IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 June 2023, accepted 28 June 2023, date of publication 3 July 2023, date of current version 7 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3291592

RESEARCH ARTICLE

Sentiment Analysis in Turkish Question Answering Systems: An Application of Human-Robot Interaction

KADIR TOHMA[1], HALIL IBRAHIM OKUR[1], YAKUP KUTLU[1], AND AHMET SERTBAS[2]
[1]Department of Computer Engineering, Iskenderun Technical University, 31200 Iskenderun, Turkey
[2]Department of Computer Engineering, Istanbul University-Cerrahpasa, 34452 Istanbul, Turkey

Corresponding authors: Kadir Tohma (kadir.tohma@iste.edu.tr) and Halil Ibrahim Okur (hibrahim.okur@iste.edu.tr)

ABSTRACT The use of the sentiment analysis technique, which aims to extract emotions and thoughts from texts, has become a remarkable research topic today, where the importance of human-robot interaction is gradually increasing. In this study, a new hybrid sentiment analysis model is proposed using machine learning algorithms to increase emotional performance for Turkish question and answer systems. In this context, as a first, we apply text preprocessing steps to the Turkish question-answer-emotion dataset. Subsequently, we convert the preprocessed question and answer texts into text vector form using Pretrained Turkish BERT Model and two different word representation methods, TF-IDF and word2vec. Additionally, we incorporate pre-determined polarity vectors containing the positive and negative scores of words into the question-answer text vector. As a result of this study, we propose a new hybrid sentiment analysis model. We separate vectorized and expanded question-answer text vectors into training and testing data and train and test them with machine learning algorithms. By employing this previously unused method in Turkish question-answering systems, we achieve an accuracy value of up to 91.05% in sentiment analysis. Consequently, this study contributes to making human-robot interactions in Turkish more realistic and sensitive.

INDEX TERMS Artificial intelligence, natural language processing, question answering, sentiment analysis.

I. INTRODUCTION

People are emotional beings who can interact socially using their emotions and have the ability to perceive emotions of other people. Emotions need to play an important role in the development of social and humanistic information systems. In this context, it is important to apply sentiment analysis methods for effective human-robot interactions [1].

Sentiment Analysis is an important subtask of Natural Language Processing (NLP), aiming to extract emotions and opinions from texts [2]. Opinions and emotions are critical pieces of information that influence decision-making processes. Sentiment analysis can be applied in various areas such as determining public opinions on political policies, business intelligence and market analysis, measuring customer satisfaction, predicting movie sales, and ensuring realism and empathy in human-robot interactions based on questions and answers [3].

In sentiment analysis studies, sentiment detection can be performed at sentence, paragraph, and document levels. Although extended texts offer greater suitability for emotional representation, shorter texts are less affected by noisy data, thereby facilitating sentiment analysis. Therefore, while the Twitter platform is frequently used in social media analyses, short dialogue sentences are preferred in question-answering systems [4]. Additionally, it is observed that these studies are generally examined according to negative and positive sentiment classes [5], [6], [7].

The dominant researches in sentiment analysis are based on English. There appears to be a need for further development of its methods and libraries for analyzing other languages. When examining Turkish studies in this context, it becomes evident that three main classification approaches are employed: classical machine learning [8], [9], deep learning [10], [11], and lexicon-based [12], [13]. Moreover, similar

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues.

classification models are favored in question-answering systems [14]. Upon reviewing relevant studies, it has been observed that sentiment analysis and question-answering methods can be employed concurrently [15], [16], [17], [18], [19]. However, no sentiment analysis study was found in Turkish question-answer systems.

The integration of sentiment analysis and question-answering systems is crucial for a better understanding of children's emotional states and the generation of appropriate responses. This approach could be particularly valuable for children with special educational needs, as they often face challenges in expressing their emotional states. Consequently, the fusion of sentiment analysis and question-answering systems could provide interaction that better addresses the emotional needs of these children. This, in turn, could improve their learning process, social skills, and overall quality of life [20]. Human-robot interaction could be significantly enhanced with the ability to understand emotional responses, providing a more effective and sensitive experience in sectors such as education, healthcare, and customer service [21]. Hence, the fundamental motivation of our study is the notion that effective sentiment analysis could facilitate more meaningful and responsive interaction between robots and humans.

In addition to the aforementioned considerations, the role of empathetic response merits further examination. Empathy, defined as the ability to perceive and resonate with the feelings of others [22], [23], is a critical aspect of any dialogue system. An empathetic dialogue system is designed to discern the emotional shifts in the user and formulate suitable responses imbued with corresponding sentiment [24]. This facet of dialogue systems is pivotal, as it directly influences the emotional experience of the user and, to a certain extent, dictates the quality of the response.

In this study, sentiment analysis is conducted using machine learning algorithms for Turkish question-answering systems with the aim of increasing emotional performance in human-robot interaction. Initially, text preprocessing steps (tokenization, removal of stop words, removal of punctuation marks and numbers, normalization, stemming) are applied to the original Turkish question-answer-sentiment dataset. Subsequently, the preprocessed question and answer texts are combined and transformed into a single vector using Pretrained Turkish BERT Model and two different word representation methods, TF-IDF and word2vec. Additionally, pre-determined polarity vectors containing positive and negative scores of words are added to the question-answer text vector. As a result, a new hybrid sentiment analysis model is proposed. Vectorized and expanded question-answer text vectors are separated into training and testing data. They are trained and tested with different machine learning (ML) algorithms. Precision, Recall, F1-Score and Accuracy values are used for comparison.

Contributions: This study provides several significant contributions to the field of Turkish sentiment analysis.

1) A new question-answer sentiment analysis model for human-robot interaction has been proposed using a unique Turkish question-answer data set. This model takes advantage of both current word representation methods (word2vec, TF-IDF and Pretrained Turkish BERT Model) and the performance of various classifiers. Furthermore, this study can provide a valuable foundation for other applications in the fields of Turkish natural language processing and human-robot interaction.
2) To increase the success rate of this model in sentiment analysis, a hybrid system has been created that adds polarity score values to word vectors. As a result, the accuracy rate of the model has risen up to 91.05% compared to existing systems.
3) In Turkish sentiment analysis studies, especially when compared with another Turkish sentiment analysis study using the same data sets, the proposed sentiment analysis model stands out with its unique data set and successful accuracy rate.

In the I section of the study, an introduction is provided, setting the foundation for the research. The II section consists of a thorough review of the literature and the presentation of related works in detail. The III section explains the methodology used in this study, including a detailed description of the proposed sentiment analysis model. In the IV section, we present our experimental studies conducted using different datasets and discuss their results. The V section is dedicated to an in-depth discussion on the findings, their implications, and the potential for future work. Finally, the VI section concludes the study, summarizing the key points and findings.

II. RELATED WORKS

Sentiment analysis is the process of determining the sentiment category and polarity of a text by systematically examining the semantic information contained in texts [25]. With the widespread use of the Internet and social networks, analyzing big data and converting it into meaningful information has become crucial. Sentiment analysis studies are frequently conducted in fields such as finance [26], marketing [27], social media analysis [28], [29], [30], [31] and question-answering systems [15], [16], [17]. In this context, numerous academic studies have been conducted using Machine Learning (ML) [32], [33], [34], dictionary-based methods [35], [36], and deep learning-based techniques [37], [38]. When examining the studies in Turkish, it is observed that there is a gap in sentiment analysis studies for question-answering systems. In this direction, the studies examined are presented in Table 1.

As shown in Table 1, Kaya et al. [39] applied sentiment classification techniques using political news obtained from Turkish news sites in their studies. The texts were preprocessed with the Zemberek framework [40] and vectorised with bag-of-words and N-gram methods. Machine learning

TABLE 1. Accuracy Table of Stated Sentiment Analysis Studies in the Turkish language And Sentiment Analysis with Questions Answering Studies in other languages.

Reference	Dataset	Vector Model	Classification	Results
[39]	Turkish Political News	BOW, N-gram	Naïve Bayes,SVM	77% Accuracy
[43]	Turkish Movie Reviews	BOW, TF-IDF	Naïve Bayes,SVM	83% F1-Score
[44]	Turkish Tweets	BOW, N-gram	SVM, NB, KNN	66% Accuracy
[45]	Turkish Tweets	Polarity Score with SentiwordNet	-	80% Accuracy
[46]	Product Reviews Rating	Polarity score with SentiTurkNet	SVM,NB, RF, KNN	74% F1-Score
[5]	Tweets on Global Warming	N-Gram	KNN, SVM, NB	74% Accuracy
[47]	The Turkish Tweets with Sentiment Symbols	Skip-Gram ve CBOW	SVM	80% Accuracy
[15]	QA Style Reviews in an Web Site	Skip-Gram	LSTM	86% Accuracy
[16]	QA Style Reviews in an Web Site	FastText and BERT Pretrained Vectors	CNN	88% Accuracy

algorithms such as Naive Bayes, Maximum Entropy, SVM, and the Character-Based N-Gram Language Model were used for sentiment classification, and a maximum accuracy value of 77% was achieved.

Akba et al. [43], pre-processed a text dataset of Turkish movie reviews using the Zemberek framework [40] and used Bag of Words and TF-IDF vector models. By conducting sentiment analysis with Naive Bayes and SVM classification models, they reached a maximum F1-Score value of 83.9

Coban et al. [44] prepared texts for sentiment analysis in their study using classic NLP preprocessing steps and Bag-of-Words and N-Gram vector representations in Turkish Twitter data. Machine learning algorithms such as SVM, Naive Bayes, Multinomial Naive Bayes, and KNN were used as classification models. A maximum accuracy value of 66.06% was achieved with the N-Gram vector representation and Multinomial Naive Bayes.

Karamollaoğlu et al. [45] preprocessed Turkish Twitter data with the help of the Zemberek framework [40] in their study and obtained Polarity Scores of Twitter texts using the Turkish SentiwordNet [41] library. By determining the negative, positive, or neutral values of the texts with the Polarity Score, an accuracy value of 80% was achieved.

Rumelli et al. [46], performed preprocessing with Zemberek [40] on a dataset consisting of product reviews and evaluations from an e-commerce site. With the help of the SentiTurkNet library [42], the polarity scores for each word of the texts were determined and displayed as a vector. Sentiment analysis was carried out using Naive Bayesian (NB), Random Forest (RF), SVM, and KNN classification models. The Naive Bayesian model showed the best performance with a 0.747 F1-Score.

Kirelli et al. [5], preprocessed a dataset obtained from Turkish tweets about global warming with the Zemberek framework and then used the N-Gram vector representation model. Sentiment analysis results were determined using K-NN, SVM, and NB classification models. The K-NN classifier showed the best performance with 74.63% accuracy.

Hayran et al. [47] in their study, preprocessed the Turkish Twitter data labeled according to the emotion symbol used in the content and cleaned it through NLP preprocessing steps. The data were then vectorised using Skip-Gram and CBOW models. Sentiment analysis was performed using the SVM classification model, and an accuracy value of 80.05% was achieved.

Human-like robots can use question-answering systems, and emotion classification is becoming applicable for many question-answering studies [48], [49]. Shen et al. [50] conducted an emotion classification study on question-answering datasets in three different categories, achieving a maximum accuracy value of 82%.

Yu et al. focused on the emotional meanings (positive-negative) of the keywords in questions and answers and performed classification accordingly [51].

Wang et al. [15] defined emotions in question-answers obtained from an e-commerce site in their study. Question-answers were used in two different ways: with their categories or with the terms they contained. Question-answers were vectorized with Skip-Gram and subsequently classified using an LSTM-based deep learning model for sentiment analysis. The best performance was achieved with an accuracy of 86%, according to the metric values shown in the results.

Zhang et al. [16] obtained a question-answering dataset from an e-commerce site in their study and defined the emotions of the terms in the dataset. Question-answers were vectorized using BERT and fasttext-based pre-trained models. Emotion classification was performed using CNN-based deep learning models, with the best performance achieved at 88% accuracy.

In conclusion, it can be seen that sentiment analysis studies are progressing with a focus on various datasets and languages, successfully applying machine learning, dictionary-based methods, and deep learning techniques. It is understood that more effort is needed in sentiment analysis studies for Turkish question-answering systems. The popularization of such studies will contribute to the development of language and culture-specific sentiment analysis techniques, providing significant support for making human-machine interaction more natural and effective.

III. METHODOLOGY

In this study, a new hybrid sentiment analysis model is developed to evaluate question-answer texts from an emotional perspective during human-robot interaction. Within the scope of the proposed model, our original question-answer sentiment dataset and sample datasets from the literature

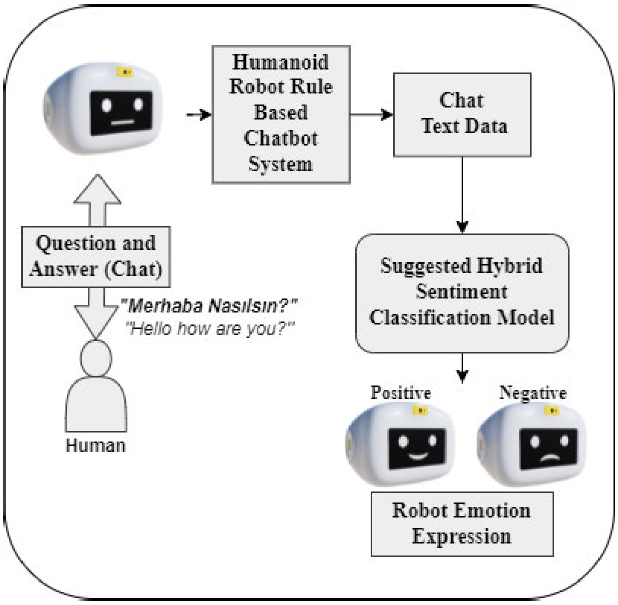


FIGURE 1. Datasets category ratios.

(Twitter sentiment datasets) are emotionally classified using various machine learning methods with the help of their sentiment labels. The aim is to accurately determine sentiment expressions during the robot's conversation process.

Figure 2 presents the general system architecture of our study. Accordingly, dialogue sentences consisting of 13,500 original and rule-based generated question-answer pairs were classified using our sentiment analysis model presented in Figure 4, obtaining sentiment expressions for the robot. As a result, users interacting with the humanoid robot experienced a more realistic and effective conversation with the sad or happy emotion expressions displayed on the robot's face animation screen.

A. DATASETS

The training of the proposed hybrid sentiment analysis model is carried out using the datasets shown in Table 2, and the human-robot-emotion model with the general architecture shown in Figure 2 is obtained.

Since there is no Turkish question-answer-sentiment dataset in the literature, our original question-answer dataset, which we named SCD, is compared with DS1 and SentimentSet, which are Twitter-sentiment datasets, through model performance metrics. Thus, it is demonstrated that our SCD dataset can be used in terms of sentiments.

1) SCD DATASET

This new dataset has been specifically created for the development and education of children with down syndrome. The dataset, containing a total of 13,500 Turkish question-answer pairs, has "positive" and "negative" emotion labels. In the context of human-robot interaction, accurately identifying and addressing positive and negative emotions has a significant impact on user experience and satisfaction. Neutral questions and answers provide less information in terms of sentiment analysis and are less relevant to the purpose of this study. Therefore, focusing on positive and negative emotions in our research was preferred in order to achieve more specific and effective results. In this context, situations of neutral questions and answers that approach positive or negative meanings have been appropriately manually corrected.

The dataset has been carefully and originally prepared, based on various topics that can be covered in conversations between robots and children. In this dataset, the question class can include any type of sentence, whether it is a single word or meant to maintain the continuity of the conversation. Similarly, the answer class is designed in the same manner. The entire set of question-answer pairs, created with 200 volunteer students from the Computer Engineering Department of Iskenderun Technical University, has been labeled with emotion tags and verified for accuracy by experts in the field. Table 3 presents a few example question-answer-emotion pairs from our dataset.

TABLE 2. Datasets information.

Dataset	Number of Data	Data Labels
SCD (Our Dialogue Dataset)	13.500	Positive - Negative Turkish QA Pairs
SentimentSet (Twitter)	2600	Positive - Negative Turkish Tweets
DS1 (Twitter)	11.119	Positive - Negative Turkish Tweets

2) DS1 DATASET

This dataset consists of tweets collected from Twitter and labeled according to positive or negative sentiment classes [53]. Containing a total of 11,119 Turkish tweets, this dataset can be used for sentiment analysis studies.

3) SENTIMENTSET DATASET

This dataset consists of 2,600 tweets collected from the Twitter platform and labeled for sentiment analysis purposes [54]. Containing both positive and negative sentiment, this dataset is considered a useful resource for Turkish sentiment analysis studies.

Figure 1 shows the ratio values for the label information of the datasets used in the study; these ratios represent the distribution of negative and positive sentiment categories in the dataset. This distribution is considered an important factor in assessing the extent to which algorithms and methods can distinguish different sentiment classes. A more balanced label distribution enables a more accurate and reliable measurement of classification performance. Therefore, taking the label distribution into account and using appropriate preprocessing techniques during the analysis process have a positive impact on the success of sentiment analysis studies.

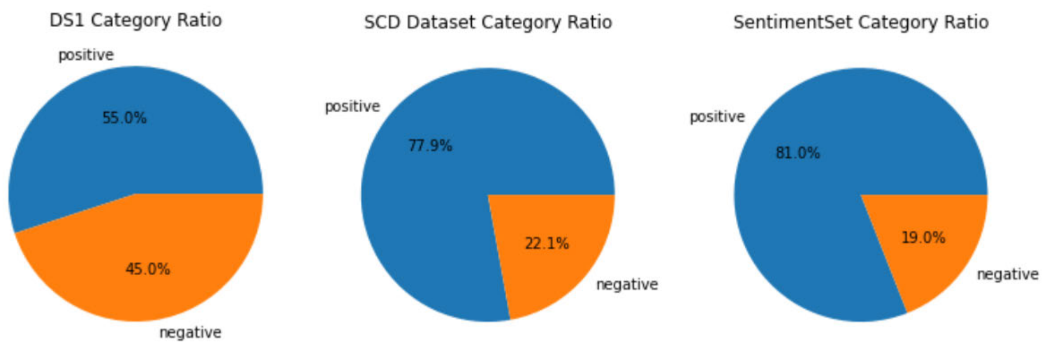


FIGURE 2. General architecture of human robot interaction.

TABLE 3. SCD dataset question-answer-sentiment examples.

Question	Answer	Sentiment
Hello, may I know your name? ("Merhaba, senin adını öğrenebilir miyim?")	Hi yes, my name is robot. ("Merhaba evet, benim adım robot.")	Positive (Olumlu)
Do you like to paint? ("Resim yapmayı sever misin?")	Yes, but I'm not very talented. ("Evet, ama pek yetenekli değilim.")	Positive (Olumlu)
Can you cook? ("Yemek yapabilir misin?")	I don't think I can. ("Yapabileceğimi sanmıyorum.")	Negative (Olumsuz)
How are you today? ("Bugün nasılsın?")	I'm very good, how are you? ("Çok iyiyim, sen nasılsın?")	Positive (Olumlu)
How are you? ("Nasılsın?")	I'm tired because I worked hard today. ("Yorgunum, çünkü bugün çok çalıştım.")	Negative (Olumsuz)

B. WORD REPRESENTATION METHODS

In this study, we used the word representation methods "word2vec", "TF-IDF", and "Pre-trained Turkish BERT" for vectorizing text data. The main function of these methods is to convert words or groups of words in texts into vector formats, enabling machine learning algorithms to process these data.

1) Word2Vec

Word2Vec is a popular method used for vectorizing text data. Known for its ability to capture semantic meanings of words, this model is used to determine similarities and relationships between words [59]. Word2Vec learns the context of a word by associating it with the words around it. As a result, words with similar contexts have similar vector representations [55].

2) TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) is a statistical method used to determine the importance of a word in a document [60]. TF-IDF finds the balance between the frequency of the word in the document (TF) and its rarity in the entire set of documents (IDF). This helps determine how important a word is within a specific document or text [56].

3) PRE-TRAINED TURKISH BERT

BERT (Bidirectional Encoder Representations from Transformers) is a groundbreaking model in deep learning-based NLP (Natural Language Processing) tasks [61]. BERT dissects the meaning of a word by looking at the context on both the right and the left of the word. This is a unique and powerful ability in language modeling. The Turkish BERT model used in this study has been pre-trained on a large Turkish text data set, hence it is capable of forming vector representations of Turkish texts [57].

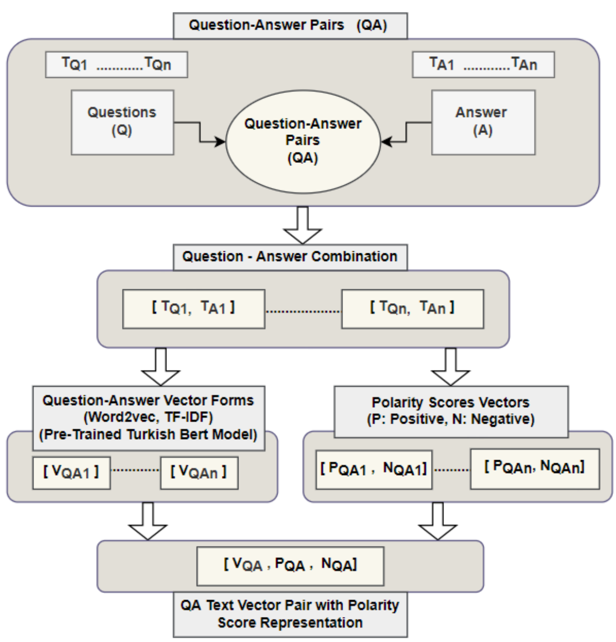


FIGURE 3. Question-answer vector formation model.

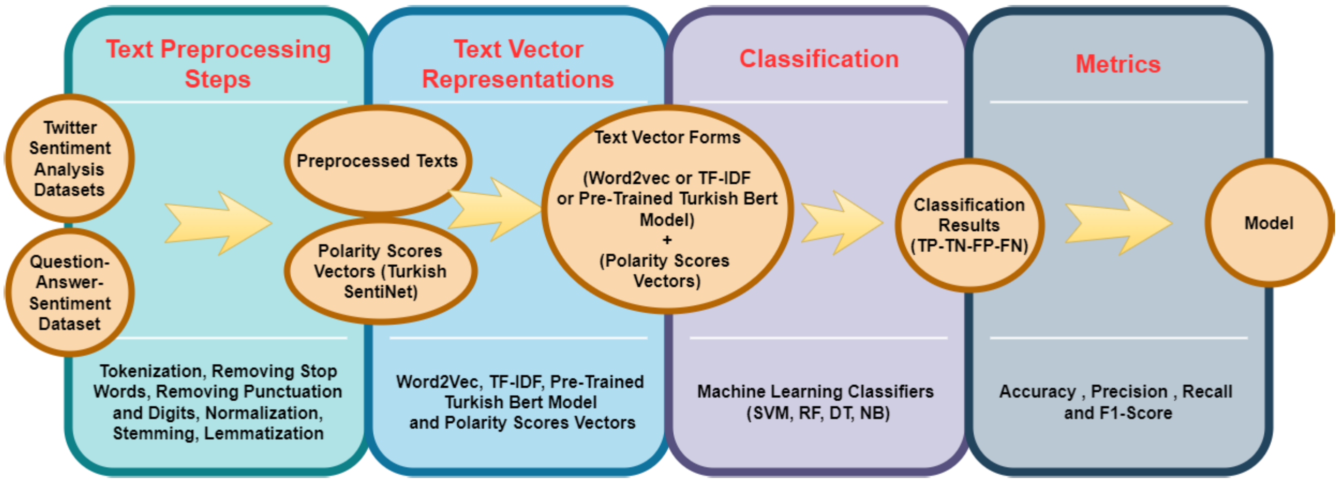


FIGURE 4. Suggested hybrid sentiment analysis model architecture.

Each of these representation methods has their unique strengths in text classification tasks such as text vectorization and sentiment analysis. Additionally, **polarity scores** representing the positive and negative emotions of texts were added to these text representations vectors [52] and **a hybrid vector form** was obtained (Figure 3).

C. MACHINE LEARNING MODELS FOR CLASSIFICATION

In our study on text classification tasks, we developed a hybrid model that benefits from both the success of different classifiers and the power of word representation methods. Four different machine learning algorithms were used: Linear Support Vector Machines (Linear SVM), Logistic Regression, Decision Trees, and Random Forest.

We began with Linear SVM, a powerful algorithm known for its robustness in handling high-dimensional feature spaces and resilience to noise and outliers. It operates by finding the optimal hyperplane that separates the data into distinct classes, proving to be an effective tool for classifying text into positive or negative categories based on emotional tone [58].

Building upon the foundation set by Linear SVM, we incorporated Logistic Regression into our methodological framework. Logistic Regression, despite its simplicity, demonstrates significant effectiveness in estimating the probability of an instance belonging to a particular class [59]. Thus, it complements Linear SVM by providing the ability to handle both binary and multiclass classification tasks.

Furthermore, we utilized Decision Trees in our research. Decision Trees work by recursively splitting the data into subsets based on the value of certain attributes. This algorithm's interpretability and versatility in handling both categorical and numerical data proved invaluable for classifying text based on its emotional tone [60]. Lastly, we employed Random Forest, an ensemble learning method known for its high accuracy. Random Forest combines multiple Decision Trees to improve the overall classification performance, showing exceptional capability in handling complex classification tasks [61].

In summary, we used these machine learning algorithms - Linear SVM, Logistic Regression, Decision Trees, and Random Forest - due to their proven effectiveness in handling classification tasks, their distinct strengths in dealing with various data characteristics, and their high accuracy in predicting outcomes. By comparing their performance in terms of sentiment analysis and classification accuracy, we were able to identify and select the most suitable model for our study.

D. SUGGESTED HYBRID SENTIMENT ANALYSIS MODEL

According to the hybrid sentiment analysis model presented in Figure 4, text datasets were first subjected to tokenization and text cleaning processes in the text pre-processing stages. Text cleaning operations such as correcting spelling errors, stemming words, expanding abbreviations, and cleaning special characters were applied to convert question-answer texts and texts obtained from social media into a standard language form. Zemberek and NLTK Snowball libraries were used for these operations.

Subsequently, these text data were transformed into vector form to be subjected to classification. In this process, word2vec and TF-IDF word representation methods and Pre-trained Turkish BERT Model were used to vectorize texts. Additionally, polarity scores representing the positive and negative emotions of texts were added to text vectors [52], and a hybrid vector form was obtained (Figure 3). Finally, our model was trained with various machine learning classification algorithms, and its performance was tested by determining performance metrics (60% training and 40% testing data).

Thanks to the developed model, the emotional evaluation of human-robot interaction is successfully carried out. Since the emotion labels in the datasets used during the training phase are divided into two categories as positive and negative, human-robot chat evaluation also represents two different emotions. A sad expression is used for negative emotions, while a happy expression is used for positive emotions.

E. PERFORMANCE MEASUREMENT

In our study, we used various metrics to evaluate the performance of our model. These metrics included accuracy, precision, recall, and the F1 score.

Accuracy represents the proportion of correctly classified instances out of the total number of instances. Precision measures how many of the instances predicted as positive are indeed positive. It is calculated by dividing the number of true positive predictions by the total of true positive and false positive predictions.

Recall, on the other hand, measures how many of the actual positive instances were correctly predicted. It is calculated by dividing the number of true positive predictions by the total of true positive and false negative predictions.

The F1 score is a measure of the overall performance of the model, calculated by taking the harmonic mean of precision and recall. This metric is particularly useful in classification problems where there is class imbalance or when both positive and negative prediction accuracy significantly affect the results.

These metrics were calculated for each combination of classifier and word representation method, allowing us to compare the performance of different models. We also used confusion matrices to examine true positive(TP), true negative(TN), false positive(FP) and false negative(FN) values, identifying areas where the model struggled.

In summary, these evaluation techniques helped us identify the most suitable model for our study. By comparing the performance of each model based on different metrics, we were able to select the model that demonstrated the highest level of accuracy and consistency in classifying text based on emotional tone.

IV. EXPERIMENTAL STUDIES AND RESULTS

Within the scope of our study, it is aimed to evaluate the dialogue system created with our original Turkish question-answer dataset in terms of sentiment, compare system performances using other Turkish sentiment datasets, develop a hybrid sentiment analysis model, and investigate the effect of adding polarity scores to text vectors on the model.

In the Figures 5, 6 the confusion matrices of the results with the highest values obtained in the sentiment analysis of the study are shown (TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) values).

All performance metrics obtained for the two different situations, where polarity score values are not added and added to the text vectors, are presented in Table 4 and Table 5, respectively. Considering that "word2vec", "TF-IDF" and Pretrained Turkish BERT representation models are used in both tables, Table 4 shows that the highest accuracy value of 90.95% was reached on the SCD dataset, while Table 5, where polarity scores of the texts were added, indicates that the highest accuracy value of 91.05% was achieved.

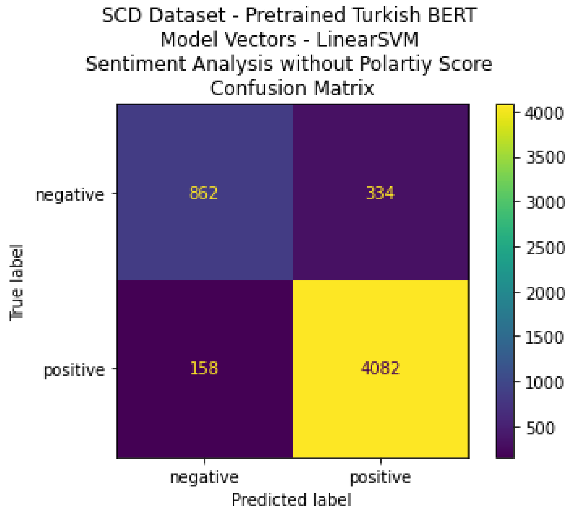


FIGURE 5. SCD dataset - pretrained turkish BERT model vectors - LinearSVM sentiment analysis without polartiy score confusion matrix.

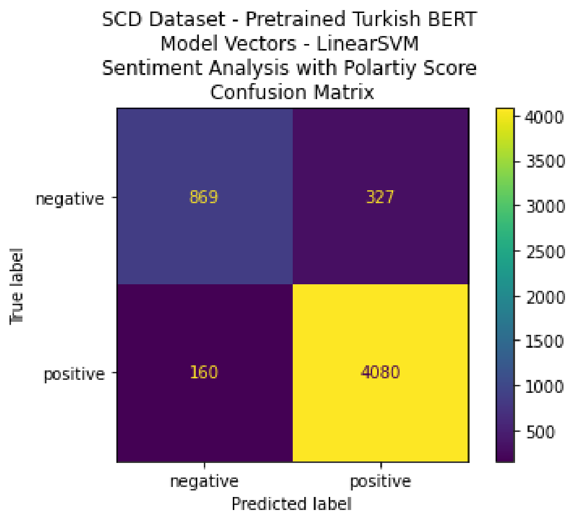


FIGURE 6. SCD dataset - pretrained turkish BERT model vectors - LinearSVM sentiment analysis with polartiy score confusion matrix.

As seen in Table 6, in light of these results, it is observed that the proposed hybrid sentiment analysis model achieves an average improvement of 2.40% using word2vec, 0.11% using TF-IDF and 1.39% using Pretrained Turkish BERT on the SCD dataset, an average of 0.35% using word2vec, 0.51% using TF-IDF and 0.75% using Pretrained Turkish BERT on the SentimentSet dataset, and an average of 6.74% using word2vec, 2.39% using TF-IDF and 1.31% using Pretrained Turkish BERT on the DS1 dataset. These results suggest that incorporating polarity scores into the model may further enhance sentiment analysis performance.

Upon examining Table 6, we see how the use of Word2vec, TF-IDF, and Pretrained Turkish BERT Model Vectors

TABLE 4. Sentiment analysis performance metrics comparison using Word2vec, TF-IDF and pretrained turkish BERT model vectors representation model without polarity score.

Datasets	Classification Algorithms	Using Word2vec Vectors without Polarity Scores Vectors				Using TF-IDF Vectors without Polarity Scores Vectors				Using Pretrained Turkish BERT Model Vectors without Polarity Scores Vectors			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
SCD	Log.Regression	83.17	84.20	82.51	84.20	88.43	88.36	88,37	**88.86**	88.54	88.69	87.79	88.69
	RandomForest	82.45	83.67	81.97	83.67	88.39	88.23	88.49	**88.73**	87.28	87.16	85.66	87.16
	LinearSVM	86.51	86.98	85.86	**86.98**	88.55	88.47	88.52	**88.97**	90.70	90.95	90.69	**90.95**
	DecisionTree	77.40	76.72	77.03	76.72	86.80	86.22	86.76	86.72	79.13	78.78	78.94	78.78
SentimentSet	Log.Regression	66.09	81.30	72.91	81.30	82.89	84.63	82.56	**84.63**	83.92	82.77	76.52	82.77
	RandomForest	83.62	84.24	80.24	**84.24**	82.78	84.14	83.18	**84.14**	87.29	86.88	84.32	86.88
	LinearSVM	78.91	81.59	74.12	81.59	82.56	84.04	82.95	**84.04**	87.64	87.47	85.32	**87.47**
	DecisionTree	76.44	76.20	76.32	76.20	81.67	81.59	81.63	81.59	78.43	77.38	77.86	77.38
DS1	Log.Regression	69.88	66.61	63.48	66.61	73.84	71.96	70.57	**71.96**	83.92	83.77	83.65	83.77
	RandomForest	65.65	65.76	65.06	65.76	72.54	71.40	70.29	**71.40**	81.51	81.07	80.81	81.07
	LinearSVM	67.68	67.17	65.94	**67.17**	73.85	71.94	70.53	**71.94**	85.65	85.64	85.59	**85.64**
	DecisionTree	55.00	54.94	54.97	54.94	68.75	68.55	67.76	68.55	64.28	64.12	64.17	64.12

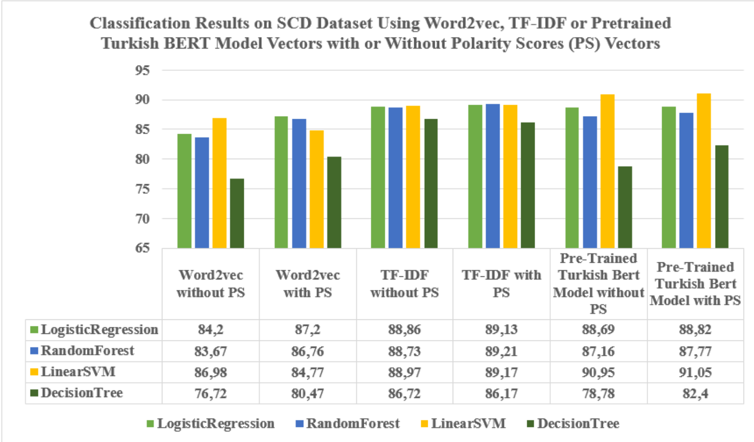
TABLE 5. Sentiment analysis performance metrics comparison using Word2vec, TF-IDF and pretrained turkish BERT model vectors representation model with polarity score.

Datasets	Classification Algorithms	Using Word2vec Vectors with Polarity Scores Vectors				Using TF-IDF Vectors with Polarity Scores Vectors				Using Pretrained Turkish BERT Model Vectors with Polarity Scores Vectors			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
SCD	Log.Regression	86.60	86.70	86.43	**87.20**	88.73	88.63	88.74	**89.13**	88.41	88.82	88.23	88.82
	RandomForest	86.12	86.26	85.84	86.76	88.97	88.71	89.06	**89.21**	87.69	87.77	86.57	87.77
	LinearSVM	86.14	84.27	85.26	84.77	88.78	88.67	88.79	**89.17**	90.80	91.05	90.8	**91.05**
	DecisionTree	80.87	79.97	80.65	80.47	86.32	85.67	86.24	86.17	82.57	82.40	82.48	82.40
SentimentSet	Log.Regression	73.74	81.10	73.52	81.10	82.84	84.53	82.95	84.53	84.68	83.65	78.41	83.65
	RandomForest	83.17	84.14	80.26	**84.14**	84.80	86.00	84.92	**86.00**	87.05	86.78	84.23	86.78
	LinearSVM	80.01	81.98	75.30	81.98	83.37	84.63	83.72	84.63	88.04	88.06	86.28	**88.06**
	DecisionTree	77.32	77.18	77.25	77.18	81.24	81.00	81.12	81.00	78.63	78.46	78.54	78.46
DS1	Log.Regression	70.31	69.51	68.38	69.51	75.37	74.55	73.86	**74.55**	83.78	83.68	83.57	83.68
	RandomForest	69.69	69.74	69.35	69.74	72.48	72.39	72.00	72.39	82.66	82.33	82.13	82.33
	LinearSVM	71.49	71.04	70.29	**71.04**	75.80	74.64	73.81	**74.64**	85.94	85.95	85.92	**85.95**
	DecisionTree	60.93	60.88	60.90	60.88	68.97	69.11	68.89	69.11	66.36	66.32	66.34	66.32

TABLE 6. Improvement effect of using Word2vec, TF-IDF and pretrained turkish BERT model vectors with or without polarity scores on model accuracy.

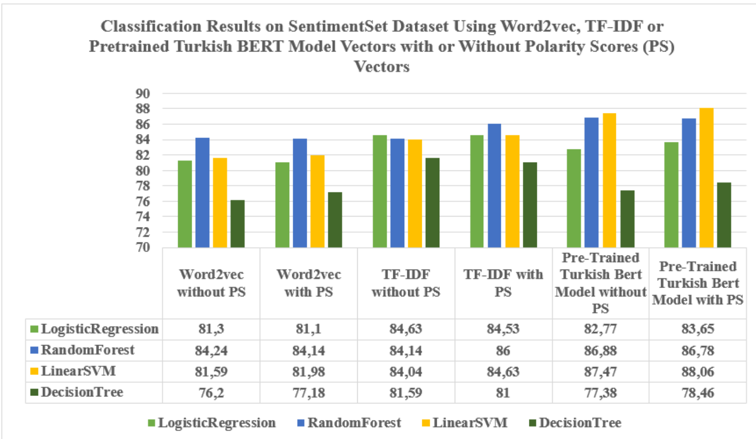
Datasets	Classification Algorithms	Improvement Effect of Using Word2vec Vectors with or Without Polarity Scores on Model Accuracy			Improvement Effect of Using TF-IDF Vectors with or Without Polarity Scores on Model Accuracy			Improvement Effect of Using Pretrained Turkish BERT Model Vectors with or Without Polarity Scores on Model Accuracy		
		Without Polarity	With Polarity	Rate (%)	Without Polarity	With Polarity	Rate (%)	Without Polarity	With Polarity	Rate (%)
SCD	LogisticRegression	84.20	87.20	**3.57**	88.86	89.13	**0.31**	88.69	88.82	**0.15**
	RandomForest	83.67	86.76	**3.70**	88.73	89.21	**0.55**	87.16	87.77	**0.70**
	LinearSVM	86.98	84.77	**-2.55**	88.97	89.17	**0.23**	90.95	91.05	**0.11**
	DecisionTree	76.72	80.47	**4.89**	86.72	86.17	**-0.64**	78.78	82.40	**4.60**
SentimentSet	LogisticRegression	81.30	81.10	**-0.25**	84.63	84.53	**-0.12**	82.77	83.65	**1.07**
	RandomForest	84.24	84.14	**-0.12**	84.14	86.00	**2.22**	86.88	86.78	**-0.12**
	LinearSVM	81.59	81.98	**0.48**	84.04	84.63	**0.71**	87.47	88.06	**0.68**
	DecisionTree	76.20	77.18	**1.29**	81.59	81.00	**-0.73**	77.38	78.46	**1.40**
DS1	LogisticRegression	66.61	69.51	**4.36**	71.96	74.55	**3.60**	83.77	83.68	**-0.11**
	RandomForest	65.76	69.74	**6.06**	71.40	72.39	**1.39**	81.07	82.33	**1.56**
	LinearSVM	67.17	71.04	**5.77**	71.94	74.64	**3.76**	85.64	85.95	**0.37**
	DecisionTree	54.94	60.88	**10.82**	68.55	69.11	**0.82**	64.12	66.32	**3.44**

impacts classification accuracy. From the experiments conducted on various datasets and with different classification algorithms, it is observed that the use of polarity values generally enhances the model's accuracy. Looking at the "rate" values in the table, we can see that the use of polarity generally increases the accuracy rate. Considering that



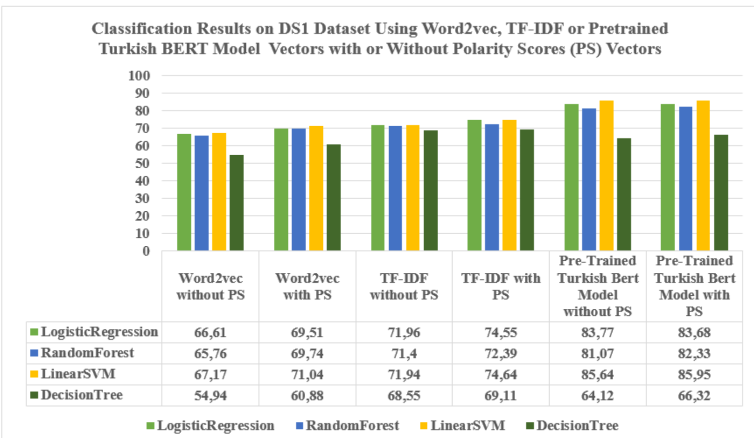
	Word2vec without PS	Word2vec with PS	TF-IDF without PS	TF-IDF with PS	Pre-Trained Turkish Bert Model without PS	Pre-Trained Turkish Bert Model with PS
LogisticRegression	84,2	87,2	88,86	89,13	88,69	88,82
RandomForest	83,67	86,76	88,73	89,21	87,16	87,77
LinearSVM	86,98	84,77	88,97	89,17	90,95	91,05
DecisionTree	76,72	80,47	86,72	86,17	78,78	82,4

FIGURE 7. Sentiment analysis accuracy results on SCD dataset using Word2vec or TF-IDF or pretrained turkish BERT model vectors with or without polarity scores (PS) vectors.



	Word2vec without PS	Word2vec with PS	TF-IDF without PS	TF-IDF with PS	Pre-Trained Turkish Bert Model without PS	Pre-Trained Turkish Bert Model with PS
LogisticRegression	81,3	81,1	84,63	84,53	82,77	83,65
RandomForest	84,24	84,14	84,14	86	86,88	86,78
LinearSVM	81,59	81,98	84,04	84,63	87,47	88,06
DecisionTree	76,2	77,18	81,59	81	77,38	78,46

FIGURE 8. Sentiment analysis accuracy results on sentimentset dataset using Word2vec or TF-IDF or pretrained turkish BERT model vectors with or without polarity scores (PS) vectors.



	Word2vec without PS	Word2vec with PS	TF-IDF without PS	TF-IDF with PS	Pre-Trained Turkish Bert Model without PS	Pre-Trained Turkish Bert Model with PS
LogisticRegression	66,61	69,51	71,96	74,55	83,77	83,68
RandomForest	65,76	69,74	71,4	72,39	81,07	82,33
LinearSVM	67,17	71,04	71,94	74,64	85,64	85,95
DecisionTree	54,94	60,88	68,55	69,11	64,12	66,32

FIGURE 9. Sentiment analysis accuracy results on DS1 dataset using Word2vec or TF-IDF vectors or pretrained turkish BERT model with or without polarity scores (PS) vectors.

approximately 80% of these rates are positive, we can say that polarity values have a decisive effect on the success of the model.

In Figure 7, 8, 9 the accuracy values of the proposed sentiment analysis model on different datasets are presented in graphs using 4 different machine learning algorithms.

Moreover, situations where polarity score values are added and not added to the text vectors are shown separately on the graph, allowing for comparison.

When these results are examined, compared to the similar studies in the literature with examples shown in Table 1, it is observed that our proposed hybrid sentiment analysis model demonstrates high performance. When examining the limited number of studies on the evaluation of question answering systems in terms of emotions, it was found that there is no such study in Turkish. In this context, performance success comparisons were made in comparison to emotion classification studies in question-answer systems conducted in other languages. In these studies, accuracy values up to a maximum of 88% have been reached. When examining Turkish sentiment analysis studies, comparisons could be made with classification studies conducted on subjects such as political news, movie reviews, social media texts, and e-commerce user reviews. In these studies, accuracy values up to a maximum of 83.9% have been reached. Especially when compared to another Turkish sentiment analysis study that used the same datasets, our study stands out with the proposed hybrid sentiment analysis model, unique dataset, and impressive accuracy rate of 91.05%. This indicates that our study delivers more potent and effective results [54].

V. DISCUSSION

In this study, a new hybrid model for Turkish sentiment analysis has been proposed, utilizing a unique question-answer dataset. The study employs word representation methods such as word2vec, TF-IDF, and the Pretrained Turkish BERT Model, demonstrating their successful application in the sentiment analysis of Turkish texts.

The proposed model's high accuracy scores are a result of the combination of these word representation methods and various classifiers. The results indicate that the model's accuracy can be further enhanced by adding polarity score values to word vectors.

These findings underscore the effectiveness and success of the proposed hybrid model for sentiment analysis in Turkish question-answer systems. This is particularly evident when compared to another Turkish sentiment analysis study that employs the same datasets.

This study is believed to lay a valuable foundation for other applications in the fields of human-robot interaction and natural language processing. It is also noted that future studies could further increase the success rate using more comprehensive datasets and improved algorithms.

Despite certain limitations, this study forms a significant basis. One such limitation is that our unique dataset prepared only addresses the positive and negative emotional aspects in human-robot interactions. However, emotional expression is not limited to these two opposing poles and can encompass a much broader spectrum. Therefore, future research has the opportunity to explore this broader range of emotional expression (including emotional states such as angry, surprised, neutral, etc.) and further enhance the realism and comprehensiveness of emotional analysis. Additionally, in future studies, ablative analyses will be conducted to determine which components contribute more to the performance of our model.

This study aims to contribute to technological advancements in special education by taking the complex interactions between emotional state analysis and Human-Robot Interaction as a foundation, with the goal of expanding the boundaries of this field. Consequently, to exemplify the application areas of this study, qualitative examples are provided below:

An In-depth Examination of the Relationship between Human-Robot: To illustrate the function of the Human-Robot Interaction (HRI) segment of our research, consider a scenario in which a robot engages with a child who communicates, "I experienced immense joy at school today!" The HRI component discerns the expression of joy within the child's statement and transmits this vital data to other modules. Consequently, this information helps in generating related queries based on joy, thereby granting a comprehensive understanding of the emotional condition.

Augmentation of Emotion Recognition through Interactive Dialogue: The role of our inquiry-response module is to present relevant questions to the child, in order to delve deeper into their emotional state. For instance, interrogating about the day's activities or inquiring about the child's favorite school event. The responses gleaned provide additional information about the emotional state, thus enhancing the precision of emotion recognition.

Compilation of Emotion Classification Findings: The emotion categorization module makes use of the gathered data to deduce the child's emotional state. For instance, having learnt that the child enjoyed school and appreciated specific activities, this module can categorize the child as experiencing joy. Without the proposed model, the analysis of the child's expression of joy would be solely based on the statement "I had so much fun at school today!" However, with the application of the aforementioned framework, a more detailed understanding of the child's emotional state can be achieved, along with insights into the factors contributing to their joy.

VI. CONCLUSION

Within the scope of this study, a new hybrid sentiment analysis model for use in human-robot interaction has been proposed using a unique Turkish question-answer dataset and sample datasets with machine learning classifiers. The primary aim of the study is to achieve high accuracy values in sentiment analysis, making human-robot interactions more realistic and empathetic. In this context, significant emphasis has been placed on the performance and applicability of the proposed model.

The word2vec, TF-IDF and Pretrained Turkish BERT Model word representation methods used in the study have been successfully employed in the sentiment analysis of Turkish texts, resulting in richer and more meaningful vectors compared to existing systems. Additionally, by adding polarity score values to the word vectors, a hybrid system was created, further enhancing the accuracy of the model.

Machine learning algorithms such as Linear SVM, Logistic Regression, Decision Tree, and Random Forest have been used for sentiment analysis classification. Each of these algorithms has been compared in terms of sentiment analysis and classification performance and the most suitable model has been selected. The hybrid model, benefiting from both the power of word representation methods and the success of different classifiers, has achieved an accuracy value of up to 91.05% in sentiment analysis.

These results demonstrate that the proposed hybrid model is an effective and successful method for sentiment analysis in Turkish question-answering systems. Furthermore, this study can serve as a valuable foundation for other applications in the fields of human-robot interaction and natural language processing.

Our study's findings indicate a positive advancement in sentiment analysis within the context of Turkish question-answer systems. However, as inherent in any research, this study encapsulates certain limitations and presents opportunities for future research.

Utilization of Expanded Datasets: Natural Language Processing (NLP) and sentiment analysis modeling is built on the requirement of structured and meaningful data. The success of this study correlates directly with the quality and volume of the datasets we utilized. However, the unique dataset used in our study accounts only for positive and negative emotional responses in human-robot interactions, neglecting a broad spectrum of emotional expression such as anger, surprise, or neutrality. Future research can enhance the inclusivity and realism of sentiment analysis by using more diverse Turkish text-enriched datasets and addressing a wider spectrum of emotional expression.

Using Advanced Natural Language Processing Techniques: With the rapid advancement of language models and natural language processing techniques, there exists the potential to further enhance the sentiment analysis performance of the model. Integration of these new techniques and approaches could provide the opportunity to refine sentiment analysis results further.

Multi-Language Support: Expanding the model's capability to perform sentiment analysis on Turkish texts and support other languages can both broaden the application scope of the model and allow for the examination of differences in sentiment analysis performance across different languages and cultures.

Real-Time Applications: The model proposed in this study provides a framework for real-time sentiment analysis. This brings to the fore the potential for automating sentiment analysis in specific scenarios such as the use of robots in the education of children with special needs. Strategic focus on these areas could deepen the findings of our study and aid in the development of more complex and effective sentiment analysis solutions.

REFERENCES

[1] S. H.-W. Chuah and J. Yu, "The future of service: The power of emotion in human-robot interaction," *J. Retailing Consum. Services*, vol. 61, Jul. 2021, Art. no. 102551.

[2] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl.-Based Syst.*, vol. 226, Aug. 2021, Art. no. 107134.

[3] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowl.-Based Syst.*, vol. 89, pp. 14–46, Nov. 2015.

[4] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Netw. Anal. Mining*, vol. 11, no. 1, p. 81, Dec. 2021.

[5] Y. Kirelli and S. Arslankaya, "Sentiment analysis of shared tweets on global warming on Twitter with data mining methods: A case study on Turkish language," *Comput. Intell. Neurosci.*, vol. 2020, pp. 1–9, Sep. 2020.

[6] H. A. Shehu, Md. H. Sharif, Md. H. U. Sharif, R. Datta, S. Tokat, S. Uyaver, H. Kusetogullari, and R. A. Ramadan, "Deep sentiment analysis: A case study on stemmed Turkish Twitter data," *IEEE Access*, vol. 9, pp. 56836–56854, 2021.

[7] A. T. Kabakus, "A novel COVID-19 sentiment analysis in Turkish based on the combination of convolutional neural network and bidirectional long-short term memory on Twitter," *Concurrency Computation: Pract. Exp.*, vol. 34, no. 22, Oct. 2022, Art. no. e6883.

[8] B. Erşahin, Ö. Aktaş, D. Kilinç, and M. Erşahin, "A hybrid sentiment analysis method for Turkish," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 27, no. 3, pp. 1780–1793, May 2019.

[9] H. A. Shehu and S. Tokat, "A hybrid approach for the sentiment analysis of Turkish Twitter data," in *Artificial Intelligence and Applied Mathematics in Engineering Problems*. Springer, 2020, pp. 182–190.

[10] Ö. Çoban, S. A. Özel, and A. Inan, "Deep learning-based sentiment analysis of Facebook data: The case of Turkish users," *Comput. J.*, vol. 64, no. 1, pp. 473–499, Nov. 2019.

[11] B. Ciftci and M. S. Apaydin, "A deep learning approach to sentiment analysis in Turkish," in *Proc. Int. Conf. Artif. Intell. Data Process. (IDAP)*, Sep. 2018, pp. 1–5.

[12] M. A. Toçoğlu and A. Alpkoçak, "Lexicon-based emotion analysis in Turkish," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 27, pp. 1213–1227, Mar. 2019.

[13] G. Yurtalan, M. Koyuncu, and Ç. Turhan, "A polarity calculation approach for lexicon-based Turkish sentiment analysis," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 27, pp. 1325–1339, Mar. 2019.

[14] P. K. Roy, S. Saumya, J. P. Singh, S. Banerjee, and A. Gutub, "Analysis of community question-answering issues via machine learning and deep learning: State-of-the-art review," *CAAI Trans. Intell. Technol.*, vol. 8, no. 1, pp. 95–117, Mar. 2023.

[15] J. Wang, C. Sun, S. Li, X. Liu, L. Si, M. Zhang, and G. Zhou, "Aspect sentiment classification towards question-answering with reinforced bidirectional attention network," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics*, 2019, pp. 3548–3557.

[16] W. Zhang, Y. Deng, X. Li, L. Bing, and W. Lam, "Aspect-based sentiment analysis in question answering forums," in *Proc. Findings Assoc. Comput. Linguistics (EMNLP)*, 2021, pp. 4582–4591.

[17] G. Hu, C. Shi, S. Hao, and Y. Bai, "Residual-duet network with tree dependency representation for Chinese question-answering sentiment analysis," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2020, pp. 1725–1728.

[18] H.-H. Nguyen and M.-T. Nguyen, "Emotion-cause pair extraction as question answering," 2023, *arXiv:2301.01982*.

[19] J. Ni, T. Young, V. Pandelea, F. Xue, and E. Cambria, "Recent advances in deep learning based dialogue systems: A systematic survey, *Artif. Intell. Rev.*, vol. 56, no. 4, pp. 3055–3155, 2023.

[20] D. M. Gezgin and C. Mıhcı, "Robotıc technology in specıal educatıon: Practıces, challenges and ethıcal consıderatıons," in *Proc. 4th Int. Istanbul Current Sci. Res. Congr.*, May 2023, pp. 214–222.

[21] Y. Karal, D. Tasdemir, and S. Öngöz, "Özel gereksinimli bireylerin egitiminde bir ögrenme materyali olarak robotlar," *Erzincan Üniversitesi Egitim Fakültesi Dergisi*, vol. 25, no. 1, pp. 170–183, 2023.

[22] Y. Ma, K. L. Nguyen, F. Z. Xing, and E. Cambria, "A survey on empathetic dialogue systems," *Inf. Fusion*, vol. 64, pp. 50–70, Dec. 2020.
[23] W. Li, W. Shao, S. Ji, and E. Cambria, "BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis," *Neurocomputing*, vol. 467, pp. 73–82, Jan. 2022.
[24] G. Tu, J. Wen, C. Liu, D. Jiang, and E. Cambria, "Context- and sentiment-aware networks for emotion recognition in conversation," *IEEE Trans. Artif. Intell.*, vol. 3, no. 5, pp. 699–708, Oct. 2022.
[25] S. Tokcaer, "Türkçe metinlerde duygu analizi," *J. Yaşar Univ.*, vol. 16, no. 63, pp. 1516–1536, Sep. 2021.
[26] K. Mishev, A. Gjorgjevikj, I. Vodenska, L. T. Chitkushev, and D. Trajanov, "Evaluation of sentiment analysis in finance: From lexicons to transformers," *IEEE Access*, vol. 8, pp. 131662–131682, 2020.
[27] C. Aguwa, M. H. Olya, and L. Monplaisir, "Modeling of fuzzy-based voice of customer for business decision analytics," *Knowl.-Based Syst.*, vol. 125, pp. 136–145, Jun. 2017.
[28] A. Ceron, L. Curini, S. M. Iacus, and G. Porro, "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 Soc.*, vol. 16, no. 2, pp. 340–358, Mar. 2014.
[29] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowl. Inf. Syst.*, vol. 60, no. 2, pp. 617–663, Jul. 2018.
[30] F. Neri, C. Aliprandi, F. Capeci, M. Cuadros, and T. By, "Sentiment analysis on social media," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2012, pp. 919–926.
[31] S. Vashishtha and S. Susan, "Fuzzy rule based unsupervised sentiment analysis from social media posts," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112834.
[32] A. Alslaity and R. Orji, "Machine learning techniques for emotion detection and sentiment analysis: Current state, challenges, and future directions," *Behav. Inf. Technol.*, vol. 41, pp. 1–26, Dec. 2022.
[33] M. A. Jassim, D. H. Abd, and M. N. Omri, "A survey of sentiment analysis from film critics based on machine learning, lexicon and hybridization," *Neural Comput. Appl.*, vol. 35, no. 13, pp. 9437–9461, May 2023.
[34] P. K. Jain, R. Pamula, and G. Srivastava, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Comput. Sci. Rev.*, vol. 41, Aug. 2021, Art. no. 100413.
[35] M. Darwich, S. A. M. Noah, N. Omar, and N. A. Osman, "Corpus-based techniques for sentiment lexicon generation: A review," *J. Digit. Inf. Manage.*, vol. 17, no. 5, p. 296, Oct. 2019.
[36] P. Ray, "Document level sentiment analysis for product review using dictionary based approach," *Int. J. Comput. Sci. Eng.*, vol. 4, no. 6, pp. 24–29, Jun. 2017.
[37] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: A review," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4335–4385, Aug. 2020.
[38] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. A. Chenaghlu, and J. Gao, "Deep learning–based text classification," *ACM Comput. Surv.*, vol. 54, no. 3, pp. 1–40, Apr. 2021.
[39] M. Kaya, G. Fidan, and I. H. Toroslu, "Sentiment analysis of Turkish political news," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, vol. 1, Dec. 2012, pp. 174–180.
[40] A. A. Akin and M. D. Akın, "Zemberek, an open source NLP framework for Turkic languages," *Structure*, vol. 10, pp. 1–5, Jun. 2007.
[41] F. Sebastiani and A. Esuli, "SentiWordNet: A publicly available lexical resource for opinion mining," in *Proc. 5th Int. Conf. Lang. Resour. Eval., Eur. Lang. Resour. Assoc. (ELRA)*, Genoa, Italy, 2006, pp. 417–422.
[42] R. Dehkharghani, Y. Saygin, B. Yanikoglu, and K. Oflazer, "SentiTurkNet: A Turkish polarity lexicon for sentiment analysis," *Lang. Resour. Eval.*, vol. 50, no. 3, pp. 667–685, Sep. 2016.
[43] F. Akba, "Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews," in *Proc. 8th Eur. Conf. Data Mining*, 2014, pp. 180–184.
[44] Ö. Çoban, B. Özyer, and G. T. Özyer, "Sentiment analysis for Turkish Twitter feeds," in *Proc. 23nd Signal Process. Commun. Appl. Conf. (SIU)*, May 2015, pp. 2388–2391.
[45] H. Karamollaoglu, I. A. Dogru, M. Dörterler, A. Utku, and O. Yildiz, "Sentiment analysis on Turkish social media shares through lexicon based approach," in *Proc. 3rd Int. Conf. Comput. Sci. Eng. (UBMK)*, Sep. 2018, pp. 45–49.
[46] M. Rumelli, D. Akkus, Ö. Kart, and Z. Isik, "Sentiment analysis in Turkish text with machine learning algorithms," in *Proc. Innov. Intell. Syst. Appl. Conf. (ASYU)*, Oct. 2019, pp. 1–5.
[47] A. Hayran and M. Sert, "Sentiment analysis on microblog data based on word embedding and fusion techniques," in *Proc. 25th Signal Process. Commun. Appl. Conf. (SIU)*, May 2017, pp. 1–4.
[48] M. Atzeni and D. R. Recupero, "Deep learning and sentiment analysis for human–robot interaction," in *Proc. Eur. Semantic Web Conf.*, in Lecture Notes in Computer Science, 2018, pp. 14–18.
[49] M. Szabóová, M. Sarnovský, V. M. Kresnáková, and K. Machová, "Emotion analysis in human–robot interaction," *Electronics*, vol. 9, no. 11, p. 1761, 2020.
[50] C. Shen, C. Sun, J. Wang, Y. Kang, S. Li, X. Liu, L. Si, M. Zhang, and G. Zhou, "Sentiment classification towards question-answering with hierarchical matching network," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 3654–3663.
[51] G. Yu, J. Li, L. Luo, Y. Meng, X. Ao, and Q. He, "Self question-answering: Aspect-based sentiment analysis by role flipped machine reading comprehension," in *Proc. Findings Assoc. Comput. Linguistics (EMNLP)*, 2021, pp. 1331–1342.
[52] M. Özçelik, B. N. Arican, Ö. Bakay, E. Sarmiş, Ö. Ergelen, N. G. Bayezit, and O. T. Yıldız, "HisNet: A polarity lexicon based on WordNet for emotion analysis," in *Proc. 11th Global Wordnet Conf.*, Jan. 2021, pp. 157–165.
[53] M. Beyaz. (Oct. 23, 2020). Türkçe sosyal medya paylasimi veri seti. Kaggle. Accessed: Mar. 25, 2023. [Online]. Available: https://www.kaggle.com/datasets/mrtbeyz/trke-sosyal-medya-paylam-veri-seti
[54] C. Balli, M. S. Guzel, E. Bostanci, and A. Mishra, "Sentimental analysis of Twitter users from Turkish content with natural language processing," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–17, Apr. 2022.
[55] T. Mikolov, I. C. K. Sutskever, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 1–9.
[56] J. Ramos, "Using TF-IDF to determine word relevance in document queries," in *Proc. 1st Instruct. Conf. Mach. Learn.*, Dec. 2003, vol. 242, no. 1, pp. 29–48.
[57] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.
[58] M. A. Fauzi, "Word2Vec model for sentiment analysis of product reviews in Indonesian language," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 1, p. 525, Feb. 2019, doi: 10.11591/ijece.v9i1.pp525-530.
[59] P. Upadhyay, S. Saifi, R. Rani, A. Sharma, and P. Bansal, "Machine learning-based sentiment analysis for the social media platforms," in *Proc. 6th Int. Conf. Inf. Syst. Comput. Netw. (ISCON)*, Mar. 2023, pp. 1–5, doi: 10.1109/iscon57294.2023.10112120.
[60] M. Rathi, A. Malik, D. Varshney, R. Sharma, and S. Mendiratta, "Sentiment analysis of tweets using machine learning approach," in *Proc. 11th Int. Conf. Contemp. Comput. (IC3)*, Aug. 2018, pp. 1–3, doi: 10.1109/ic3.2018.8530517.
[61] M. Guia, R. Silva, and J. Bernardino, "Comparison of Naïve Bayes, support vector machine, decision trees and random forest on sentiment analysis," in *Proc. 11th Int. Joint Conf. Knowl. Discovery, Knowl. Eng. Knowl. Manage.*, 2019, pp. 525–531, doi: 10.5220/0008364105250531.

KADIR TOHMA was born in Antakya, Hatay, Turkey, in 1990. He is currently pursuing the Ph.D. degree with the Department of Electrical and Electronics Engineering, Iskenderun Technical University. He is a Research Assistant with the Department of Computer Engineering, Iskenderun Technical University. His current research interests include wireless sensor networks, natural language processing, humanoid robots, and sentiment analysis. He has numerous publications in these areas and continuously strives to contribute to these fields of study.

HALIL IBRAHIM OKUR was born in Iskenderun, Hatay, Turkey, in 1991. He received the B.S. degree in computer engineering from Cukurova University, Adana, Turkey, in 2014, and the M.S. degree in computer engineering from Istanbul University-Cerrahpasa, Istanbul, Turkey, in 2018, where he is currently pursuing the Ph.D. degree. Since 2014, he has been a Research Assistant with Iskenderun Technical University. Since 2015, he has been continuing his academic studies under the supervision of Prof. Dr. Ahmet Sertbas in computer engineering from Istanbul University-Cerrahpasa. He has conference publications on image processing on MR mammography images, hardware security, and text classification in natural language processing. He continues his academic studies on text classification, deep learning, and sentiment analysis in human–robot interaction.

YAKUP KUTLU received the B.S. degree in electrical and electronics engineering from Dokuz Eylül University, the master's degree in electrical and electronics engineering from Mustafa Kemal University, and the Ph.D. degree in electrical and electronics engineering from Dokuz Eylül University. He is currently an Associate Professor with the Department of Computer Engineering, Iskenderun Technical University. He has an extensive background in academia, specializing in electrical and electronics engineering. His research interests include machine learning, artificial intelligence, natural language processing, and deep learning. He has numerous publications in these fields and continues to make significant contributions to these areas of study.

AHMET SERTBAS was born in Istanbul, Turkey, in 1965. He received the B.Sc. and M.Sc. degrees in electronic engineering from Istanbul Technical University, Istanbul, in 1986 and 1990, respectively, and the Ph.D. degree in electric-electronic engineering from Istanbul University, Istanbul, in 1997. Since 2000, he has been an Assistant Professor, an Associate Professor, and a Professor with the Computer Engineering Department, Istanbul University, and a Professor with the Computer Engineering Department, Istanbul University-Cerrahpasa, since 2018. He has 25 articles in indexed SCI-SCIE journals and many journal articles not indexed SCI-SCIE and international conference papers. His research interests include image processing, artificial intelligence, computer arithmetic, and hardware security.