Computers, Materials & Continua

Tech Science Press

Doi:10.32604/cmc.2025.065566

ARTICLE

Chinese DeepSeek: Performance of Various Oversampling Techniques on Public Perceptions Using Natural Language Processing

Anees Ara[1], Muhammad Mujahid[1], Amal Al-Rasheed[2,*], Shaha Al-Otaibi[2] and Tanzila Saba[1]

[1]Artificial Intelligence & Data Analytics Lab, CCIS, Prince Sultan University, Riyadh, 11586, Saudi Arabia

[2]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11671, Saudi Arabia

*Corresponding Author: Amal Al-Rasheed. Email: aaalrasheed@pnu.edu.sa

Received: 16 March 2025; Accepted: 28 April 2025; Published: 03 July 2025

ABSTRACT: DeepSeek Chinese artificial intelligence (AI) open-source model, has gained a lot of attention due to its economical training and efficient inference. DeepSeek, a model trained on large-scale reinforcement learning without supervised fine-tuning as a preliminary step, demonstrates remarkable reasoning capabilities of performing a wide range of tasks. DeepSeek is a prominent AI-driven chatbot that assists individuals in learning and enhances responses by generating insightful solutions to inquiries. Users possess divergent viewpoints regarding advanced models like DeepSeek, posting both their merits and shortcomings across several social media platforms. This research presents a new framework for predicting public sentiment to evaluate perceptions of DeepSeek. To transform the unstructured data into a suitable manner, we initially collect DeepSeek-related tweets from Twitter and subsequently implement various preprocessing methods. Subsequently, we annotated the tweets utilizing the Valence Aware Dictionary and sentiment Reasoning (VADER) methodology and the lexicon-driven TextBlob. Next, we classified the attitudes obtained from the purified data utilizing the proposed hybrid model. The proposed hybrid model consists of long-term, short-term memory (LSTM) and bidirectional gated recurrent units (BiGRU). To strengthen it, we include multi-head attention, regularizer activation, and dropout units to enhance performance. Topic modeling employing KMeans clustering and Latent Dirichlet Allocation (LDA), was utilized to analyze public behavior concerning DeepSeek. The perceptions demonstrate that 82.5% of the people are positive, 15.2% negative, and 2.3% neutral using TextBlob, and 82.8% positive, 16.1% negative, and 1.2% neutral using the VADER analysis. The slight difference in results ensures that both analyses concur with their overall perceptions and may have distinct views of language peculiarities. The results indicate that the proposed model surpassed previous state-of-the-art approaches.

KEYWORDS: DeepSeek; prediction; natural language processing; deep learning; analysis; TextBlob; imbalance data

1 Introduction

DeepSeek, a Chinese firm based in Hangzhou, is intensifying its research and development efforts in artificial intelligence. DeepSeek is pleased to introduce the competitively priced R1 Model, unlike others who are required to allocate substantial resources towards artificial intelligence research. The model utilises optimisation techniques such as dimension reduction, weight minimisation, and shape simplification. The most advantageous aspect of DeepSeek is that the model is open source, allowing anyone to utilise and enhance it [1,2]. It uses human reinforcement learning (HRL) to make it more responsive, rotational positioning for accuracy, and a mix of experts (MOE) to achieve outclass accuracy. DeepSeek, after its launch date, has many more downloads for the Apple Store that ensure its predictability [3]. In the financial

sector, Flyer employs artificial intelligence, especially in the creation of investment models and trading strategies [4,5].

The performance of DeepSeek is analogous to that of OpenAI's GPT. This is facilitated by DeepSeek's open-source approach, which promotes a cost-effective infrastructure model and provides a more flexible infrastructure solution [6]. Due to United States (US) restrictions on NVIDIA's export of advanced GPUs to China and India, companies have been necessitated to create robust AI models in China. Despite these challenges, DeepSeek successfully developed two AI models utilizing cloud storage and AI processors produced in China. To maintain competitiveness in the rapidly evolving AI market, DeepSeek is focusing on the development of models that are faster, more precise, and more sophisticated [7,8]. The optimal solutions are covered using the GPT and LLaMA models [9].

Sentiment analysis, sometimes referred to as opinion mining, is the study of text to ascertain its sentimentality, that is, its polarity, positive or negative. Knowing from their online reviews how current and potential clients regard your model will be quite beneficial. This might help researchers to have a better experience and improved advertising strategies as well [10,11]. Sentiment analysis methods sort user reviews, comments, and social media entries. Track what others have to say about your model online to learn about their level of features, methods, and service satisfaction. This study improves the efficacy of reputation management by helping one identify and deal with unpleasant emotions. Sentiment analysis will help in this study to avoid going over every comment by hand. Companies that get a lot of comments via multiple marketing channels could discover this process to be a time-consuming chore. Topic modelling can extract latent semantic patterns from the vast text for thematic analysis. This statistical modeling method searches a text corpus for groups of like terms using unsupervised machine learning. The subjects of a corpus of work can help one to subtly characterise it. When writing about a given subject, authors could find that a particular vocabulary term or phrase appears more frequently and jokes using GPT [12,13]. The study [14] analyzed tweets with topic modeling and labeling.

We extract data from Twitter and employ several ways to preprocess it, thereby cleansing and structuring unstructured data to comprehend and analyse public sentiment and discourse around DeepSeek, including its features, usage, and other facets. Subsequently, we convert tweets or derive sentiments from them through a lexicon-based analysis. This study's principal contributions are as follows:

Contributions

- This study collects tweets from the twitter using DeepSeek hashtag only to extract the relevant tweets in English.
- We are analyzing the tone of tweets from stakeholders to evaluate how well DeepSeek accomplishes the target results.
- In order to understand the thoughts and feelings of AI users regarding the transition to DeepSeek, we use natural language processing (NLP) methods for text processing, sentiment analysis and lexicon based approaches.
- We predict how the public would perceive DeepSeek using the proposed hybrid model.
- We use topic modeling to identify the DeepSeek issues related to complexity, interaction, and technology.
- In addition, we compare the different sampling techniques to address the imbalance issues in the structured data and validate the proposed model performance with several metrics.

2 Literature Review

The authors analyzed sentiment analysis models with the goal to assess baseline study methods' efficacy across several fields and with several datasets. Their results show that the suggested model performs

consistently better than transformers, a well-known model with recent developments having been trained on a large range of datasets. The suggested paradigm became quite successful in several spheres. Given the shortcomings in the "MAMS" dataset [15], they should investigate more trustworthy language models including GPT and possible enhancements or modifications. Another aim of the research was to increase the efficiency of intelligence collecting by analyzing public opinion sentiment using a few-shot learning framework'a key component of open-source intelligence. By combining comprehensive language model knowledge with contrastive suggestions, they present a new way to address the data-driven inadequacies of sentiment detection models in low-resource environments [16].

Social media have removed geographical and chronological constraints, therefore, enabling the free flow of knowledge among people all around the world. Simultaneously, it has been proved to be a great tool for detecting and rapidly lowering national and organizational hazards [17]. Combining sentiment analysis techniques with deep learning algorithms was the aim of the paper [18] to satisfy consumers during message exchanges. Using deep learning methods helps the proposed chatbot to predict the objectives of its users and provide a reasonable and useful response. Deep learning against transfer learning presents a novel approach on how sarcasm detection and sentiment analysis could be included into the conversion process. This work makes use of deep learning, a transformer model well-known for success in NLP-related challenges. The results reveal that the RoBERTa model adequately classified a spectrum of review comments and performed well in sentiment identification.

The assessment technique [19] helped identify important elements affecting the functionality of the instrument and provide recommendations for future improvements. Although there are many different communication models at hand, each has its benefits and drawbacks [20]. The authors enhance sentiment classification performance in an independent study by means of a customised deep learning model that comprises an LSTM network and a modified word embedding technique. They also suggested a combined model that merges our basic classifier with new and creative sentiment analysis classifiers [21].

The research [22] intends to improve tweet prediction precision by studying approaches for refining Machine Learning (ML) models that capture tweets' complicated linkages and contextual nuances. Four famous machine learning models were used to predict disaster-related tweets: Logistic Regression (LR), XGBoost, Random Forest (RF), and Support Vector Machine (SVM). Their algorithms were capable of extracting sentiment, semantic nuances, and relevant information from tweets, allowing them to provide robust predictions. The LR, XGBoost, and RF models produced average F1-scores of 78% and 79%, respectively, while the SVM model performed substantially better, with accuracies of 81% and 82%. A basic model capable of differentiating between positive and negative emotions was also accessible [23]. A sophisticated meta-learning dynamic reward-sharing system utilizing LLMs was created to enhance collaboration among users who are not co-located [24].

3 Materials and Methods

The public's perceptions of Deepseek's Chinese are presented in this study. The principal concepts covered in this section include data collection, preprocessing, lexicion-based sentiment analysis techniques, topic modeling (including cluster-based and LDA modeling), and classification using the proposed approach. The proposed methodology's detailed flow is shown in Fig. 1.

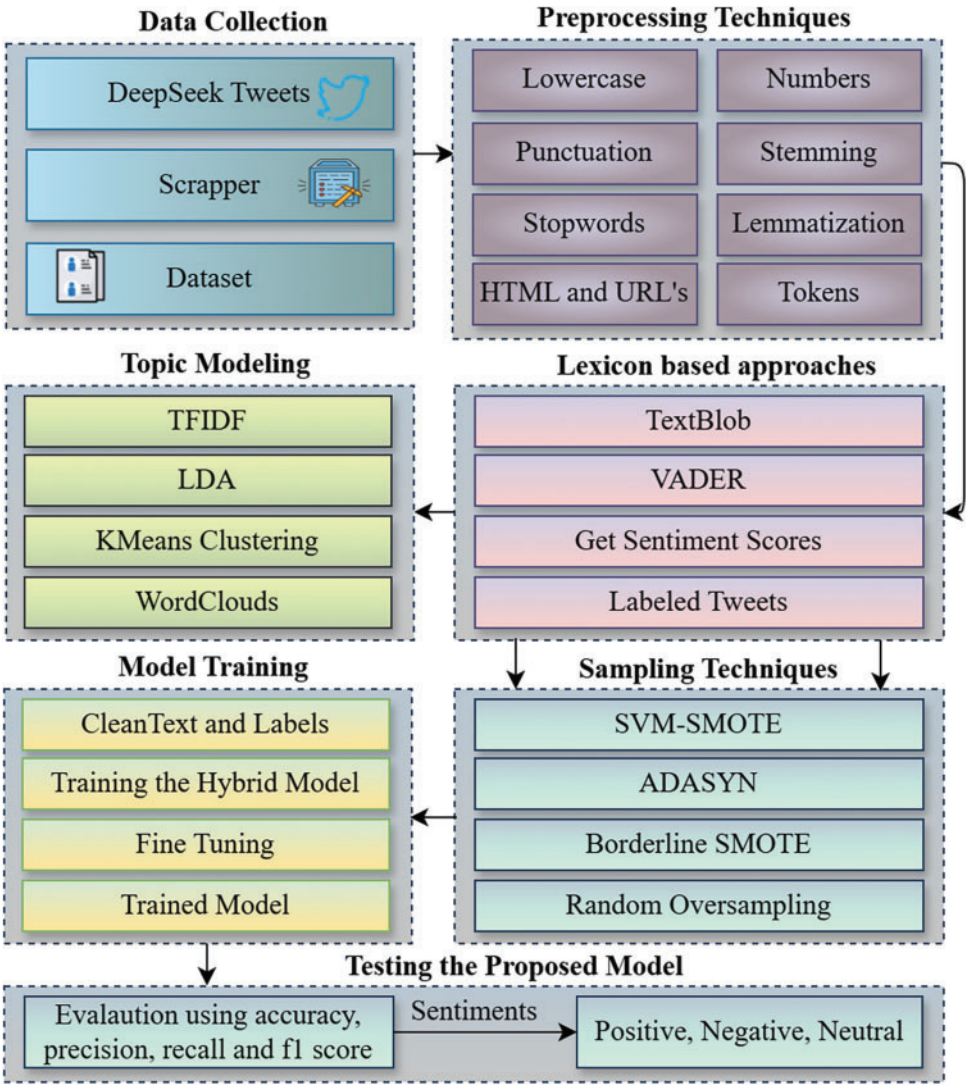

Figure 1: The work flow of the proposed methodology for the perception of public reactions. The approach includes data collection, preprocessing, lexicon-based methodologies such as TextBlob and VADER, LDA topic modeling, sampling techniques, and the training and testing of the proposed model

3.1 Dataset Collection and Preprocessing

The DeepSeek dataset is collected from the twitter using Hashtag. The scrapped tweets consists of irrelevant information and unstructured. The dataset has more than 6000 tweets. Twitter data frequently contains a lot of random words, keywords, phrases, hashtags, URLs, and special characters, and is disorganized and unstructured. Tweets can contain formatting, typography, and user input that makes reliable analysis challenging, in contrast to well-written articles.

To understand raw information, we must first analyze it. The consistent structure and clarity of the content can enhance case summaries, subject modeling, and video analysis. By removing superfluous data from the document, we may derive valuable content and improve the precision of deep learning predictions. Spaces, underlines, and redundant words are instances of unnecessary information that could distort results

and affect model performance if not properly incorporated [25]. A sample of tweets after preprocessing is shown in Table 1. The crucial steps of preprocessing are discussed below:

Table 1: Sample of tweets after preprocessing

Raw tweets	Cleaned tweets
A thread to earn money online Goodbye #ChatGPT	Thread earn money online goodbye chatgpt
It's only been 5 days since China's #Deepseek R1 was launched, and the world is already amazed by its capabilities. . .	days since china deepseek launched world already amazed capabilities
There's been a market shock due to China's #AI initiative, #DeepSeek, affecting #Crypto prices, particularly AI tokens. #Nvidia has also experienced a drop in relation to this news	Market shock due china initiative deepseek affecting crypto prices particularly tokens nvidia also experienced drop relation news
Privacy Alert! Downloaded China's new LLM DeepSeek Android app and saw this message: "We will use your personal info." This raises big concerns about how and why our data is being used. Your data matters! #DeepSeek #DataPrivacy #ChinaDataConcerns #PrivacyMatters	Privacy alert downloaded china new llm deepseek android app saw message use personal info raises big concerns data used data matters deepseek dataprivacy chinadataconcerns privacymatters
DeepSeek is coming for Nvidia! With its powerful new AI model, this Chinese startup might just challenge Nvidia's dominance. The future of AI is getting real interesting. . . #DeepSeek #Nvidia #AI #OpenAI #DeepSeekR	Deepseek coming nvidia powerful new model chinese startup might challenge nvidia dominance future getting real interesting deepseek nvidia openai deepseekr
Big news! We've integrated the DeepSeek into Tearline, making us even more powerful and efficient. Excited to bring you enhanced capabilities and insights! #AI #Innovation #Tearline #DeepSeek	Big news integrated deepseek tearline making even powerful efficient excited bring enhanced capabilities insights innovation tearline deepseek

Lowercase: To maintain consistency, all content should be converted to lowercase. By using this command, words like "twitter" and "Twitter" won't be used as synonyms, which could result in abuse.

HTML and URLs: Links and hyperlinks in tweets frequently don't convey the text's original meaning. The model can better concentrate on the appropriate keywords by eliminating them. Regex expressions, for instance, are used to parse URLs.

Stopwords: Words like "the," "the," "the," and "the" are not important in text analysis. We keep only words that have significance by eliminating their stopwords. But in other situations (like sentiment analysis), some endings might not be required and ought to be omitted.

Punctuation and Numbers: Punctuations, numbers, and special characters are not very important in NLP tasks. So, this study removes the punctuation and numbers. Stemming and Lemmatization: The stem converts the words into their base form but the lemmatization is more important.

Tokenization: The initial stage in text mining pipelines is tokenization, employed to transform unstructured text input into a format suitable for deep processing. It is typically one of the initial steps in natural language processing pipelines, as it is required for later preprocessing techniques.

3.2 Lexicon Based Approaches

Two most crucial approaches are used to label the structures tweets like TextBlob [26] and VADER [27]. TextBlob was able to accomplish its objectives with the assistance of NLTK. The NLTK library simplifies classification and other lexical-intensive procedures. Lexicon-based techniques define sentiment by its semantic orientation and word intensity. It demonstrates sentence subjectivity and polarity. Negative moods are represented by −1, while joyful moods are represented by 1. Negative terms reverse polarity. The subjective value spectrum is defined as [0, 1]. Subjectivity distinguishes personal opinion from factual facts.

$$\text{Polarity Score} = \frac{\sum_{i=1}^{N} \text{Sentiment}(wo_i)}{n} \tag{1}$$

$$\text{Subjectivity Score} = \frac{\sum_{i=1}^{N} \text{Intensity}(wo_i)}{n} \tag{2}$$

$$\text{Compound Score} = \frac{\sum_{i=1}^{n} Se(wo_i) \cdot We(wo_i)}{\sum_{i=1}^{N} W(wo_i)} \tag{3}$$

In Eqs. (1) and (2), wo_i represents the ith word in the tweets or sentence, Sentiment (wo_i) represents the polarity of word wo_i, Intensity (wo_i) represents the subjective of word wo_i and n number of words.

3.3 Topic Modeling

In natural language processing and machine learning, topic modeling is a flexible technique for spotting main ideas in a document corpus. An often used approach for topic modeling is LDA [28]. LDA is a generative probabilistic corpus model used unsupervised. It is based on a random mix of latent topic-containing documents whereby a topic is defined as a distribution over words. Using the statistical and visual approach of LDA helps one to recognize the connections in the word distribution over a corpus of papers. In a given collection of tweets data, a word cloud is a graphic depiction of the most often occurring terms. It can help to clarify the fundamental ideas in the information. It clarifies the structure and content of the data; hence it is vital for the study of text data. As the size of every word in the word cloud shows its frequency of occurrence, the biggest words in the word cloud reflect the most often occurring keywords in the data.

3.4 Proposed Model

The LSTM-BiGRU (long term short term memory-bi-directional gated recurrent unit) hybrid model combines the best features of multiple deep learning methods for handling continuous data, such as text. An embedding layer is the first component of the model that transforms each word or piece of data into a vector of integers that represents the word's meaning. This step is crucial for training the model to acquire words, detect patterns, and comprehend their links. A 100-dimensional vector is represented by each word when the embedding parameter is set to 100. This strikes a nice mix between being sophisticated enough to capture significant information without slowing down the model as shown in Fig. 2.

After being transformed into these vectors, the input data is then sent through a layer of local format. The mismatch helps prevent overfitting, which occurs when the model fails to effectively integrate new information and learn from the training data. In order to force the model to learn general patterns rather than precise details, some input data is randomly ignored during training. The data is subsequently sent to an LSTM layer by the model. The model can follow the complete data series throughout time with the help of the LSTM layer, which is made to return a sequence of values. An LSTM-like but quicker GRU layer is then added to the model. The GRU layer enables more effective training with fewer parameters while handling continuous data in a manner similar to that of an LSTM. Data is processed both end-to-end

and bidirectional by the most potent GRU layer. Prediction is enhanced as a result of the model's ability to comprehend historical and prospective context. A second drop layer is then added to reduce the model's reliance on particular data points. Through the integration of LSTM and GRU layers, this hybrid model leverages both short-term and long-term memory, increasing its resilience and effectiveness. The multi-head focus layer is one of this model's key components. With the use of this method, the model can simultaneously concentrate on several aspects of the incoming data. By examining multiple components simultaneously, the model is able to concentrate on the most crucial data for forecasting. Lastly, the data passes through a 32-unit dense layer. This layer makes a final conclusion by combining the features that were learned in the other layers. A dense output layer with three components and a softmax activation function makes up the final portion of the model.

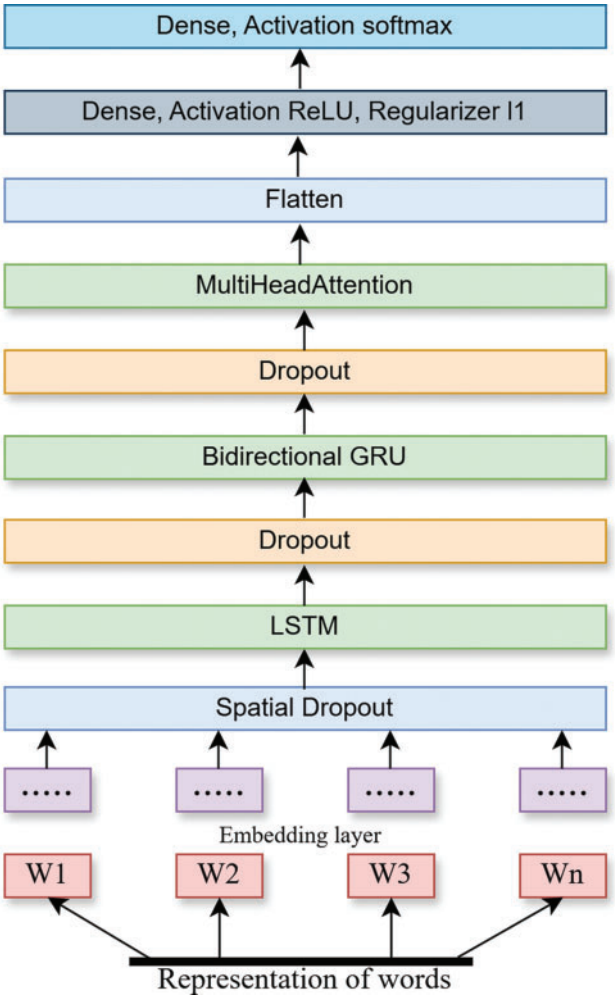

Figure 2: The proposed Hybrid Model leverages the strength of LSTM and BiGRU architectures with a minimal number of layers to develop a lightweight model

4 Results and Discussion

This section delineates the experimental outcomes utilizing TextBlob, VADER, several oversampling methodologies, and the proposed model. The model is assessed utilizing multiple performance measures. Accuracy is calculated by dividing the sum of true positives and true negatives by the total number of

sentiment predictions. Precision is calculated by dividing true positives by the sum of true and false predictions. Recall is calculated by dividing true positives by the sum of true positives and false negatives. The F1 score is the average of precision and recall.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1 - Score = 2 \times \frac{PR \times RE}{PR + RE} \tag{7}$$

4.1 Performance and Impact of SVM-SMOTE and ADASYN

This study utilized the SVM-SMOTE oversampling technique to balance the dataset and examined the impact of sampling on the proposed sentiment classification model. The model attained a maximum average precision of 0.717 and a minimum recall score of 0.706 as shown in Table 2. This research employed the ADASYN method to equilibrate the dataset and analyzed the effect of sampling on the proposed classification of sentiment model. The model achieved precision of 0.821 for sentiment 2 and 0.667 on sentient 0 as illustrated in Table 2.

Table 2: Impact of SVM-SMOTE and ADASYN on performance

	SVM-SMOTE			ADASYN		
Sentiment	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**
0	0.592	0.659	0.624	0.667	0.622	0.643
1	0.849	0.719	0.779	0.755	0.836	0.794
2	0.709	0.737	0.723	0.821	0.789	0.805
Accuracy	–	–	0.706	–	–	0.749
Macro avg.	0.717	0.706	0.709	0.748	0.749	0.747

4.2 Performance and Impact of Borderline-SMOTE and Random Oversampling

After balancing the dataset using the Borderline SMOTE approach, this study examined how sampling affected the proposed sentiment categorization model. According to Table 3, the model obtained a recall of 0.707 and f1 score of 0.734, in addition to a precision of 0.762 for emotion 2. Table 3 also indicates that the model achieved precisions of 0.975, 0.968, and 0.964 for sentiments 2, 1, and 0, respectively. The proposed model demonstrated superior outcomes for random oversampling, and this technique is more effective than others employed in sentiment classification.

Fig. 3 shows the training accuracy, validation accuracy, training loss, validation loss, and ROC-AUC for the proposed model using four different sampling methods. Fig. 3a shows results using ADASYN, Fig. 3b shows results using Borderline SMOTE, Fig. 3c shows results with random oversampling, and Fig. 3d shows results with SVM-SMOTE. Fig. 3e shows training accuracy, Fig. 3f shows validation accuracy, Fig. 3g shows training loss, and Fig. 3h shows validation loss.

Table 3: Impact of borderline SMOTE and random oversampling on performance

	Borderline SMOTE			Random oversampling		
Sentiment	**Precision**	**Recall**	**F1 score**	**Precision**	**Recall**	**F1 score**
0	0.632	0.675	0.653	0.964	0.981	0.972
1	0.787	0.790	0.789	0.975	0.970	0.972
2	0.762	0.707	0.734	0.968	0.947	0.957
Accuracy	–	–	0.724	–	–	0.966
Macro avg.	0.727	0.724	0.725	0.969	0.966	0.967

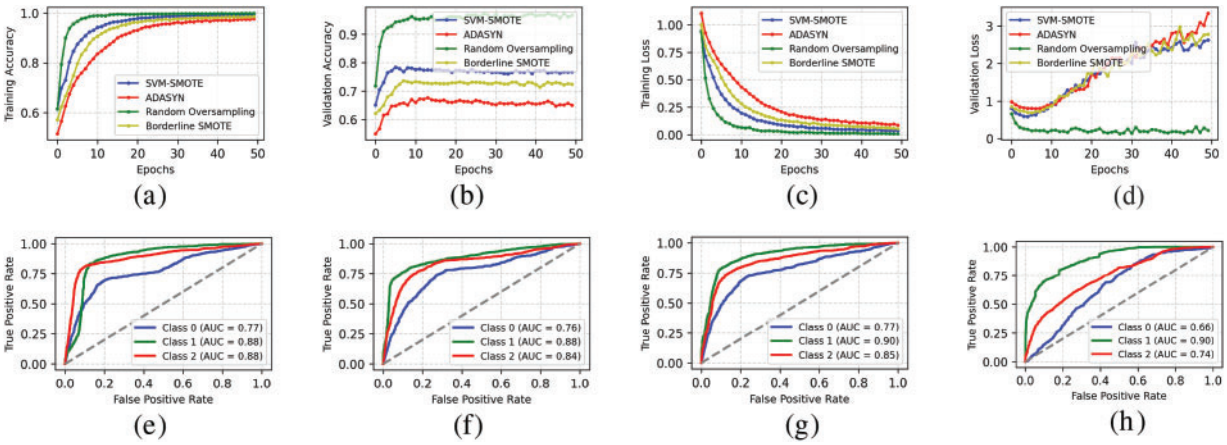

Figure 3: The learning and ROC-AUC for the proposed model utiling different sampling techniques. (**a**) shows that random oversampling has the best accuracy; (**b**) shows that ADASYN oversampling has the lowest validation accuracy; (**c**) and (**d**) show that random oversampling has the lowest training and validation loss; and the figure in the last row also shows great performance

4.3 Topic Modeling Results

LDA-based topic modeling is a flexible technique for spotting the main ideas in a document corpus and is often used as an approach to generative probabilistic corpus models that are unsupervised. Fig. 4 shows LDA-based topics, and Fig. 5 demonstrates the KMeans cluster-based topics.

Figure 4: LDA based topics extraced from the cleaned tweets using the proposed model with TFIDF

Figure 5: KMeans based topics extraced from the cleaned tweets using the proposed model with TFIDF

The sentiments of people regarding the DeepSeek are shown in Fig. 6. More than 80% of user responses in this study were positive, according to DeepSeek sentiment analysis, demonstrating the tool's high level of acceptability and favorable attitude. The increasing popularity of many internet platforms is linked to this tendency. DeepSeek's sophisticated NLP capabilities, which assist in producing intricate texts and dialogues using language that is similar to that of humans, can capture this positive mood.

Figure 6: The sentiments of people using TextBlob and VADER technique

Extensive discussions on social media, forums, and technical communities, where users show their value in comprehending context and producing pertinent comments, provide credence to this favorable appraisal and result in very high user acceptance and participation. Fig. 7 shows sentiment distribution across positive, negative, and neutral tweets with the top 10 topics extracted using LDA. Nevertheless, the sample's analysis also showed that both positive and negative words were present. A tiny percentage of users mostly students voiced reservations over the tool's applicability, pointing to issues with cost, feasibility, and validation. These results emphasize the necessity of ongoing development to resolve user issues and enhance the tool's functionality across various use cases, which will boost efficiency and adoption. The frequency of each most occurred words in positive and negative tweets is represented in Fig. 8.

This research provides an overview of the distribution of words in a dataset by combining numerous variables and revealing the meaning of words in a 3D plot. First, new words are extracted from the database, and their characteristics are examined using methods like hierarchical clustering. By doing this, a matrix of document words is produced, enabling us to determine the frequency with which each word occurs throughout the dataset.

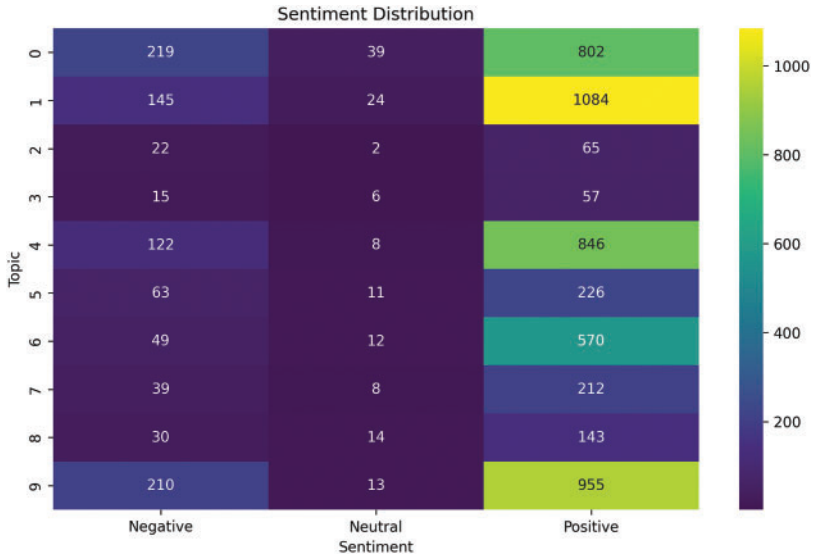

Figure 7: Sentiment distribution across positive, negative and neutral tweets with top 10 topics extracted using LDA

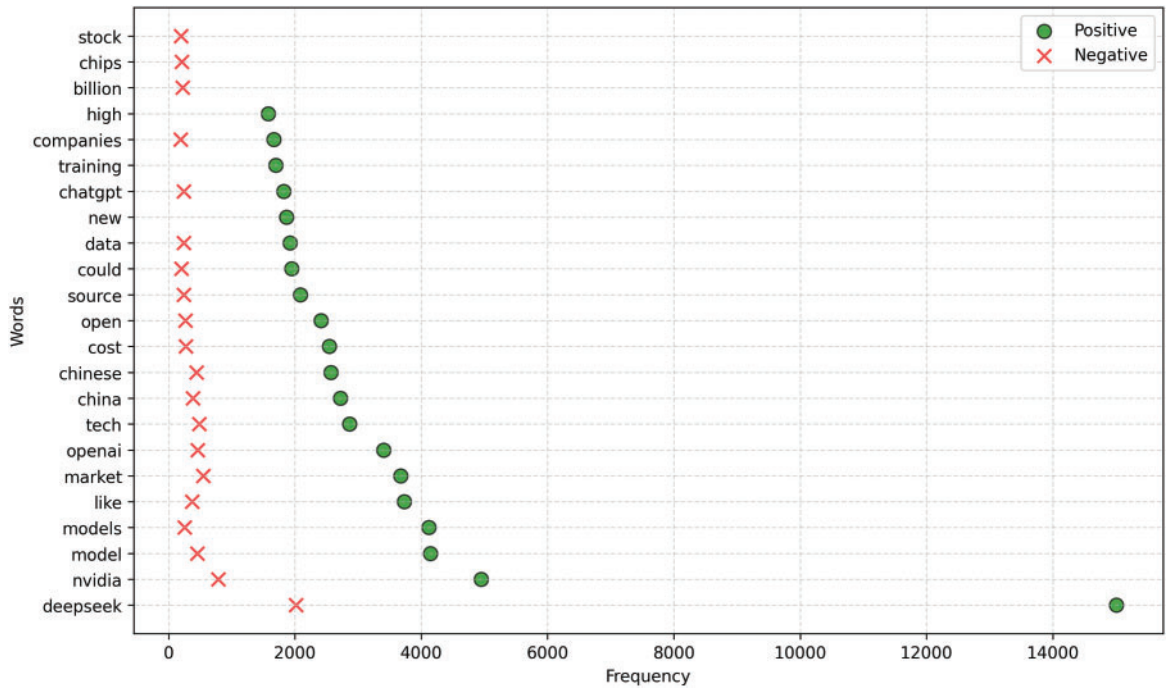

Figure 8: Frequency of each most occurred words in positive and negative tweets

VADER sentiment analysis was then used to examine the sentiment of each sentence. This tool gives a general mood score between −1 (which indicates a bad mood) and 1 (which indicates a good mood). According to this classification, words having negative meanings are positioned on the left side of the screen, while those that are deemed good are positioned on the right.

A scatter plot in three dimensions aids in the explanation of these connections. With positive words on the right and negative words on the left, the x-axis displays each word's mood. The frequency of terms in the

database is displayed on the y-axis, where greater numbers indicate higher levels of expression. Each word's length is displayed on the z-axis, which also positions shorter words farther from the root and closer to it. A greater comprehension of word frequencies, word length distributions, and stress patterns is made possible by this 3D representation, which also offers insightful information about the dataset's linguistic characteristics. Word frequency for the entire dataset is shown in Fig. 9.

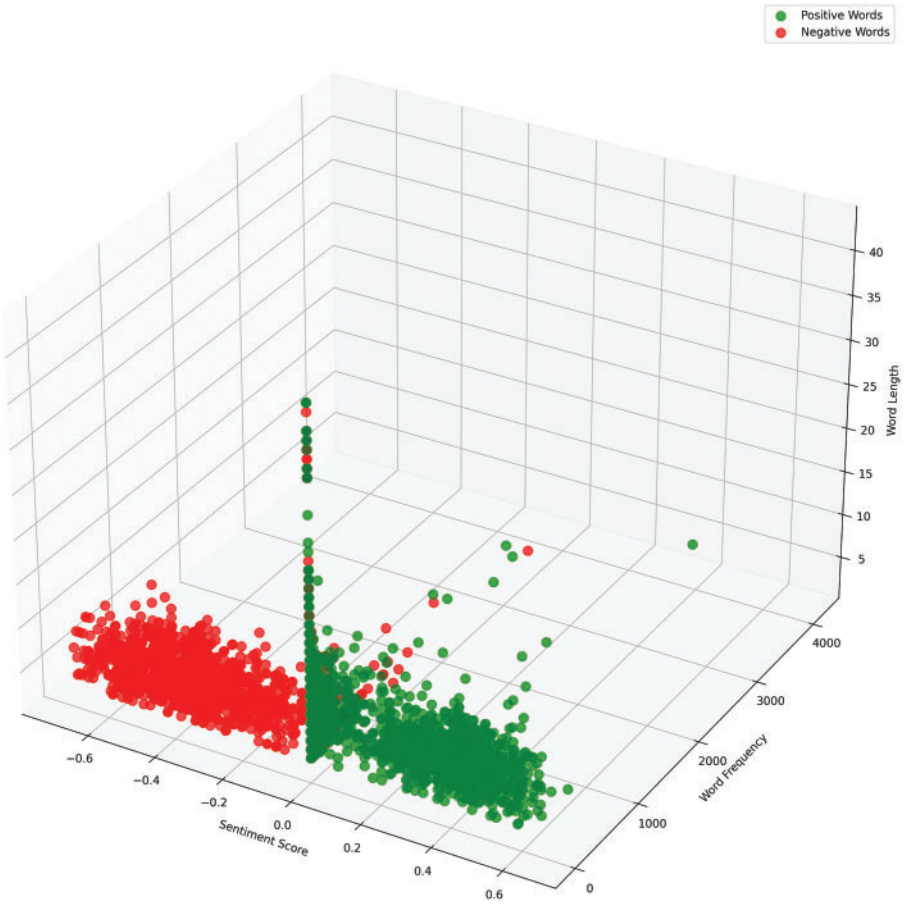

Figure 9: Word frequency for the entire dataset. The x-axis plots the sentiment score, such as 0.2, 0.4, and 0.6, which indicates positive sentiment, whereas −0.2, −0.4, and −0.6 indicate negative sentiment

A large number of users, particularly within services, express apprehension over the cost, quality, and reliability of AI technologies. Open source platforms mitigate these problems by diminishing the necessity for costly coding tools, hence enhancing the accessibility of AI, particularly inside systems of logistics. They are readily deployable on-premises or on cloud servers, rendering them appropriate for diverse environments [29]. These models promote collaboration among academics and healthcare practitioners, aiding in the resolution of clinical challenges, including regional health concerns. Employing environmental data to test models enhances their precision and applicability in practical scenarios. Open-source frameworks exhibit more stability, facilitating analysis and yielding more dependable results. Furthermore, they can be tailored with complimentary content, ensuring their continued relevance over time. Ultimately, these technologies assist individuals in making improved health decisions by uncovering information they may otherwise overlook, resulting in more precise and effective therapy [30].

5 Discussion and Conclusion

This study used unstructured tweets linked to DeepSeek that were extracted from Twitter and cleansed using several preprocessing techniques. Next, two lexicon approaches, TextBlob and VADER, are used to label preprocessed tweets. According to TextBlob, 82.5% of people have favourable opinions about DeepSeek, 15.2% have negative opinions, and only 2.3% are neutral. In contrast, the VADER study reveals that 82.8% of people have positive opinions, 16.1% have negative opinions, and 1.2% are neutral. We then used the proposed framework to classify sentiments and evaluate how different oversampling techniques affected the performance. The proposed model achieved 70.6% for SVM-SMOTE, 74.9% for ADASYN, 72.4% for Borderline SMOTE, and 96.6% for random oversampling in classifying the tweets as positive, negative, and neutral. The random oversampling is most effective for the text data. According to our study, most respondents had positive impressions about DeepSeek, and only 2% had negative ones. A large number of users greatly appreciate the technology's ability to aid in various fields. Certain individuals, though, find DeepSeek more interesting than others; this is especially true when considering its capacity to provide complex solutions and thorough answers on a marketing and industrial scale.

DeepSeek is tremendously effective due to its frequent sampling; it reduces expenses by improving resource management, reducing the need for manual labor, and reducing overall expenditures over time. The ability to make well-informed decisions offers various benefits as well, including the ability to gain a competitive edge, increase operational efficiency, and better align with strategic plans. It automates routine tasks and simplifies operations in various areas.

Every technology has some limitations. DeepSeek Chat has opened up to its limitations, including the inability to learn new facts beyond initial training and the occasional production of misinformation. There are significant concerns about job displacement as a result of DeepSeek AI's integration across industries. DeepSeek has less real-world application data and is mostly targeted at the Chinese market. The open-source chatbot only reads the first few lines of data that the user uploaded for modifications, which is another limitation for understanding the entire context. DeepSeek could be improved with sophisticated algorithms and reliable data storage technologies.

The use of transformations and training approaches has allowed DeepSeek to significantly improve speed and accuracy. Artificial intelligence-powered DeepSeek chatbots and virtual assistants quickly answer client questions. Current technology can boost customer satisfaction, minimize support staff workload, and speed up response time. More sustainable real-world applications have been made possible by the decrease in labor costs, which has been made possible by improvements in performance like cost efficiency and adaptability.

Acknowledgement: This research is supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project, Riyadh, Saudi Arabia. The authors are also thankful to AIDA Lab CCIS Prince Sultan University, Riyadh Saudi Arabia of APC support.

Funding Statement: This research is funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Anees Ara, Muhammad Mujahid, Tanzila Saba; methodology, Tanzila Saba; software, Muhammad Mujahid; validation, Amal Al-Rasheed, Shaha Al-Otaibi; formal analysis, Muhammad Mujahid; investigation, Shaha Al-Otaibi; resources, Amal Al-Rasheed; data curation, Muhammad Mujahid; writing—original draft preparation, Muhammad Mujahid, Anees Ara; writing—review and editing, Anees Ara; visualization, supervision, Tanzila Saba; project administration, Amal Al-Rasheed; funding acquisition, Amal Al-Rasheed, Shaha Al-Otaibi. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author, Amal Al-Rasheed, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Krause D. DeepSeek and FinTech: the democratization of AI and its global implications. 2025. doi:10.2139/ssrn.5116322.
2. Mondillo G, Colosimo S, Perrotta A, Frattolillo V, Masino M. Comparative evaluation of advanced AI reasoning models in pediatric clinical decision support: ChatGPT O1 vs. DeepSeek-R1. medRxiv. 2025. doi:10.1101/2025.01.27.25321169.
3. Zeff M. DeepSeek displaces ChatGPT as the App Store's top app. TechCrunch. [Internet]. 2025 [cited 2025 Apr 27]. Available from: https://techcrunch.com/2025/01/27/deepseek-displaces-chatgpt-as-the-app-stores-top-app/.
4. Chowdhury NA, Li W, Zhang M. Redefining scalability in AI: the innovations behind DeepSeek-V3's MoE and multi-token prediction. Int J Sci Res Publ. 2024;1(5):1–3. doi:10.5281/zenodo.14742743.
5. Bi X, Chen D, Chen G, Chen S, Dai D, Deng C, et al. Deepseek llm: Scaling open-source language models with longtermism. arXiv:2401.02954. 2024. doi:10.48550/arXiv.2401.02954.
6. Gibney E. China's cheap, open AI model DeepSeek thrills scientists. Nature. 2025;638(8049):13–4. doi:10.1038/d41586-025-00229-6.
7. TechTarget. DeepSeek explained: everything you need to know. [Internet]. 2025 [cited 2025 Feb 2]. Available from: https://www.techtarget.com/whatis/feature/DeepSeek-explained-Everything-you-need-to-know.
8. Krause D. DeepSeek's potential impact on the magnificent 7: a valuation perspective. [cited 2025 Jan 30]. Available from: https://ssrn.com/abstract=5117909. doi:10.2139/ssrn.5117909.
9. Kumar A, Sharma R, Bedi P. Towards optimal NLP solutions: analyzing GPT and LLaMA-2 models across model scale, dataset size, and task diversity. Eng Technol Appl Sci Res. 2024;14(3):14219–24. doi:10.48084/etasr.7200.
10. Yu D, Xiang B. Discovering topics and trends in the field of artificial intelligence: using LDA topic modeling. Expert Syst Appl. 2023;225(1):120114. doi:10.1016/j.eswa.2023.120114.
11. Semary NA, Ahmed W, Amin K, Pławiak P, Hammad M. Enhancing machine learning-based sentiment analysis through feature extraction techniques. PLoS One. 2024;19(2):e0294968. doi:10.1371/journal.pone.0294968.
12. Akbar NA, Darmayanti I, Fati SM, Muneer A. Deep learning of a pre-trained language model's joke classifier using GPT-2. J Hunan Univ (Nat Sci). 2021;48(8):235–41.
13. Alharbi A, Hai AA, Aljurbua R, Obradovic Z. AI-driven sentiment trend analysis: enhancing topic modeling interpretation with ChatGPT. In: IFIP International Conference on Artificial Intelligence Applications and Innovations; 2024 Jun 27–30; Corfu, Greece. Cham, Switzerland: Springer Nature. p. 3–17.
14. Koonchanok R, Pan Y, Jang H. Public attitudes toward ChatGPT on twitter: sentiments, topics, and occupations. Soc Netw Anal Min. 2024;14:106. doi:10.1007/s13278-024-01260-7.
15. Mughal N, Mujtaba G, Shaikh S, Kumar A, Daudpota SM. Comparative analysis of deep natural networks and large language models for aspect-based sentiment analysis. IEEE Access. 2024;12(2):60943–59. doi:10.1109/ACCESS.2024.3386969.
16. Yang H, Zi Y, Qin H, Zheng H, Hu Y. Advancing emotional analysis with large language models. J Comput Sci Softw Appl. 2024;4(3):8–15.
17. Arora A, Arora A, McIntyre J. Developing chatbots for cyber security: assessing threats through sentiment analysis on social media. Sustainability. 2023;15(17):13178.
18. Merizig A, Belouaar H, Bakhouche MM, Kazar O. Empowering customer satisfaction chatbot using deep learning and sentiment analysis. Bull Electr Eng Inform. 2024;13(3):1752–61.
19. Kebede D, Tesfai N. Ai-powered text analysis tool for sentiment analysis [bachelor thesis]. Vasteras, Sweden: Malardalen University; 2023.

20. Senthilkumar M, Chowdhary CL. An AI-based chatbot using deep learning. In: Intelligent systems. Palm Bay, FL, USA: Apple Academic Press; 2019. p. 231–42.
21. Alsayat A. Improving sentiment analysis for social media applications using an ensemble deep learning language model. Arab J Sci Eng. 2022;47(2):2499–511. doi:10.1007/s13369-021-06227-w.
22. Fattah M, Haq MA. Tweet prediction for social media using machine learning. Eng Technol Appl Sci Res. 2024;14(3):14698–703.
23. Wang H, Qiu X, Tan X. Multivariate graph neural networks on enhancing syntactic and semantic for aspect-based sentiment analysis. Appl Intell. 2024;54(22):11672–89. doi:10.1007/s10489-024-05802-6.
24. Qiu X, Wang H, Tan X, Qu C. ILTS: inducing intention propagation in decentralized multi-agent tasks with large language models. In: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM'24); 2024; Boise, ID, USA. p. 3989–93. doi:10.1145/3627673.3679942.
25. Chai CP. Comparison of text preprocessing methods. Nat Lang Eng. 2023;29(3):509–53. doi:10.1017/s1351324922000213.
26. Diyasa IGSM, Mandenni NMIM, Fachrurrozi MI, Pradika SI, Manab KRN, Sasmita NR. Twitter sentiment analysis as an evaluation and service base on python textblob. IOP Conf Series Mat Sci Eng. 2021;1125:012034.
27. Elbagir S, Yang J. Twitter sentiment analysis using natural language toolkit and VADER sentiment. In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2019 (IMECS 2019); 2019 Mar 13-15; Hong Kong, China.
28. Tajbakhsh MS, Bagherzadeh J. Semantic knowledge LDA with topic vector for recommending hashtags: twitter use case. Intell Data Anal. 2019;23(3):609–22.
29. Mikhail D, Farah A, Milad J, Nassrallah W, Mihalache A, Milad D, et al. Performance of DeepSeek-R1 in ophthalmology: an evaluation of clinical decision-making and cost-effectiveness. medRxiv. 2025. doi:10.1101/2025.02.10.25322041.
30. Temsah A, Alhasan K, Altamimi I, Jamal A, Al-Eyadhy A, Malki KH, et al. DeepSeek in healthcare: revealing opportunities and steering challenges of a new open-source artificial intelligence frontier. Cureus. 2025;17(2):e79221. doi:10.7759/cureus.79221.