinde Dengeli Ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması. *Xıx. Türkiye'de İnternet Konferansı*.

Oğuz, B. (2009). *Metin Madenciliği Teknikleri Kullanılarak Kulak Burun Boğaz Hasta Bilgi Formlarının Analizi*. [Yüksek Lisans Tezi, Akdeniz Üniversitesi Sağlık Bilimleri Enstitüsü].

Özdeş, M. (2017). *Büyük Veri Araçlarını Kullanarak Duygu Analizi Gerçekleştirimi* [Master's Thesis, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü].

Pak, A., & Paroubek, P. (2010, May). Twitter As A Corpus For Sentiment Analysis and Opinion Mining. In *Lrec* (Vol. 10, No. 2010, Pp. 1320-1326).

Pang, B., & Lee, L. (2005, June). Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics* (Pp. 115-124). Association For Computational Linguistics.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1-135.

Parlar, T., Saraç, E., & Özel, S. A. (2017, May). Comparison of Feature Selection Methods for Sentiment Analysis on Turkish Twitter Data. In *2017 25th Signal Processing and Communications Applications Conference (Siu)* (Pp. 1-4). Ieee.

Pennacchiotti, M., & Popescu, A. M. (2011, July). A Machine Learning Approach To Twitter User Classification. In *Fifth İnternational Aaai Conference On Weblogs And Social Media*.

Pilavcılar, İ. F. (2007). Metin Madenciliği ile Metin Sınıflandırma.

Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M., & Kivikangas, M. (2006). Spatial Presence and Emotions During Video Game Playing: Does it Matter with Whom You Play? *Presence: Teleoperators and Virtual Environments*, *15*(4), 381-392.

Ravi, K., & Ravi, V. (2015). A Survey On Opinion Mining And Sentiment Analysis: Tasks, Approaches and Applications. *Knowledge-Based Systems*, *89*, 14-46.

Sağlam, F. (2019). Otomatik Duygu Sözlüğü Geliştirilmesi ve Haberlerin Duygu Analizi.

Sahayak, V., Shete, V., & Pathan, A. (2015). Sentiment Analysis on Twitter Data. *International Journal of Innovative Research in Advanced Engineering (Ijirae)*, *2*(1), 178-183.

Saif, H., He, Y., Fernandez, M., & Alani, H. (2016). Contextual semantics for sentiment analysis of Twitter. *Information Processing & Management, 52*(1), 5-19.

Saleh, M. R., Martín-Valdivia, M. T., Montejo-Ráez, A., & Ureña-López, L. A. (2011). Experiments with Svm to Classify Opinions in Different Domains. *Expert Systems with Applications*, *38*(12), 14799-14804.

Shukri, S. E., Yaghi, R. I., Aljarah, I., & Alsawalqah, H. (2015, November). Twitter Sentiment Analysis: A Case Study in the Automotive İndustry. In *2015 Ieee Jordan Conference on Applied Electrical Engineering and Computing Technologies (Aeect)* (Pp. 1-5). Ieee.

Shukri, S. E., Yaghi, R. I., Aljarah, I., & Alsawalqah, H. (2015, November). Twitter Sentiment Analysis: A Case Study in the Automotive İndustry. In *2015 Ieee Jordan Conference on Applied Electrical Engineering and Computing Technologies (Aeect)* (Pp. 1-5). Ieee.

Şeker, S. E. (2016). İş Planı (Business Plan). *Yönetim Bilişim Sistemleri Ansiklopedisi, 3*(3), 5.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, *37*(2), 267-307.

Tahiroğlu, B. T. (2008). Derlem Dilbilimi ve Güncel Gelişmeler. *Uluslararası Türk Dili Kurultayı*, 20-25.

Tetsuya, N. (2003). Jeonghee Yi. Sentiment Analysis: Capturing Favorabillty Using Natural Language Processing. In *Proceedings of the 2nd International Conference on Knowledge Capture, America*.

Tong, R. M. (2001). An Operational System for Detecting and Tracking Opinions in On-Line Discussion. In *Working Notes of the Acm Sıgır 2001 Workshop on Operational Text Classification* (Vol. 1, P. 6).

Tunalı, V. (2011). Metin Madenciliği İçin İyileştirilmiş Bir Kümeleme Yapısının Tasarımı ve Uygulaması.

Türkmenoğlu, C. (2015). *Türkçe Metinlerde Duygu Analizi* [Doctoral Dissertation, Fen Bilimleri Enstitüsü].

Qaisi, L. M., & Aljarah, I. (2016, July). A Twitter Sentiment Analysis For Cloud Providers: A Case Study of Azure Vs. Aws. In *2016 7th International Conference on Computer Science and Information Technology (Csıt)* (Pp. 1-6). Ieee.

Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A System For Real-Time Twitter Sentiment Analysis of 2012 Us Presidential Election Cycle. In *Proceedings of the Acl 2012 System Demonstrations* (Pp. 115-120). Association for Computational Linguistics.