Acta Infologica
acin 2024, 8 (1), 23–33
DOI: 10.26650/acin.1418834

Research Article / Araştırma Makalesi

Sentiment Analysis on GPT-4 with Comparative Models Using Twitter Data*

Mustafa Özel[1], Özlem Çetinkaya Bozkurt[2]

[1]Burdur Mehmet Akif Ersoy University, Social Sciences Institute, Burdur, Türkiye
[2]Burdur Mehmet Akif Ersoy University, Bucak Faculty of Business Administration, Department of Business Administration, Burdur, Türkiye

Corresponding author : Mustafa Özel
E-mail : mozel@mehmetakif.edu.tr

*Bu çalışma 10. Uluslararası Yönetim Bilişim Konferansında (IMISC 2023) sözlü olarak sunulan bildirinin geliştirilmiş halidir.

ABSTRACT
Every day, people from all over the world use Twitter to talk about many different topics using hashtags. Since ChatGPT was launched, researchers have been studying how people perceive it in society. This research aims to find out what Turkish Twitter users think about OpenAI's latest AI model called Generative Pre-trained Transformer 4 (GPT-4). The quantitative data used in this study consist of hashtags on the topic of GPT-4 and involve 2,978 tweets on this topic that were shared on Twitter between March 14-April 9, 2023. The study uses TextBlob sentiment scores to classify the tweets and support vector machines, logistic regression, XGBoost, and random forest algorithms to classify the sentiment of the dataset. The results from the logistic regression, XGBoost, and support vector methods are in close alignment. All parameter findings indicate dependable machine learning, emphasizing the models' success in classifying tweet sentiment.

Keywords: Sentiment analysis, social media, Twitter, X, natural language processing

Submitted : 12.01.2024
Revision Requested : 20.03.2024
Last Revision Received : 15.04.2024
Accepted : 16.04.2024
Published Online : 03.06.2024

1. INTRODUCTION

Social media has become an integral part of life as it allows people to communicate, share, learn, and express themselves over common interests in real time. Just like food, water, and home, social media has become a basic need (Dandekar et al., 2018, p. 882). People often share their intentions, troubles, solutions, and moods on social media. Numerous users actively use social media platforms, and through these platforms, users with various opinions express their opinions and thoughts via text (Rahman et al., 2023, p. 1069). Many social media platforms exist and are available today, such as Twitter, Facebook, WhatsApp, Instagram, and TikTok. Among these social media platforms, Twitter is one of the most effective for getting ideas about issues and events. Many people share their perspectives on different issues on Twitter, making social media platforms such as Twitter an open source of data.

Twitter users around the world discuss various topics through hashtags every day. Recently, the most interesting and intriguing of these hashtags has been ChatGPT, with GPT meaning Generative Pre-trained Transformer. Since the introduction of ChatGPT, researchers have started to investigate the public's attitude toward it. Numerous studies have discussed the broad social implications of ChatGPT (Abdullah et al., 2022) and its domain-specific potentials (Munggaran et al., 2023; Botchu & Iyengar, 2023). These studies have predominantly utilized research methods from the social sciences, including interviews, user experience, and expert-based perspectives. Other studies in the literature that are more relevant to the current study have used different computational techniques to explore the public sentiment of ChatGPT using social media data. As an illustration, Liu et al. (2023) endeavored to furnish a thorough examination of extant studies on ChatGPT and its prospective implementations in diverse domains. In pursuit of this objective, they carried out an extensive examination of ChatGPT-related papers in the JarXiv repository and attempted to provide understanding into ChatGPT's skills, potential ramifications, ethical considerations, and prospects for future progress in this domain. The results indicate a notable and increasing fascination with ChatGPT/GPT-4 investigations across several disciplines such as education, history, mathematics, medicine, and physics, that are primarily centered around direct applications of natural language processing. Feng et al. (2023) conducted a study on Twitter and Reddit users to investigate the potential for ChatGPT in code generation and attitudes toward ChatGPT. The study revealed fear as the prevailing emotion linked to ChatGPT code generation, surpassing such emotions as happiness, anger, surprise, and sadness. Lampropoulos et al. (2023) used Twitter data to report on the perspectives, attitudes, emotions, and discourses surrounding the use of ChatGPT for general and educational purposes. Their results demonstrated the broad applicability of artificial intelligence (AI) tools, as well as the versatility of ChatGPT. Li et al.'s (2023) study analyzed Twitter data to ascertain the primary apprehensions regarding the use of ChatGPT in the field of education and found that, while a generally positive sentiment was present, concerns also occurred in five areas: learning outcomes and skill development, academic integrity, skill limitations, political and social impacts, and workforce challenges. Meanwhile, Keskin's (2023) study analyzed news publications to identify the main topics, focusing on how ChatGPT is addressed in Türkiye's Internet agenda. As a result of the analysis, prominent themes were identified such as education, new developments in ChatGPT, business life, information gathering, coding and development, the IT sector, daily life, investment consultancy, and creative content production. Korkmaz et al. (2023) performed a sentiment analysis on Twitter posts related to ChatGPT to thoroughly assess the emotions and opinions expressed over the initial two months following the announcement of ChatGPT. The results showed the majority of users who'd used ChatGPT for the first time to have found the experience successful and to be satisfied with ChatGPT; however, it also aroused negative emotions such as fear and anxiety in some users.

Unlike other research, this study collects and translates into English Turkish tweets about GPT-4, an artificial intelligence application that is on the agenda all over the world before performing the sentiment analyses in an attempt to investigate the general public attitude toward ChatGPT in Türkiye. The study then categorizes these public attitudes according to their sentiment scores using machine learning techniques. In line with this, it has collected tweets sent between March 14-April 9, 2023 and classified the preprocessed tweets based on their sentiment scores using the TextBlob dictionary. The study also subjected the dataset that had been classified according to emotion scores to emotion classification using the logistic regression, support vector machines, and random forest machine learning algorithms. Finally, the study has used Python programming language for data preprocessing and other operations. The findings obtained from this exploratory study can be useful for both the public interested in ChatGPT as well as the developers of ChatGPT-related technology.

Table 1. Some of the Studies on Chat-GPT

Author(s)	Title of the Study	Year
Abdullah et al.	Fundamentals, Applications and Social Impacts	2022
Munggaran et al.	Sentiment Analysis of Twitter Users' Opinion Data Regarding the Use of ChatGPT in Education	2023
Botchu & Iyengar	Will ChatGPT Drive Radiology in the Future?	2023
Liu et al.	Summary of Chat-GPT/GPT-4 Research and Perspective Towards the Future of Large Language Models	2023
Feng et al.	Investigating Code Generation Performance of ChatGPT with Crowdsourcing Social Data	2023
Lampropoulos et al.	A Social Media Data Analysis of General and Educational Use of ChatGPT: Understanding Emotional Educators	2023
Li et al.	ChatGPT in Education: A Discourse Analysis of Worries and Concerns on Social Media	2023
Keskin, E.	Yapay zekâ sohbet robotu ChatGPT ve Türkiye internet gündeminde oluşturduğu temalar	2023
Korkmaz et al.	Analysing the User's Sentiments of ChatGPT Using Twitter Data	2023

2. SOCIAL MEDIA AND DATA ANALYSIS

2.1. Twitter and Chat-GPT

Twitter ranks highly among social media platforms for staying informed about current events and trending topics. Many people share their perspectives on different topics on Twitter, making social media platforms such as Twitter an open source of data. Three basic symbols are used in this communication that are realized through a common universal tweet system. Using the @ symbol followed by a Twitter account name tags (mentions) the person or organization being tagged. Retweet (i.e., RT) is the sharing of an interesting tweet by another Twitter user. A hashtag (i.e., #) is a largely user-generated mechanism for labelling and collating messages (i.e., tweets) on a particular topic. Users who want to send messages add short words, sentences, or abbreviations to their messages, preceded by # to indicate that their messages address certain themes (Aladwani, 2015, p. 16; Bruns & Burgess, 2011, p. 2; Suh et al., 2010, p. 177).

Launched on November 30, 2022, Chat-GPT is an interactive chatbot developed by the AI company OpenAI (Kirmani, 2022, p. 574). Chat-GPT understands what is requested by the user, interprets it, and produces appropriate responses in almost natural human language. Besides practical applications, Chat-GPT's ability to successfully perform complex tasks has made it a major innovation in the fields of natural language processing and AI (Lund & Wang, 2023, p. 26). Finally, OpenAI has introduced GPT-4, the latest member of the GPT family. Many users have praised GPT-4's most recent improvements and distinctive capabilities (Koubaa, 2023, p. 1). Unlike its previous version, GPT-4 is a multimodal and large-scale model that also accepts images as input and can produce text output. GPT-4 has outperformed many traditional natural language processing (NLP) tests, as well as older large language models and more advanced systems (OpenAI, 2023; Aydın & Karaarslan, 2023, p. 4).

2.2. Machine learning

Machine learning (ML) is the realization of knowledge transfer similar to that of humans. In machine learning, a training model is created using data, and the decision-making quality of the system is improved. This learning method is use to try and ensure that the system makes successful predictions or successful classifications against similar data in the future (Doğan, 2022, p. 914).

2.2.1. Logistic Regression

Logistic regression analyzes data to estimate the probability of a certain outcome (dependent variable) based on its relationship with other factors (independent variables; Bircan, 2004, p. 186). It is an algorithm used to solve both regression and classification problems with both numerical and textual data. Three methods are found for applying the logistic regression classification algorithm in real life: binary (binomial), ordinal, and multinominal, with the multiclass (multinomial) approach allowing the dependent variable to have three or more different values (Ulaş & Karabay, 2020, p. 271).

2.2.2. *Support Vector Machines*

Support vector machines are usually divided into linear and non-linear problems. The purpose of using support vector machines in linear problems is to separate the features of the classes as far as possible from each other, with a hyperplane passing through the features (Metlek & Kayaalp, 2020, p. 2217). Nonlinear classifiers are used in non-linear situations. In such cases, the dataset is shifted from a two-dimensional to a three-dimensional space, and mapping is performed. The non-linear mapping approach moves the two-dimensional dataset to the three-dimensional feature space, enabling the linear separation of the dataset (Ayhan & Erdoğmuş, 2014, p. 185).

2.2.3. *Random Forest*

Random forest is an algorithm that creates more than one decision tree during the classification process and thus increases the classification rate. Randomly selected decision trees together form the decision forest (Aydın, 2018, p. 172). Random forest classifier is a prediction tool that uses the average to improve prediction accuracy and prevent overfitting by applying a set of decision tree classifiers to different subsamples of the dataset. The subsample size is always equal to the original input sample size (Veranyurt et al., 2020, p. 279).

2.2.4. *XGBoost*

XGBoost is a decision tree-based machine learning algorithm and a supervised learning algorithm used for classification and regression, with high-value results being obtainable in the shortest amount of time with less resource consumption. XGBoost operates similarly to the random forest algorithm. While bagging is applied in the random forest algorithm, boosting is applied in the XGBoost algorithm (Turan & Polat, 2024, p. 99; Tekin & Yaman, 2023, p. 156).

2.3. Sentiment Analysis

Sentiment analysis is the process of collecting and analyzing people's opinions, thoughts, and impressions on various topics (Wankhade et al., 2022, p. 5731). The beginning and rapid growth of the field coincides with the beginning of social media on the Web, forum discussions, blogs, microblogs, microblogging, Twitter, and social networks. Since the early 2000s, sentiment analysis has emerged as a highly dynamic field of study within the NLP domain. Sentiment analysis has been disseminated beyond the field of computer science to the realm of management sciences and various other disciplines, including marketing, finance, and health. This is because ideas are at the heart of almost all human activities and significantly influence human behavior. The values one believe in, one's reality, and the choices one makes depend to a large extent on how others view and evaluate the world. Therefore, one often considers the opinions of others when making a decision. This is true not only for humans but also for organizations (Zhang et al., 2018, p. 1). Due to the daily increase in user-generated data on the Web, this content needs to be analyzed in order to know users' opinions, thus increasing the demand for sentiment analysis research (Agarwal & Mittal, 2013, p. 14).

Sentiment analysis falls into three main categories: dictionary-based, ML, and hybrid approaches. Dictionary-based methods leverage pre-existing sentiment lexicons for unsupervised classification, while ML methods rely on training data labeled for supervised learning. As the name suggests, hybrid approaches combine elements of both dictionary and ML techniques (Biltawi et al., 2016, p. 339).

3. MATERIALS AND METHODS

This section shows the model of the study (see Figure 1) and explains the study steps, such as obtaining the dataset for sentiment analysis on Twitter using ML, data preprocessing steps, data labeling, data separation and modeling, model comparison s, and performance measurements

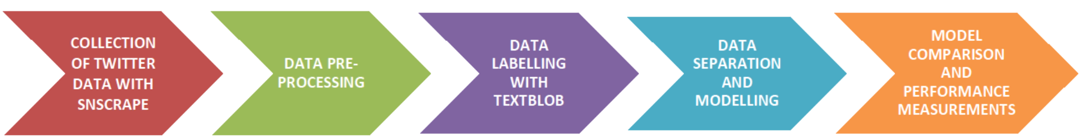

Figure 1. Application steps.

3.1. Dataset

The study benefits from Turkish tweets containing the keyword "GPT-4" that were shared on Twitter between March 14-April 9, 2023. A total of 3,041 tweets were accessed using the Snscrape Library, and a Python library was used to collect data from Twitter. The obtained dataset contains Datetime, Tweet ID, and Text information. Figure 2 shows examples of the dataset that is used.

...	Datetime	Tweet Id	Text
0	2023-04-08 23:29:23+00:00	1644844920593690112	GPT-5 yakın zamanda hayatımıza girecek ve etki...
1	2023-04-08 21:34:34+00:00	1644816024909360128	@Gentlem4nJack GPT-3.5+Midjourney V5:_x000D_\n...
2	2023-04-08 21:10:05+00:00	1644809864017779968	GPT-4: _x000D_\nKuantum mekaniğinde "gözlemci...
3	2023-04-08 21:09:23+00:00	1644809689224420096	@dayiekonomi Dayı bu hangi sürümü şu yeni çık...
4	2023-04-08 20:51:09+00:00	1644805098277669888	@K2adir GPT-4'e sorarım. Muhakkak bi bildiği v...

Figure 2. Tweet dataset example.

3.2. Data Preprocessing

The next preprocessing step removes the hashtags, mentions, URLs, and emojis. To remove hashtags (words starting with #), mentions (words starting with @), emojis, and URLs, the study uses the Python Regular Expression Syntax (RE) module, which is a powerful tool for finding, matching, replacing, or removing specific patterns in text. These operations are performed using the function *re.sub()*.

After removing hashtags, mentions, URLs, and emojis, the Turkish tweets are converted into normal text and then translated into English in the next step. For this process, the study uses the deep_translator library, which is used for simple translations between different languages. The next step converts uppercase letters in the tweets into lowercase letters. Afterward, the punctuation marks are removed. The conversion from uppercase to lowercase is done with the previously used *lower()* method. To remove punctuation marks, the *string. punctuation* command in the Python string module is used to remove punctuation marks or replace them with specific punctuation marks.

Textual data consist of many redundant words (e.g., this, that, of, or, and, with, the) that are not context-related and will not help in classifying the textual contexts of a tweet but do help humans understand it correctly. These are called stop words (Verma et al., 2019). This next stage filters out the frequently used stop words in a language using the *stop_words* tool, after which the tokenization process is applied. Tokenization involves segmenting the text according to its features, such as spaces and punctuation marks (Küçükkartal, 2020, p. 11; Kahya, 2021, p. 12). Stemming (finding the root) is performed next. Stemming and lemmatization are different methods. Stemming is applied to remove inflectional prefixes or suffixes from words. Words with the same meaning and spelling are considered to be different words according to a prefix or suffix. Stemming is used to avoid this. For instance, words used with different inflections, such as come, coming, and was coming, can all be reduced to the root "come". Lemmatization is the process that takes into account the morphological analysis of words and accordingly separates the meaningful word into its roots (Ağralı & Aydın, 2021, p. 28). For these processes, the Natural Language Toolkit (NLTK) library of the Python programming language is most commonly used in NLP and thus is also used here. For stopwords, the stop_words sub-module of the NLTK library is used, and the NLTK library's *nltk*. stem module is used for stemming.

3.3. Data Labeling

The study has adopted a dictionary-based approach to determine whether the preprocessed tweets contain a positive, negative, or neutral meaning. The dictionary-based approach uses a sentiment dictionary containing opinion words which are then matched with the data to determine polarity. As a result of this matching, sentiment scores are assigned to the opinion words that define the positive, negative, and neutral scores of the words in the dictionary (Hardeniya & Borikar, 2016, p. 318).

This study also uses the TextBlob Library as a dictionary-based approach for labeling the tweets. TextBlob is an open-source Python library built upon NLTK and analyzes textual content to assign polarity scores ranging from -1 (negative) to 1 (positive). It achieves this by meticulously examining each word within a text fragment and assigning semantic scores to individual words. These scores are then meticulously weighted, effectively calculating a weighted average to determine a comprehensive score for the entire sentence based on the polarity contributions of each word (Zahoor & Rohilla, 2020, p. 538). The Textblob library can only classify English texts into three types: positive, negative, and neutral. A polarity value greater than 0 is considered positive, less than 0 is considered negative; and equal to 0 is considered neutral (Diyasa et al., 2021, p. 4). With the TextBlob library, tweets are labeled as positive, negative, or neutral according to their polarity values. As a result of TextBlob's sentiment outputs, 724 (37.17%) of the user tweets were determined to be positive, 304 (15.61%) to be negative, and 920 (47.22%) to be neutral (Figure 3). Thus, the dataset has been prepared for comparing the models.

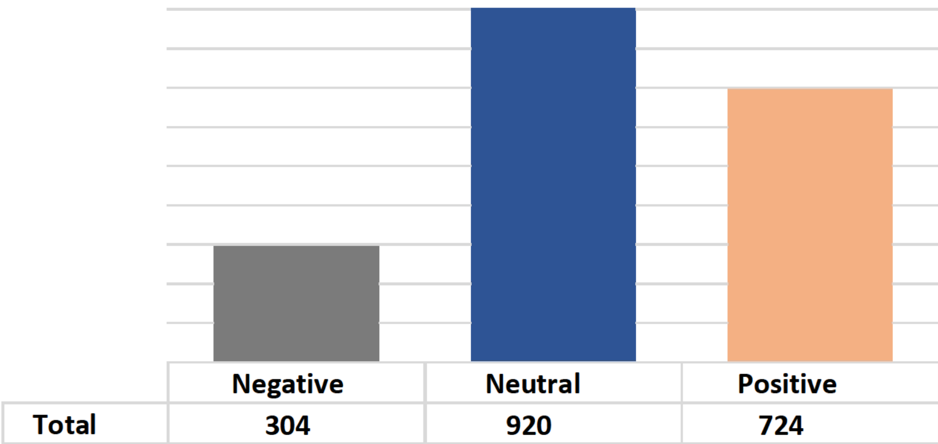

Figure 3. Textblob emotion distributions.

3.4. Word Cloud

This study's tweets labeled as negative, neutral, or positive using sentiment analysis are now visualized with a word cloud. Word clouds are the easiest and most preferred visualization method, as it allows one to visualize the most frequently repeated words in a dataset and to comment on the dataset by looking at these words.

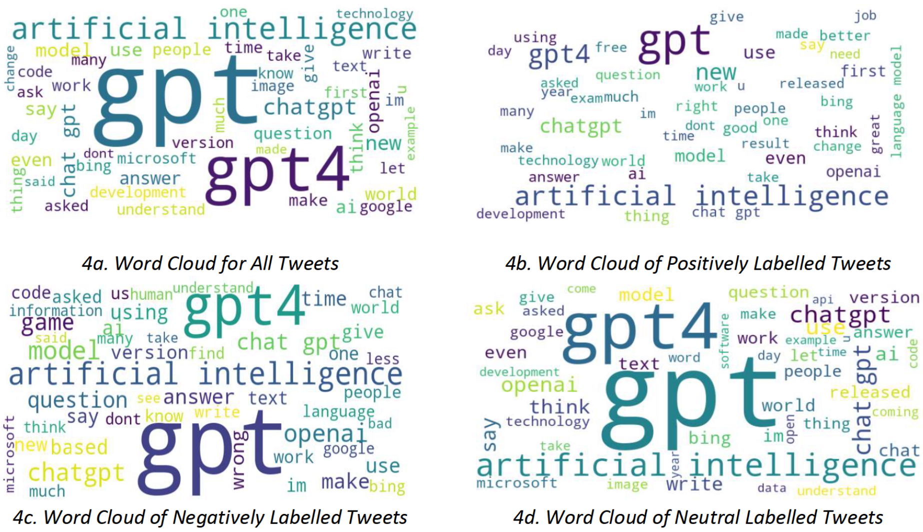

Figure 4. Word clouds of the tweets; a) all tweets, b) positive tweets, c) negative tweets, d) neutral tweets.

When analyzing the distributions of the words used in all the Twitter posts, the most frequently used word is *gpt*, followed by *gpt4, intelligence, artificial, chat, chatgpt, new, model, and openai*. The word clouds containing all the tweets and the sentiments are shown in Figures 4a-4d.

3.5. Data Separation and Modeling

The dataset was prepared for analysis after preprocessing and labeling. The next step splits the dataset into two parts: 80% for training and 20% for testing using the *train_test_split* method in the Python scikit-learn library. For obtaining the optimal hyperparameters, the Grid search technique was additionally applied in order to work with the correct parameters. Hence, the classification performances of the ML algorithms are compared using labeled text data. In this case, the text data should be converted into numerical vectors first. This is done using the *CountVectorizer* class in the *sklearn.feature_extraction.text* module of the scikit-learn library. This allows for the text data to be able to be used with the ML algorithms.

3.6. Comparison Models and Performance Measures

The logistic regression (LR), support vector machines (SVM), XGBoost, and random forest (RF) algorithms have been used to classify the Turkish tweets that were obtained from Twitter and that had been subjected to the GPT-4-themed preprocessing steps. The study uses the confusion matrix, accuracy, precision, sensitivity, and F-1 score performance measures to evaluate the success of the models that have been developed within the scope of the research.

4. FINDINGS

This study has subjected the dataset to sentiment classification using four different ML algorithms: SVM, LR, XGBoost, and RF. The study then examined the performance measures listed in Table 2. The ML techniques applied to the text data labeled with TextBlob were seen to provide successful results. The accuracy rates obtained from the LR, XGBoost, and SVM ML methods were observed to be very close to one another, with XGBoost providing slightly more successful results and RF having the lowest accuracy rate. The same situation is observed when analyzing the other parameters. These values indicate the classification performance of the model to have been successful. The obtained parameter results indicate the ML algorithms that were used to have provided successful classification results. The performances of the ML techniques are listed in Table 2.

Table 2. Performance of Machine Learning Techniques

	Accuracy	Sensitivity	Responsiveness	F-1Score
XGBoost	0.87	0.87	0.87	0.86
LR	0.86	0.87	0.86	0.85
DVM	0.85	0.85	0.85	0.84
RF	0.81	0.83	0.81	0.79

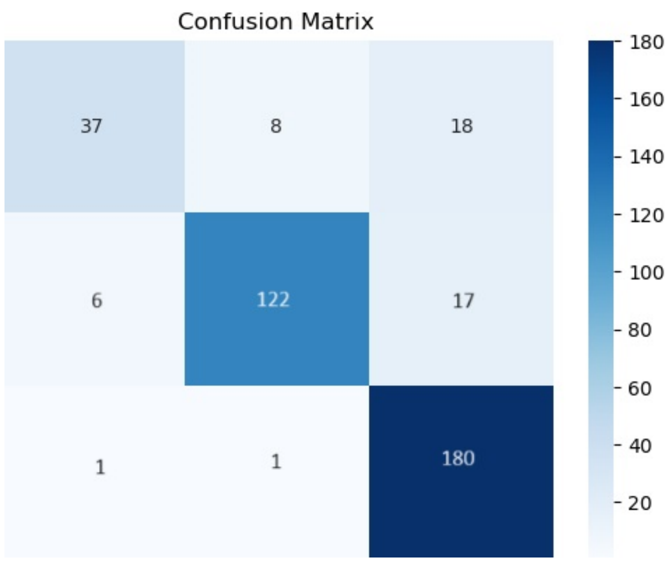

Figure 5. Complexity matrix of the XGBoost algorithm.

Analyzing the complexity matrix of the XGBoost algorithm shows 98.9% (180) of the neutral values, 84.13% (122) of the positive values, and only 58.73% (37) of the negative values to have been correctly classified (see Figure 5).

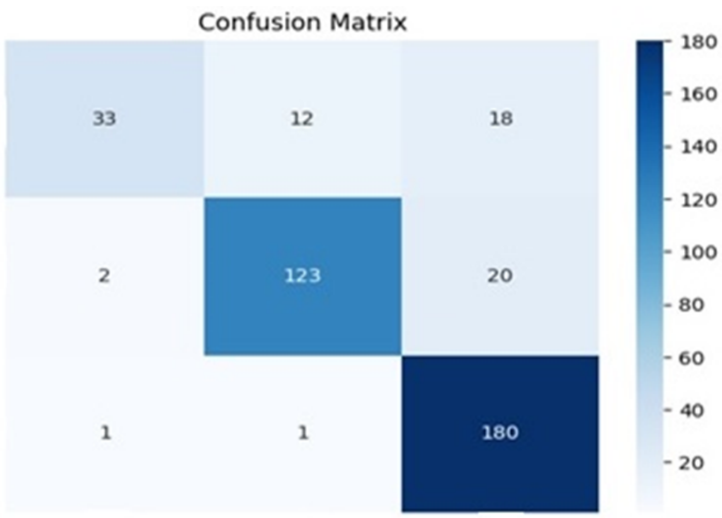

Figure 6. Complexity matrix of the logistic regression algorithm.

Figure 6 depicts the complexity matrix of the LR algorithm. When analyzing the complexity matrix of the LR, 98.9% (180) of the neutral values, 84.83% (123) of the positive values, and 52.38% (33) of the negative values are seen to have been classified correctly.

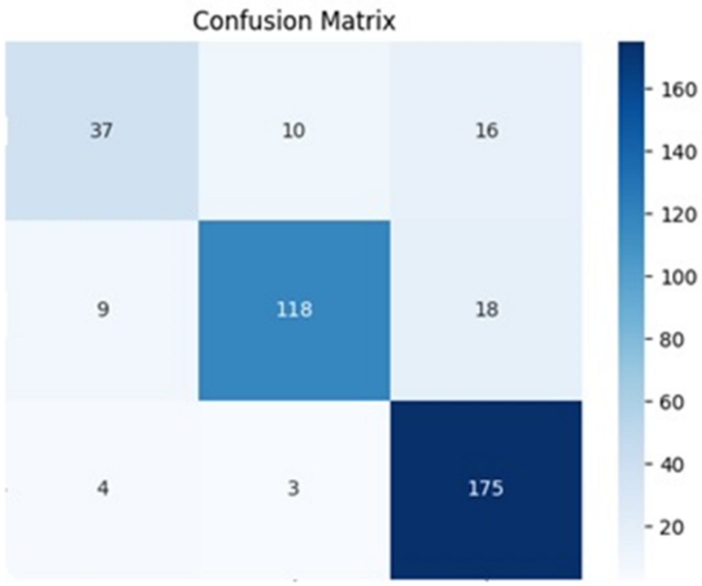

Figure 7. Complexity matrix of the support vector machines algorithm.

Figure 7 represents the complexity matrix of the SVM algorithm. When analyzing the complexity matrix of the SVM, 96.15% (175) of the neutral values, 81.38% (118) of the positive values, and 58.73% (37) of the negative values are observed to have been classified correctly.

When analyzing the complexity matrix of the RF classification, 98.35% (179) of the neutral values, 82.07% (119) of the positive values, and only 28.57% (18) of the negative values were determined to have been classified correctly (see Figure 8).

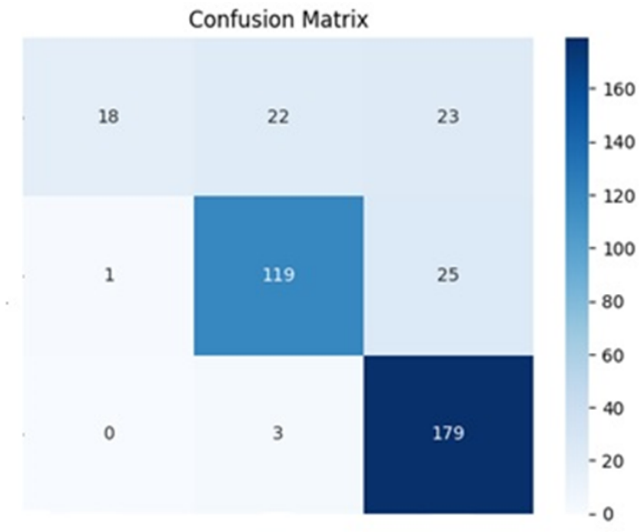

Figure 8. Complexity matrix of the random forest algorithm.

5. DISCUSSION AND CONCLUSION

This study presents the sentiment analysis of Turkish tweets about GPT-4, which was released by OpenAI on March 14, 2023. The study preprocessed the obtained data and then labeled them using the TextBlob dictionary. As a result of the sentiment analysis, 15.61% of the tweets Twitter users posted expressed negative opinions, 37.17% expressed positive opinions, and 47.22% were neutral. The results of the study are consistent with the previous studies by Lampropoulos et al. (2023) and Li et al (2023). The numerically expressed and class-labeled dataset was separated for training and testing. The study also compared the classification performances of the following ML algorithms: SVM, LR, XGBoost, and RF. The results indicate the XGBoost, LR, and SVM algorithms to scored close to one another. However, the XGBoost achieved the highest accuracy rate with an accuracy value of 0.87. All parameter results show that the ML algorithms used in the study to have provided reliable results. Based on these results, the models are seen to have successfully performed the sentiment classification of tweets. When analyzing the complexity matrices related to the results, the XGBoost and LR algorithms were found to have yielded the most successful results when classifying positive and neutral values, while the XGBoost and SVM algorithms were more successful classifying negative values.

Future studies can consider using different ML algorithms, such as Naive Bayes, K-nearest neighbor, and decision trees for the classification of tweets. In addition, tweets on different topics can be analyzed using the same method. Labeled data can also be compared using different sentiment analysis approaches based on a dictionary or on machine learning.

The information obtained from the study can be beneficial for both the general public interested in ChatGPT as well as the developers of ChatGPT-related technology. This information can help these groups create a broader perception of ChatGPT and decide whether they want to use the technology or not. In this way, developers can also understand the social context around ChatGPT and better optimize this technology.

Peer Review: Externally peer-reviewed.
Author Contributions: Conception/Design of Study- M.Ö., Ö.Ç.B.; Data Acquisition- M.Ö., Ö.Ç.B.; Data Analysis/Interpretation- M.Ö., Ö.Ç.B.; Drafting Manuscript- M.Ö., Ö.Ç.B.; Critical Revision of Manuscript- Ö.Ç.B.; Final Approval and Accountability- M.Ö.
Conflict of Interest: The authors have no conflict of interest to declare.
Grant Support: The authors declared that this study has received no financial support.

ORCID IDs of the authors

Mustafa Özel	0009-0001-7384-7486
Özlem Çetinkaya Bozkurt	0000-0002-6218-2570

REFERENCES

Abdullah M., Madain A., & Jararweh Y. (2022). ChatGPT: fundamentals, applications and social impacts. *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*. 1-8. http://dx.doi.org/10.1109/SNAMS58071.2022.10062688

Ağralı, Ö., & Aydın, Ö. (2021). Tweet classification and sentiment analysis on metaverse related messages,. *Journal of Metaverse, 1*(1), 25-30.

Aydın, C. (2018). makine öğrenmesi algoritmalari kullanilarak itfaiye istasyonu ihtiyacinin siniflandirilmasi. *Avrupa Bilim ve Teknoloji Dergisi*, (14), 169-175.

Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. 1-23. *Social Science Research Network (SSRN):* https://ssrn.com/abstract=4341500 or http://dx.doi.org/10.2139/ssrn.4341500.

Ayhan, S., & Erdoğmuş, Ş. (2014). Destek vektör makineleriyle siniflandirma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilim Dergisi, 9*(1), 175–198.

Agarwal, B., Mittal, N., Bansal, P., & Garg, S. (2015). Sentiment analysis using common-sense and context information. Computational *İntelligence and Neuroscience, 2015*, 1-9. http://dx.doi.org/10.1155/2015/715730

Bayram, İ., & Turan, A. (2022). Türkiye'de kripto para farkindaliği ve tutumu: duygu analizi ve istatistiksel analiz ile bir değerlendirme. *Yönetim Bilişim Sistemleri Dergisi, 8*(2), 20-35. https://dergipark.org.tr/en/pub/ybs/issue/75943/1197985

Bengesi S., Oladunni T., Olusegun R., & Audu H. (2023). A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets, IEEE Access, 11(11811-11826). http://dx.doi.org/10.1109/ACCESS.2023.3242290

Bircan, H. (2004). Lojistik regresyon analizi: Tıp verileri üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi*, (8), 185-208.

Biltawi, M., Etaiwi, W., Tedmori, S., Hudaib, A., & Awajan, A. (2016). Sentiment classification techniques for arabic language: A survey. *7th International Conference on Information and Communication Systems (ICICS)*, 339-346.

Botchu, R., & Iyengar, K. P. (2023). Will ChatGPT drive radiology in the future?. *Indian Journal of Radiology and Imaging*. https://doi.org/10.1055/s-0043-1769591

Bruns, A., & Burgess, J. E. (2011). The use of Twitter hashtags in the formation of ad hoc publics. *Paper presented at the 6th European Consortium for Political Research General Conference*, 25 - 27 August 2011, University of Iceland, Reykjavik. http://eprints.qut.edu.au/46515/

Dandekar, A. R., Shrotri, A.P., Hargude, N.V., Awate, P.P., & Patil, N.D. (2018,August). The rise of social media and its impact. *Internatıonal Journal of Research And Analytıcal Revıews 5*(3), 882-886.

Diyasa, I. G. S. M., Mandenni, N. M. I. M., Fachrurrozi, M. I., Pradika, S. I., Manab, K. R. N., & Sasmita, N. R. (2021). Twitter sentiment analysis as an evaluation and service base on python textblob. *In IOP Conference Series: Materials Science and Engineering*, 1125(1), 1-12. https://doi.org/10.1088/1757-899X/1125/1/012034

Doğan, G. (2022). Makine öğrenmesi algoritmalari ile betonarme kirişlerin burulma momenti tahmini. *El-Cezeri El-Cezerî Fen ve Mühendislik Dergisi, 9*(2), 912-924.

Feng, Y., Vanam, S., Cherukupally, M., Zheng, W., Qiu, M., & Chen, H. (2023). Investigating code generation performance of chatgpt with crowdsourcing social data. *In Proceedings of the 47th IEEE Computer Software and Applications Conference*, 1-10.

Hardeniya, T., & Borikar, D. A. (2016). Dictionary based approach to sentiment analysis-a review. *International Journal of Advanced Engineering, Management and Science, 2*(5), 317-322.

Kahya, A. N. (2021). Wikipedia'daki verilere metin madenciliği yöntemlerinin uygulanması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi (ESTUDAM) Bilişim Dergisi, 2*(1), 11-14.

Kaur C., & Sharma A. (2022). Social issues sentiment analysis using python. *2020 5th International Conference on Computing, Communication and Security (ICCCS)*, 1-6. http://dx.doi.org/10.1109/ICCCS49678.2020.9277251

Keskin, E. K. (2023). Yapay zekâ sohbet robotu ChatGPT ve Türkiye internet gündeminde oluşturduğu temalar. *Yeni Medya Elektronik Dergisi, 7*(2), 114-131.

Kirmani, A. R. (2022), Artificial Intelligence-Enabled Science Poetry. ACS Energy Letters, 8, 574–576. https://doi.org/10.1021/acsenergylett.2c02758

Korkmaz, A., Aktürk, C., & Talan, T. (2023). Analyzing the user's sentiments of ChatGPT using twitter data. *Iraqi Journal for Computer Science and Mathematics, 4*(2), 202-214.

Koubaa, A. (2023). GPT-4 vs. GPT-3.5: A concise showdown. Preprint. https://doi.org/10.20944/preprints202303.0422.v1

Küçükkartal, H. K. (2020). Twitter'daki verilere metin madenciliği yöntemlerinin uygulanması. *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 1*(2), 10-13.

Lampropoulos, G., Ferdig, R. E., & Kaplan-Rakowski, R. (2023). A social media data analysis of general and educational use of chatgpt: understanding emotional educators. *Available at SSRN*: https://ssrn.com/abstract=4468181 or http://dx.doi.org/10.2139/ssrn.4468181

Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). ChatGPT in education: A discourse analysis of worries and concerns on social media, 1-35. arXiv preprint arXiv:2305.02201.

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., ... & Ge, B. (2023). Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models. 1-35. arXiv preprint arXiv: 2304.01852. https://doi.org/10.48550/arXiv.2304.01852.

Lund, B.D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News, 40*(3), 26-29. https://doi.org/10.1108/LHTN-01-2023-0009.

Metlek, S., & Kayaalp, K. (2020). Derin öğrenme ve destek vektör makineleri ile görüntüden cinsiyet tahmini. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 8*(3), 2208-2228.

Munggaran, J. P., Alhafidz, A. A., Taqy, M., Agustini, D. A. R., & Munawir, M. (2023). Sentiment analysis of twitter users' opinion data regarding the use of ChatGPT in education. *Journal of Computer Engineering, Electronics and Information Technology, 2*(2), 75-88.

Rachman F. H., Imamah, & Rintyarna B. S. (2022). Sentiment analysis of madura tourism in new normal era using text blob and KNN with hyperparameter tuning. *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, 23-27. http://dx.doi.org/10.1109/ISMODE53584.2022.9742894

Ulaş, M., & Karabay, B. (2020). Terör saldırılarını içeren büyük verinin makine öğrenmesi teknikleri ile analizi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 32*(1), 267-277.

Okoloegbo, C. A., Eze, U. F., Chukwudebe, G. A., & Nwokonkwo, O. C. (2022). Multilingual Cyberbullying Detector (CD) Application for Nigerian Pidgin and Igbo Language Corpus. *In 2022 5th Information Technology for Education and Development (ITED)*, 1-6. https://doi.org/10.1109/ITED56637.2022.10051345

OpenAI (2023). GPT-4 Technical Report. https://cdn.openai.com/papers/gpt-4.pdf.

Rahman, S., Jahan, N., Sadia, F., & Mahmud, I. (2023). Social crisis detection using twitter based text mining-a machine learning approach, BULLETIN OF ELECTRICAL ENGINEERING AND INFORMATICS, 12(2), 1069-1077. https://doi.org/10.11591/eei.v12i2.3957

Santra, A. K., & Christy, C. J. (2012). Genetic algorithm and confusion matrix for document clustering. *International Journal of Computer Science Issues (IJCSI), 9*(1), 322-328.

Shinde, P. P., & Shah, S. (2018). *A review of machine learning and deep learning applications.* In 2018 Fourth İnternational Conference on Computing Communication Control And Automation (ICCUBEA), 1-6. https://doi.org/10.1109/iccubea.2018.8697857

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. *In 2010 Second İnternational Conference on Social Computing*, 177-184.

Tekin, R., & Yaman, O. (2023). Akıllı ev sistemleri için XGBoost tabanlı saldırı tespit yöntemi. *Journal of Intelligent Systems: Theory and Applications, 6*(2), 152-158. https://doi.org/10.38016/jista.1075054

Turan, A. K., & Polat, H. (2024). Yarı denetimli makine öğrenmesi yöntemini kullanarak müzik türlerinin tespiti. *Gazi University Journal of Science Part C: Design and Technology, 12*(1), 92-107. https://doi.org/10.29109/gujsc.1352477

Veranyurt, Ü., Deveci, A., Esen, M. F., & Veranyurt, O. (2020). Makine öğrenmesi teknikleriyle hastalık sınıflandırması: Random forest, k-nearest neighbour ve adaboost algoritmaları uygulaması. *Uluslararası Sağlık Yönetimi ve Stratejileri Araştırma Dergisi, 6*(2), 275-286.

Verma, A., Mittal, V., & Dawn, S. (2019). *FIND: Fake information and news detections using deep learning*. In 2019 Twelfth İnternational Conference On Contemporary Computing (IC3), 1-7. https://doi.org/10.1109/IC3.2019.8844892

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A Survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review, 55*(7), 5731-5780.

Zahoor, S., & Rohilla, R. (2020). Twitter sentiment analysis using lexical or rule based approach: a case study. *In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 537-542.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery, 8*(4), 1-25.

How cite this article

Özel, M., & Çetinkaya Bozkurt, Ö. (2024). Sentiment analysis on GPT-4 with comparative models using Twitter data. *Acta Infologica, 8*(1), 23-33. https://doi.org/10.26650/acin.1418834