International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 September 2023

Enhancing Sentiment Analysis for Autistic Children: A Hybrid Approach Using SBERT and Ensemble Learning

Dr. T. Dheepak[1], K. Saraswathi[2], Dr. T. Suresh [3]

[1]Assistant Professor of Computer Science, CDOE, Bharathidasan University Tiruchirappalli-620 024, Tamilnadu, India.
[2]Assistant Professor, Dhanalakshmi Srinivasan College of Arts and Science for Women (Autonomous) (Affiliated to Bharathidasan University, Tiruchirappalli), Perambalur, Tamilnadu, India.
[3]Assistant Professor in AIML, K. Ramakrishnan College of Engineering, Samayapuram, Trichirappalli, Tamilnadu, India

Abstract

Autistic children utilise both verbal and nonverbal means of communication. Children diagnosed with autism predominantly communicate their feelings through verbal expression rather than physical gestures or behaviours. Hence, this work suggests a blending model that integrates Sentence-BERT as a language embedding model with Voting Classifier as a machine learning ensemble model. The proposed mixture model enhances the efficacy of sentiment analysis models for autistic children by incorporating both sentence embedding and semantic weights. This research presents a prospective paradigm for analysing the emotions of children with autism.

Keywords: Autistic children, NLP, SBERT, Machine Learning, Ensemble Learning.

1. INTRODUCTION

Developmental disabilities encompass impairments and deviations in social interactions, communication, and cognitive growth. Autistic children, a subset of individuals with developmental disorders, exhibit a delay in the development of language and communication skills. Children diagnosed with autism communicate their feelings through the use of language and engage in repeated behaviours. Immediate or delayed echo terms are used to characterise them. Hence, carers find it challenging to comprehend the emotional experiences of children with autism.The user's text is [1]

This research study employs a language model and a machine learning classification technique to categorise emotions based on linguistic patterns. Sentence-BERT (SBERT) is a language model that enhances the performance of sentence embedding by incorporating pooling operations into the output of BERT. SBERT generates vector representations for each input sentence and computes sentence embeddings by measuring their semantic similarity using cosine similarity computations [3]. The Voting Classifier ensemble model in machine learning involves the use of two or more machine learning models for the learning process. Ultimately, the Voting Classifier makes its decision based on the machine learning outcomes obtained during the learning process in order to generate the final prediction [9].

This research effort introduces an S-BEL (Sentence BERT – Ensemble Learning) Mixture model that merges the Voting Classifier model between SBERT and ensemble models. The S-BEL Mixture model generates fixed vectors for each input sentence and assigns semantic weights. This study constructs a model utilising emotional dialogue data from individuals without disabilities due to challenges in gathering data specifically for children diagnosed with autism. The proposed S-BEL Mixture model allows for the emotional categorization of individuals without disabilities. Moreover, the S-BEL Mixture model that is being suggested has the capability to produce responses in the form of suitable chatbots that align with the analysed emotions.

2. BACKGROUND OF AUTISM SPECTRUM DISORDER

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition that significantly impacts the lives of children, shaping their social interactions, communication abilities, and behavioral patterns. As a spectrum disorder, ASD manifests in a wide range of presentations, from mild to severe, encompassing diverse challenges that make each affected child's experience unique. This introduction seeks to provide a comprehensive overview of Autism Spectrum Disorder in children, exploring the definition, diagnostic criteria, prevalence, and the intricate interplay of factors that contribute to the complexity of this condition.

Defining Autism Spectrum Disorder: Autism Spectrum Disorder is characterized by persistent deficits in social communication and interaction, coupled with restricted and repetitive patterns of behavior, interests, or activities. This umbrella term acknowledges the diverse ways ASD can manifest, emphasizing the unique strengths and challenges each individual may possess. The spectrum encompasses several previously distinct disorders, including autistic disorder, Asperger's syndrome, and pervasive developmental disorder-not otherwise specified (PDD-NOS).

The core features of ASD encompass challenges in social reciprocity, difficulty in non-verbal communication, and the presence of repetitive behaviors or intense interests. These characteristics may emerge as early as infancy, but a formal diagnosis is typically made during early childhood, often before the age of three.

Diagnostic Criteria and Assessment: The diagnostic criteria for ASD, as outlined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), involve a thorough evaluation of a child's behaviors and developmental history. Clinicians consider deficits in social-emotional reciprocity, nonverbal communication, and developing and maintaining relationships. Repetitive behaviors, stereotyped movements, and rigid adherence to routines or interests are also key indicators. A comprehensive assessment may involve input from parents, teachers, and various healthcare professionals, ensuring a holistic understanding of the child's functioning.

Prevalence and Demographics: The prevalence of Autism Spectrum Disorder has increased over the past few decades, with greater awareness, improved diagnostic criteria, and a broader understanding contributing to more accurate identification. According to the Centers for Disease Control and Prevention (CDC), approximately 1 in 44 children in the United States are diagnosed with ASD. Boys are diagnosed more frequently than girls, with a ratio of around 4:1.

Prevalence rates also vary across different demographic groups, highlighting the importance of considering cultural and contextual factors in understanding the impact of ASD. It is crucial to recognize that ASD affects individuals from all ethnic, socioeconomic, and cultural backgrounds.

Heterogeneity of ASD: The term "spectrum" reflects the heterogeneity of ASD, acknowledging the vast variability in how the disorder presents in different individuals. Some children with ASD may have exceptional cognitive abilities in specific areas, such as mathematics or music, while others may face significant intellectual challenges. The spectrum also includes individuals with varying degrees of language proficiency, from non-verbal to highly articulate.

This diversity poses challenges for researchers, clinicians, and educators, emphasizing the need for personalized and tailored approaches to support each child's unique strengths and address specific challenges.

Early Signs and Developmental Milestones: Early identification of ASD is crucial for implementing timely interventions. Parents and caregivers often notice subtle signs during a child's first few years of life, such as a lack of response to their name, limited eye contact, delayed speech or language development, and atypical play behaviors. Tracking developmental milestones becomes essential, allowing for early intervention services that can significantly improve outcomes for children with ASD.

Genetic and Environmental Factors: The etiology of Autism Spectrum Disorder is complex and involves a combination of genetic and environmental factors. While a genetic predisposition is evident in many cases, environmental influences, such as prenatal complications, exposure to certain toxins, or maternal infections during pregnancy, also contribute to the risk of developing ASD.

Ongoing research aims to unravel the intricate interplay between genetic and environmental factors, providing valuable insights into the underlying mechanisms of ASD. Understanding these factors is essential for developing targeted interventions and personalized treatment approaches.

Impact on Families and Caregivers: Raising a child with ASD can have a profound impact on families and caregivers. The unique challenges associated with the disorder, including navigating the diagnostic process, accessing appropriate interventions, and addressing the daily needs of the child, can create stressors that extend beyond the affected individual. Support networks, educational resources, and community understanding play pivotal roles in helping families cope with the demands of raising a child with ASD.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Educational and Therapeutic Approaches: The educational and therapeutic landscape for children with Autism Spectrum Disorder has evolved significantly. Evidence-based interventions, such as Applied Behavior Analysis (ABA), speech and language therapy, and occupational therapy, are commonly employed to address specific challenges. Early intervention services, provided in collaboration with parents and educators, focus on developing communication skills, social interaction, and adaptive behaviors.

Societal Awareness and Inclusion: As awareness of ASD has grown, there has been a shift towards promoting inclusion and understanding in society. Advocacy efforts seek to reduce stigma, increase public awareness, and foster inclusive environments in schools, workplaces, and communities. Acceptance and accommodation of neurodiversity are essential for creating a society that values the unique contributions of individuals with ASD.

Research and Future Directions: Ongoing research endeavors continue to deepen our understanding of Autism Spectrum Disorder. Advances in neuroscience, genetics, and behavioral science contribute to the development of targeted interventions and therapeutic approaches. The quest for early biomarkers, personalized treatment plans, and a deeper understanding of the genetic underpinnings of ASD remains at the forefront of research efforts.

3. RELATED RESEARCH

Park Sang-min, Lee Jae-Yoon, Sung Yuri, and Kim Jae-Eun have put up a method for developing an SBERT model that incorporates keyword information to accurately represent semantic vectors. Park Sang-min, Lee Jae-Yoon, Sung Yuri, and Kim Jae-Eun created a dataset for training purposes that consists of positive and negative keywords. This dataset was built using sentence n-gram to ensure it contains relevant keyword information. Park Sang-min, Lee Jae-Yoon, Sung Yuri, and Kim Jae-Eun employed the SBERT model to acquire knowledge from the gathered learning sentences and created keyword pairs. In this correlated investigation, an SBERT model trained on data that mirrors keywords exhibited a 2.74% enhancement in performance in comparison to the preexisting SBERT model. The user's text is [2].

Yoon Hye-jin, Ku Ja-hwan, and Kim Eung-mo employed five distinct word embedding approaches and three different machine learning classification models to enhance sentiment classification. The CounterVectorizer, TfidfVectorizer, Word2vec (CBOW), Word2vec (Skip-gram), and Pretrain_Word2vec techniques were employed for generating five-word embeddings. Additionally, the Decision Tree, RandomForest, and Logistic Regression models were utilised as machine learning classification models. Testing revealed that CounterVectorizer outperforms TfidifVectorizer, as indicated by related studies. Yoon Hye-jin, Koo Ja-hwan, and Kim Eung-mo conducted a study comparing the accuracy based on the number of emotion categories. Consequently, they verified that the precision increased by over 30% while employing two emotion categories (positive and negative) in contrast to using seven emotion categories. Furthermore, they verified that a segmented emotion category was inappropriate for the classification of emotions using embedding techniques. The user's text is [4]

Park Sang-min, Lee Jae-yoon, Son Yu-ri, and Kim Jae-eun generated two keywords using ngram analysis. They then matched these keywords with phrases to facilitate learning. He subsequently conducted an analysis of emotions using three distinct machine learning techniques: Yoon Hye-jin, Koo Ja-hwan, and Kim Eung-mo. Furthermore, they also verified the enhancement of the accuracy of the analysis model in relation to the quantity of emotion categories. Human emotions can be conveyed through diverse means. Certain emotions defy categorization as either positive or negative. By integrating many machine learning models instead of relying on a single model, the accuracy of emotion machine learning can be enhanced due to the specificity and diversity of emotion models. Thus, this paper categorises the keywords into four distinct moods for the purpose of conducting training. Furthermore, this study presents a novel approach that integrates many machine learning models by leveraging natural language processing techniques and ensemble methods.

4. S-BEL MIXTURE MODEL

This study developed a model by utilising over 20,000 emotional datasets sourced from the National Information Society Agency. The emotional data collected by the National Information Society Agency is categorised into six key categories: "joy," "embarrassment," "anger," "anxiety," "wound," and "sadness." These categories encompass emotions, human response, and system response. This paper employs human response one and system response 1. Furthermore, this study incorporates the wound category as the representation of grief and the panic category as the representation of rage. Hence, we employ a grand total of four emotion categories, namely "joy," "anxiety," "anger," and "sadness." The schematic of the S-BEL Mixture model, as described in this paper, is seen in Figure 1 below. The suggested model comprises a Sentence Bert model, an ensemble machine learning model, and an S-BEL Mixture model.

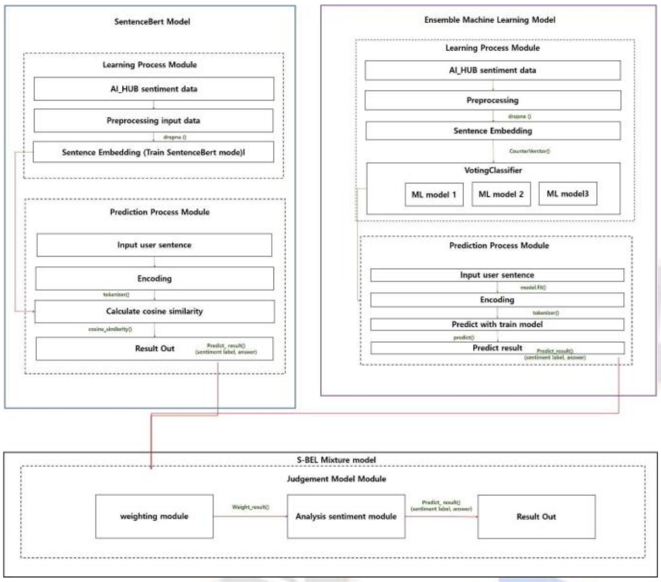

Figure 1. S-BEL Mixture Model Diagram

(ML1: MultinomialNB, ML2 : RandomForest, ML3 : XGBoost)

The S-BEL hybrid model, as presented in this study, utilises the prediction outcomes of SBERT models and ensemble machine learning models. The S-BEL Mixture model utilises these paperweights to forecast the outcome values. The SBERT model trains using the STS dataset, which quantifies sentence similarity, and the NLI dataset, which categorises the relationship between sentences. This study focuses on training the Sentence Bert model through fine-tuning. The training data includes pairs of human sentences and their corresponding emotion categories, which are used as feature values. The Sentence BERT, which has been trained, uses an encoding technique to convert the sentences entered by the user into vectors. These vectors are then used to produce predictions based on the user's input. The process of encoding sentences into vectors allows us to forecast the most similar sentence and emotion by calculating the cosine similarity using the trained model. The calculation of cosine similarity is performed using Equation 1 [8].

$$\text{similarity} = \cos(\theta) = \frac{A * B}{\|A\|\|B\|}$$

Equation 1. Cosine Similarity Formula

Ensemble learning employs multiple machine learning models from within the field of machine learning classifications. The ensemble learning model aggregates the prediction outcomes of the constituent models to get a single, consolidated prediction outcome. The VotingClassifier models employed in this study utilise XGBoost, RandomForest, and MultinomialNB models. A Hard-Voting model determines the predicted value of the class that receives the highest number of votes among the outcomes of each classifier model using majority classification. On the other hand, a Soft-Voting model selects the class with the highest mean probability among all classifier predictions.The user's text is "[7]". We utilised the Soft-Voting model for the proposed concept. In this research, we employ the CounterVectorizer technique to convert sentences into vectors based on the frequency of word occurrences. We apply this method to emotional data that has been collected. Subsequently, we proceeded to train the MultinomialNB, RandomForest, and XGBoost models and then integrated them into the VotingClassifier model via a pipeline. The suggested ensemble model involves the process of embedding user input words in order to predict the emotions associated with the user's input. The model's prediction determines the emotional outcome of the input sentence based on its probability.

The weighting formula for the S-BEL Mixture model, as described in this research, is presented in Equation 2. In the above model, b_s represents the similarity measure of the SBERT model, W_ms represents the accuracy value of the ensemble learning model, and X_mp represents the probability value assessed by the ensemble learning model. Equation 2 conducted a correlation analysis between accuracy and prediction probability using the multiplication of W_ms and X_mp. Furthermore, equation 2 demonstrates the deflection magnitude as a function of the reciprocal of b_s. In addition, Equation 3 utilised a sigmoid function formula to rescale values inside the range of 0 and 1 [10].

$$S.BEL = W * X + (b - 1)$$

Equation 2. S-BEL Mixture Model Weight Formula

($S.BEL$: Weight result value,
W : $ELaccuracyvalueofthemodel$, X :
Probability value measured through EL model,
b : $SimilaritymeasureofSBERTmodel$)

1

$S.BEL = e$______ (.)

Equation 3. Weight scaling formula of S-BEL Mixture model

The S-BEL Mixture model proposed a decision module that utilised a threshold. This study employed an accuracy comparison by selecting a range of thresholds from 0.1 to 1.0 in order to determine the most suitable threshold. According to Figure 2, the accuracy was maximum when the threshold value was set to 0.5. Hence, the suggested S-BEL mixture model examines if the S-BEL mixture weight value exceeds or falls below the designated threshold (0.5) using the assessment module. If the decision outcome of the decision module exceeds or equals the designated threshold, an emotion label and chatbot-style response that aligns with the emotion predicted by the proposed model will be generated. Alternatively, if the determination result of the decision module is below the set threshold value, the result value is not produced.

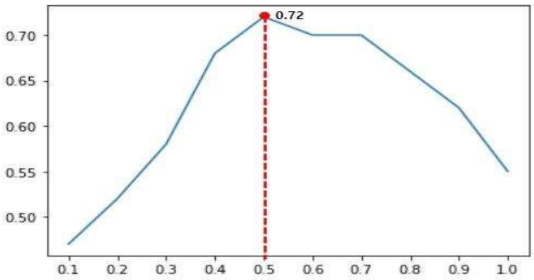

Figure 2. Performance evaluation results according to the threshold

5. EXPERIMENTS AND RESULTS

This paper compares the accuracy of the classification model and the proposed model, as shown in Figure 3. In this paper, we measured accuracy by dividing the train data and the test dataset in a 7:3 ratio and dividing the total number of samples by the correctly predicted number. As a result of the experiment, the accuracy was higher when progressing using ensemble learning than when progressing with sentiment analysis using each machine learning algorithm. Furthermore, the results ensured that the proposed S-BEL Mixture model achieved the highest accuracy by comparing the compared models. We obtained an accuracy of about 93% in the sentiment analysis performed.

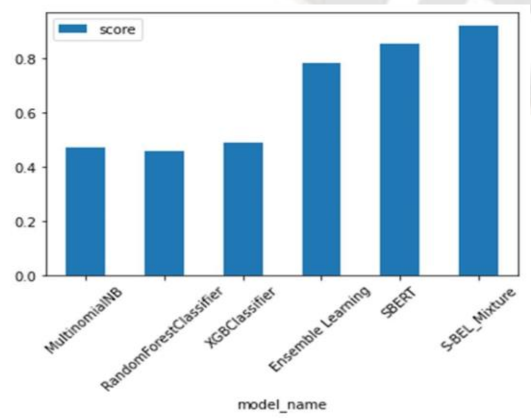

Figure 3. Benchmark Model Accuracy Comparison

For experimental evaluation, we utilized BLEU and Cosine-Similarity evaluations to ensure the accuracy of the chatbot-type response of the S-BEL Mixture system. BLEU (Bilingual Evaluation Understudy) is one of the methods for evaluating the performance of machine translation systems. [11] BLEU measures performance by comparing how similar the result of the machine translation system is to the actual result. The BLEU evaluation method means that the higher the score, the better the performance.[11] Cosine-Similarity measures the similarity between the generated sentence and the correct answer sentence. Cosine-Simlarity is measured through the angle of two vectors and ranges from -1 to +1. The closer Cosine-Similarity is to +1, the highest the similarity.[12]

To verify the evaluation method proposed, we conducted a comparative analysis of the BLEU Score and Cosine Similarity Score for each data. Figure 4,5,6,7,8,9 shows each algorithm's BLEU and Cosine Similarity Score scores according to the number of test data sets. Figure 4 is BLEU Score and Cosine Similarity Score Graph about S-BEL Mixture Model. Figure 5 is the BLEU Score and Cosine Similarity Score Graph about the Ensemble Model consisting of MultinomialNB, RandomForesClassifier, and XGBoost. Figure 6 is the BLEU Score and Cosine Similarity Score Graph about the Ensemble Model consisting of MultinomialNB and RandomForestClassifier. Figure 7 is the BLEU Score and Cosine Similarity Score Graph about the Ensemble Model consisting of MultinomialNB, XGBoost. Figure 8 is the BLEU Score and Cosine Similarity Score Graph about the Ensemble Model consisting of RandomForesClassifier, XGBoost. Figure 9 shows the correlation between the BLEU score of the S-BEL Mixture Model and the cosine similarity score. The heat map on the right side of Figure 9 shows the correlation between S-BEL Mixture Model's BLEU score and cosine similarity score using the Pearson correlation coefficient. The Pearson correlation coefficient represents a linear relationship between two variables. Pearson correlation coefficients indicate that as the correlation coefficient value approaches 1, the relationship between the two variables is higher.[13]

Figures 4, 5, 6, 7, 8, and 9 show that the evaluation score differs depending on the number of data sets. Since the text of the correct answer heavily influences BLEU and Cosine Similarity Scores, the accuracy decreases when an answer is generated by using a word that is not in the correct answer in the test data set. Figure 10 shows the correlation between the proposed S-BEL Mixture model's BLEU Score and the Cosine Similarity Score. The heat map graph on the right shows that

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

the BLEU Score and the Cosine Similarity Score have a high correlation of 0.88. The heat map graph means that the BLEU Score and the Cosine Similarity Score derive scores in a similar way. Therefore, the verification method proposed in this paper has a significant meaning.

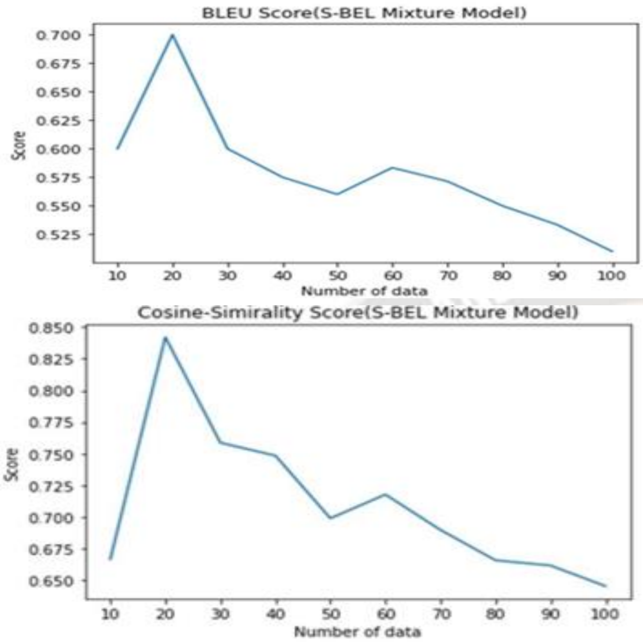

Figure 4. S-BEL Mixture Model BLEU, Cosine-Similarity Score

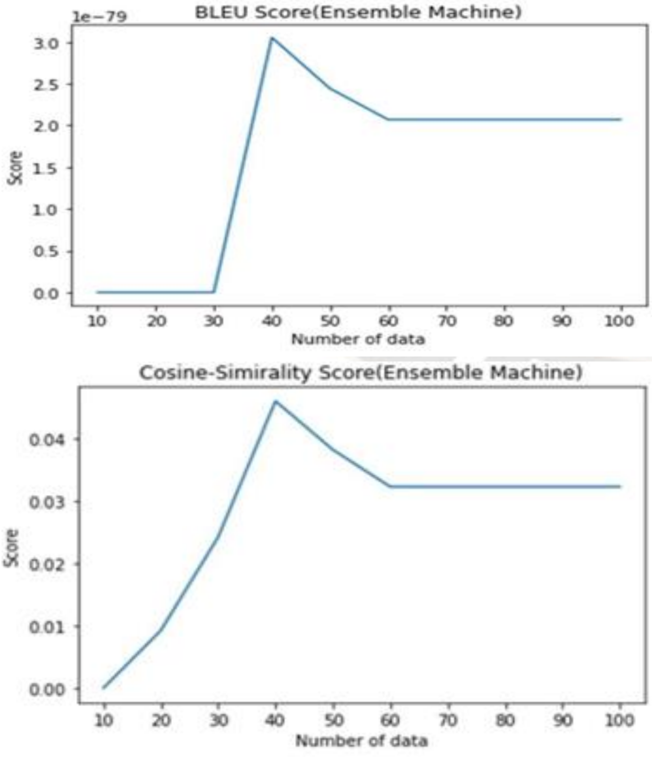

Figure 5. Ensemble Machine (MultinomialNB, RandomForesClassifier, XGBoost) BLEU, Cosine-Similarity Score

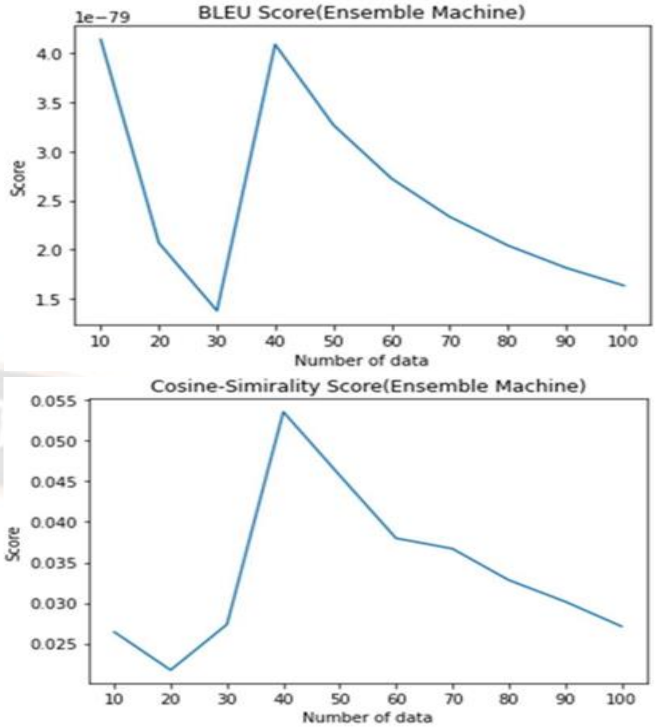

Figure 6. Ensemble Machine (MultinomialNB, RandomForesClassifier) BLEU, CosineSimilarity Score

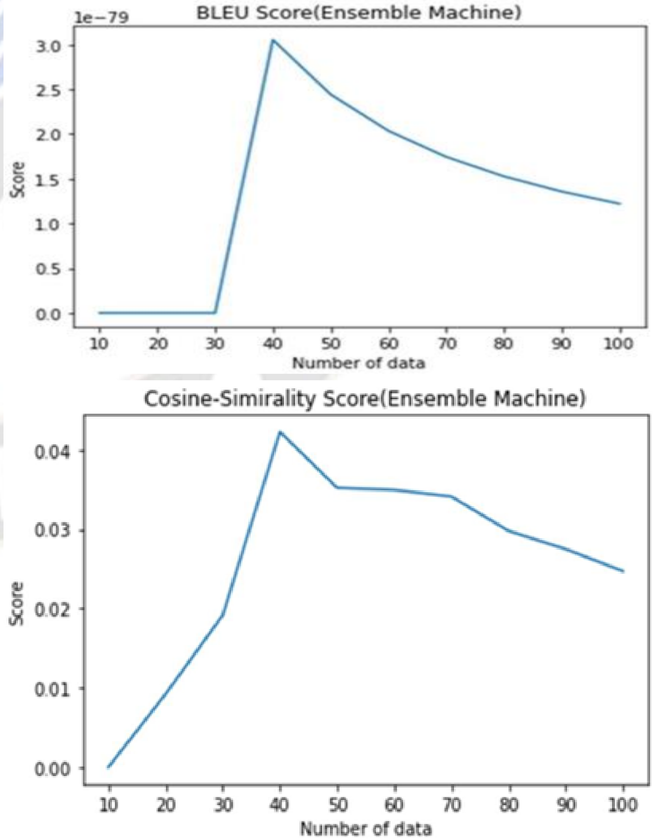

Figure 7. Ensemble Machine (MultinomialNB, XGBoost) BLEU, Cosine-Similarity Score

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

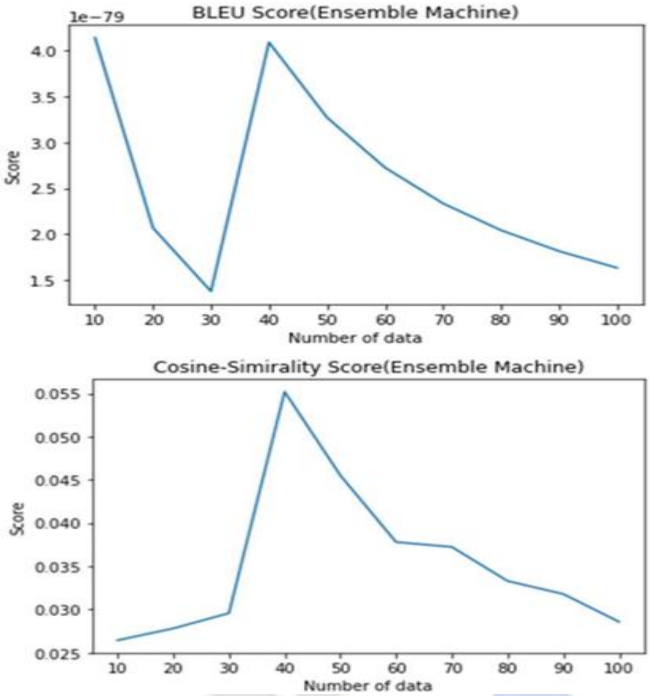

Figure 8. Ensemble Machine (RandomForesClassifier, XGBoost) BLEU, CosineSimilarity Score

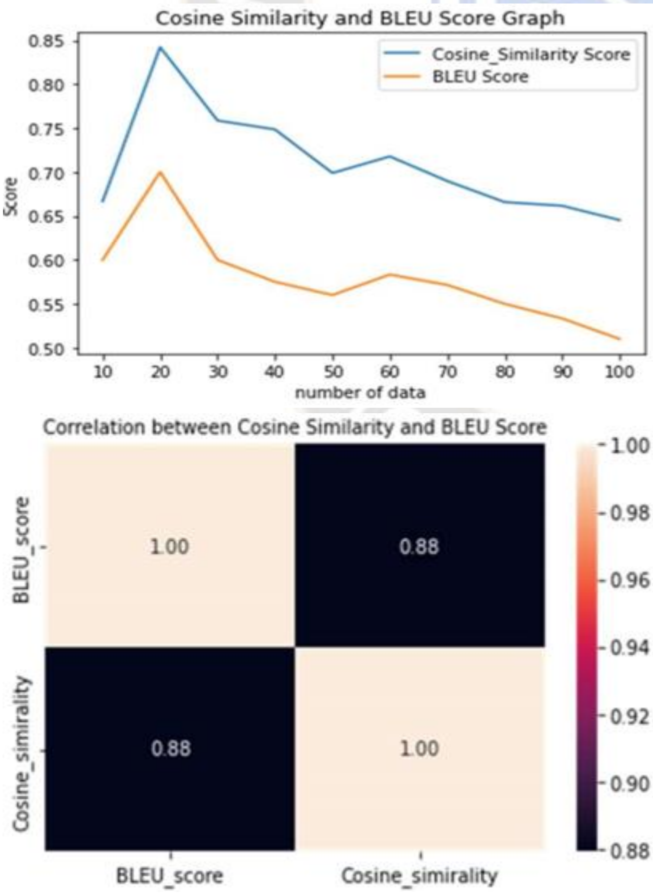

Figure 9. Correlation between Cosine Similarity and BLEU Score Graph

In this research paper, we evaluated a dataset of 60 emotion sentence tests with the highest score through the evaluation verification method. As a result of the experiment, as shown in Table 1, it can be confirmed that the proposed S-BEL Mixture model derives high performance in all performance evaluations with a BLEU value of 0.593 and a CosineSimilarity of 0.712.

Table 1. Measurement of the accuracy of each system's chatbot-type response

System name	BLEU Score	Cosine-Similarity Score
S-BEL Mixture system	0.593	0.712
Ensemble Machine (MultinomialNB, RandomForesClassifier, xgboost)	0.020	0.032 (1e-79)
Ensemble Machine (MultinomialNB, RandomForesClassifier)	0.016	0.027 (1e-79)
Ensemble Machine (MultinomialNB, xgboost)	0.011	0.024 (1e-79)
Ensemble Machine (RandomForesClassifier, xgboost)	0.016	0.028 (1e-79)

Table 2 below is a sample result of the proposed S-BEL mixture model system. Through the resulting sample, this paper confirmed whether the actual emotion label for the user input sentence matches the emotion label predicted by the proposed mixed model.

Table 2. Experience Result (Excerpts of some sentences from the test data)

	Input Sentence	Real sentiment labels	Predict sentiment labels	Answers in the form of chatbots
1	Sad	Sadness	Sadness	You must be sad.
2	That's very annoying.	Anger	Anger	You're annoying.
3	I'm angry.	Anger	Anger	I think you'll be so angry.
4	Glad.	Joy	Joy	Oh, you must be so happy!
5.	Are you living a	Joy	JOY	After getting married, my

	happy married life?			life has become so comfortable

Table 2 shows that the actual and predicted emotion labels resulting from the input sentences are almost identical. In addition, the proposed model confirms that the actual emotion labels of simple sentences such as 'I am angry' and 'I am happy' match the predicted emotion labels. Therefore, the proposed model is capable of appropriate chatbot-style response output and emotion label prediction.

6. RESULTS AND DISCUSSION

In this research paper, we proposed an emotion analysis model using SBERT, ensemble machine learning, and S-BEL Mixture weights. Our proposed S-BEL Mixture model had the highest accuracy compared to other emotion analysis models for analyzing emotion. The currently proposed sentiment analysis model used the sentiment data of non-disabled people rather than the dialogue data of autistic children for sentiment analysis. For future research, we plan to collect conversation data about children with autism and train using the data collected. Then we will fine-tune the SBERT according to the training data. Therefore, future research will improve the performance of learning models by optimal fine-tuning. In addition to models such as RandomForest and XGBoost used to improve the accuracy of the VotingClassifier model, we plan to combine various machine learning models to present the optimal combination and proceed with learning. The correct answer data set greatly influences the verification method used in this paper. Therefore, we will verify this in future studies by pre-processing the correct answer data set. We also plan to extend the proposed mixture model to improve its performance of the model.

REFERENCES

[1] JeongMi Lee. "Study on Spontaneous Communicative Behavior in Children with Autism." Emotional and Behavioral Disorder Research 23, no.4 (2007): 345-359.

[2] SangMin Park, JaeYoon LeeYu Ri Sung, JaeEun Kim. (2022). SBERT-IQ: A SentenceBERT-based embedding model considering keyword information content. Proceedings of the Korean Information Science Society, (), 1058-1060.

[3] Nils Reimers, Iryna Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," In Proceedings of Empirical Methods in Natural Language Processing, pp. 3980-3990, 2019.

[4] HyeJin Yoon, JahWan Koo, and UngMo Kim. "Comparison of Word Embedding Performance for Sentiment Classification." Proceedings of the Annual Meeting of the Korean Information Processing Society 28.2 (2021): 760-763.

[5] Y. J. Jang, H. J. Kim, and H. J. Jo, "Data Mining," KNOU PRESS, 2016, 1-200

[6] W. Yun, D.H. Seo, S. Min, H. Nam, "Breast Cancer Classification using RandomForest and XGBoost," Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 113-114, 2021.

[7] M. C. Jeong, J. Lee, and H. Oh, "Ensemble Machine Learning Model Based YouTube Spam Comment Detection," Journal of the Korea Institute of Information and Communication Engineering, vol. 24, no. 5, pp. 576–583, May 2020.

[8] SungEun Gil. (2020). Research on action-based authentication by cosine similarity measure Proceedings of the Annual Meeting of the Korean Information Processing Society, 27(1), 165-168.

[9] HuiWon Song, JuHyun Lee, SungHae Jun, SangSung Park. (2022). A Study on Inferable Patent Evaluation Model using Ensemble Method. Journal of Korean Institute of Intelligent Systems, 32(5), 379-385.

[10] Sanjeev Kumar, Mahesh Chandra). "Detection of Microcalcification Using the WaveletBased Adaptive Sigmoid Function and Neural Network." Journal of information processing systems 13.4 (2017): 703-715.

[11] Junghyun Cho, Hyunki Jung, Chanyoung Park, and Yuseop Kim. " An Autonomous Assessment of a Short Essay Answer by Using the BLEU " Korean HCI Conference 2009.2 (2009): 606-610.

[12] Eungju Kwon, Jongwoo Kim, Nojeong Heo, and Sanggil Kang. "Personalized Recommendation System using Level of Cosine Similarity of Emotion Word from Social Network" Information Technology Architecture Research 9.3 (2012): 333-344.

[13] Dongho Lee, Minseo Lee, Jaeseo Lee, Heejo Park, Hyunsung Jun, Seehwan Yoo.(2021).Real-time adaptive model training technique using Pearson correlation coefficient. Proceedings of the Korean Information Science Society Conference,15361538.