International Journal of Engineering Research and Science & Technology

IJERST

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com

Peer Reviewed,Referred & Indexed Journal

Research Paper

A Semantically Enriched Hybrid Learning Paradigm for Large-Scale Opinion Mining in Geopolitical Crisis Narratives

B. Laxmi Pathi[1*], Gujja Avinash Babu[2], Ballepu Bhargav[2], Mathangi Madhu[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering

[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: B. Laxmi Pathi (laxmipathibilla2020@gmail.com)

ABSTRACT

Among the more than 500 million tweets generated daily, a considerable number reflect public perspectives on global socio-political events such as the Russia–Ukraine War. Analyzing sentiment at this scale manually is time-consuming, inconsistent, and unsuitable for real-time decision-making. To overcome these limitations, this study presents a hybrid deep learning framework for automated sentiment analysis of Twitter discussions related to the Russia–Ukraine conflict. The approach begins with comprehensive Natural Language Processing (NLP) preprocessing, including tokenization, stopword removal, normalization, and lemmatization, followed by Exploratory Data Analysis (EDA) to analyze sentiment distribution, word frequency patterns, and textual trends. Context-aware semantic representations are generated using Distilled Robustly Optimized BERT Pretraining Approach (DistilRoBERTa), which offers efficient and lightweight transformer-based embeddings with reduced computational cost. To handle class imbalance and enhance minority class prediction, KMeans-Synthetic Minority Over Sampling Technique (SMOTE) is applied for synthetic data generation. For comparative evaluation, traditional classifiers such as Logistic Regression (LR), Support Vector Machine Classifier (SVC), and Random Forest Classifier (RFC) are utilized. The proposed model combines a Deep Neural Network (DNN) for effective feature learning with a Greedy Tree-based classifier (GTC) to improve classification performance and generalization capability. The resulting system, named "Semantic Distil Deep Tree (SDDT)," demonstrates superior performance over baseline models in terms of accuracy and F1-score. Overall, the proposed framework delivers a scalable, efficient, and dependable solution for real-time sentiment monitoring, aiding policymakers, researchers, and analysts in understanding public opinion on global conflicts.

Keywords: Sentiment Analysis, Russia–Ukraine War, Twitter Data, Natural Language Processing (NLP), DistilRoBERTa, Deep Neural Network (DNN).

1. INTRODUCTION

The ongoing Russia–Ukraine War (RUW) is one of the largest armed conflicts in Europe since 1945 as shown in Fig. 1. On 24 February 2022, Russia initiated its invasion of Ukraine by advancing into its territory and launching attacks on major cities, including Chernihiv, Berdyansk, Odesa, Sumy, and Kyiv. The war has shocked the entire globe in many ways, for example, [1] the global inflation rate has increased by 3% and the world's gross domestic product (GDP) growth reduced by 1.44%, and the USA alone saw its GDP fall by 1.28% since the war began. Studies such as [2] gave two major reasons for this war; first, Ukrainian willingness to join the North Atlantic Treaty Organization (NATO). Second, NATO's enlargement into eastern Europe. NATO is a group of 30 countries in a security alliance and, according to, "NATO's fundamental goal is to safeguard the Allies' freedom and security by political and military means". Therefore, Ukraine's prospective intentions and NATO's foreseen expansion for European security turned this political clash into an armed battle. The RUW is not only affecting eastern Europe but also damaging living standards in the whole world. Despite receiving aid from NATO, Ukraine is still in a volatile and unpredictable situation [3].

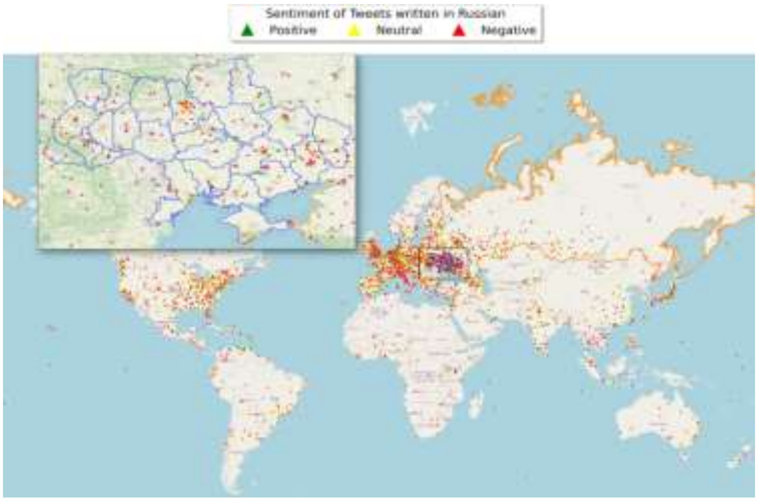


Fig. 1: Global sentiment distribution of 3.84 million Russian-language tweets.

According to the aforementioned statistics and rationales, it is essential to understand societal emotions. Such emotional expressions in the form of text are helpful to form attitudinal insights. Nowadays, major countries are using both territory and information on social media as part of the war effort. Such social media information dispersion usually turns into information warfare between citizens and politicians of countries participating in the war. For decades there has been antagonism between Ukraine and Russia but in the recent past, the turmoil has flared up belligerently [4]. The year 1991 was revolutionary as with the expiry of the Soviet Union, Ukraine proclaimed its national independence on 24 August 1991.

2. RELATED WORK

The rapid growth of social media, financial technologies, and intelligent data-driven systems has led to the development of advanced machine learning and deep learning approaches for cyber intelligence, sentiment analysis, and predictive analytics. Early approaches relied on traditional statistical and machine learning methods, which lacked the ability to capture complex patterns in large-scale, heterogeneous data. Recent research has shifted toward hybrid models, deep learning architectures, and multi-source data integration to improve prediction accuracy, scalability, and real-time decision-making.

2.1 Cyber Intelligence and Multi-Source Data Analysis

The integration of multiple data sources has significantly enhanced cyber intelligence systems. Sufi et al. [5] introduced an autonomous cyber intelligence framework utilizing four semantic agents to collect and analyze data from social media and web sources. The system incorporated techniques such as CNN, sentiment analysis, LDA, TF-IDF, and exponential smoothing to process multilingual data. Evaluated on over 37,000 tweets across 54 languages and extensive web data, the system generated comprehensive intelligence reports within seconds. The study demonstrated the effectiveness of combining multiple analytical techniques for real-time intelligence generation.

2.2 Sentiment-Aware Financial Prediction Models

Sentiment analysis has become a critical component in financial forecasting. Frohmann et al. [6] proposed a hybrid model combining time series forecasting with sentiment analysis derived from microblogs to predict Bitcoin prices. The study introduced a fine-tuned BERT model and a novel weighting scheme based on user influence, achieving improved prediction accuracy with lower MAE and RMSE values. Similarly, Simionescu et al. [10] integrated sentiment analysis into econometric ARDL models for inflation forecasting, demonstrating superior performance compared to SARIMA and machine learning techniques. These studies highlighted the importance of sentiment-driven insights in enhancing financial prediction models.

2.3 Deep Learning for Time Series Forecasting

Advanced deep learning models have been widely adopted for forecasting complex financial data. Aldabagh et al. [7] proposed a hybrid CNN–LSTM model for crude oil price prediction, effectively capturing both spatial and temporal features. The model outperformed traditional methods such as ARIMA, SVM, and standalone LSTM models in both short-term and long-term forecasting scenarios. The study validated the superiority of hybrid deep learning architectures in handling non-linear and dynamic time series data.

2.4 Topic Modeling and Social Media Analytics

Topic modeling techniques have evolved to handle large-scale short-text data from social media platforms. Mendonça et al. [8] explored the optimization of BERTopic for Twitter data, improving coherence, stability, and diversity of extracted topics. The study successfully identified key discussion themes and demonstrated enhanced performance over traditional topic modeling methods. Caratù et al. [11] further analyzed social media engagement related to energy communities, revealing low public awareness and emphasizing the need for structured communication strategies. These studies contributed to understanding user engagement and thematic analysis in digital platforms.

2.5 User Profiling and Behavioral Analysis

Understanding user behavior has become essential for personalized and predictive systems. Wang et al. [9] proposed a deep user profiling approach based on behavioral patterns using an improved stacking ensemble learning model. By incorporating high-dimensional feature extraction and a refined 3Cs evaluation principle, the model achieved significant improvements in prediction accuracy and AUC performance. The study demonstrated the effectiveness of ensemble learning in capturing complex behavioral characteristics.

2.6 Advanced Sentiment Analysis Techniques

Recent advancements in natural language processing have improved sentiment classification accuracy. Michailidis et al. [12] evaluated multiple approaches, including traditional machine learning, artificial neural networks, transfer learning, and large language models for Greek sentiment analysis. The results showed that transformer-based models such as BERT and GPT-4 achieved significantly higher accuracy compared to conventional methods. This study highlighted the growing importance of transformer architectures in handling multilingual sentiment analysis tasks.

2.7 Social Network Analysis and Information Propagation

Analyzing interaction patterns in social networks provides insights into information dissemination and polarization. Kratzke et al. [13] investigated retweet interaction networks and identified echo chambers using graph-based techniques such as greedy modularity maximization and HITS metrics. The study revealed distinct communication patterns and emphasized the role of structural analysis in detecting misinformation spread without relying on content-based filtering.

2.8 Fake News Detection and Cross-Platform Analysis

The detection of fake news has become increasingly important in maintaining information integrity. Ferdush et al. [14] proposed a multi-platform fake news detection framework using the Dempster–Shafer theory to combine evidence from different social media sources. The model analyzed relationships between comments rather than relying solely on textual content, achieving improved detection accuracy. The study demonstrated the benefits of integrating content-based and interaction-based features for robust fake news identification.

2.9 IoT and Machine Learning Integration

The convergence of IoT and machine learning has enabled the development of intelligent and connected systems. Bzai et al. [15] provided a comprehensive review of ML-enabled IoT applications, categorizing them based on data, application domains, and industry use cases. The study also highlighted emerging trends such as edge computing, Internet of Behavior (IoB), and lightweight deep learning models. Additionally, it identified key challenges across technological, business, and societal dimensions, offering insights for future research in smart and sustainable environments.

2.10 Research Gap

Although significant advancements have been made in cyber intelligence, sentiment analysis, and predictive modeling, several limitations persist. Many existing systems focus on single-domain data sources, limiting their ability to capture complex cross-domain relationships. Additionally, high computational requirements and model complexity reduce scalability in real-time applications. There is also a lack of unified frameworks that integrate sentiment analysis, behavioral modeling, and predictive analytics into a single system.

The proposed system addresses these gaps by introducing an integrated framework that combines multi-source data analysis, advanced deep learning techniques, and efficient feature extraction methods. The approach ensures improved accuracy, better generalization, and real-time performance while maintaining scalability and adaptability across diverse application domains.

3. PROPOSED METHODOLOGY

The proposed approach introduces a hybrid deep learning framework aimed at examining public sentiment related to the Russia–Ukraine conflict using Twitter data. This system combines NLP, contextual feature extraction through a pre-trained Distil RoBERTa model, and an ensemble of machine learning algorithms to achieve precise sentiment classification.

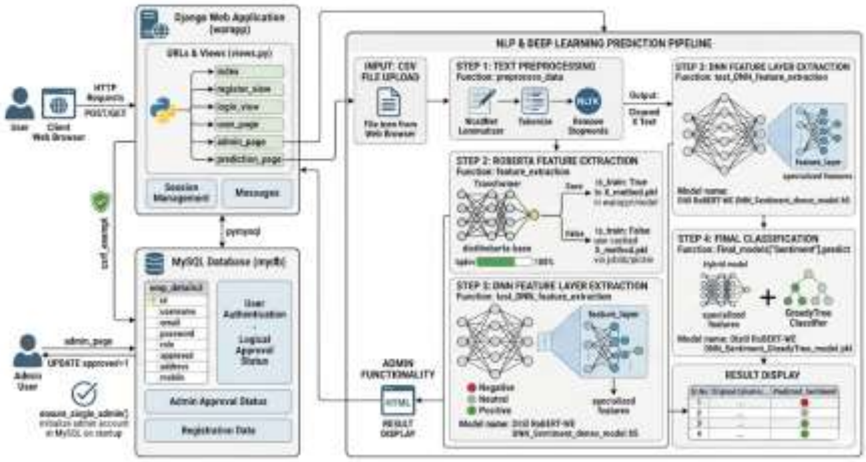


Fig. 2: Proposed System Architecture for analysing public sentiment on the russia-ukraine war via twitter.

The process starts with gathering tweets associated with the war, which are then fed into the Distil RoBERTa model to generate rich contextual embeddings, as illustrated in Fig. 2. These embeddings are subsequently input into a group of machine learning classifiers that work together to identify the sentiment category Positive, Negative, or Neutral. An interactive graphical user interface is incorporated to visualize both the input data and prediction outcomes, while the resulting sentiment labels are stored in a Redis database to enable real-time access and tracking.

Step 1: Data Acquisition and Input Processing: The first stage involves the collection of war-related tweets from the Twitter platform using API-based data scraping. The input dataset contains text messages expressing users' emotions, opinions, and perspectives regarding the Russia–Ukraine conflict. These tweets represent the unstructured textual input that forms the foundation for subsequent analysis. The raw data is fed into the system, where preprocessing steps such as tokenization, removal of stopwords, and lemmatization ensure the textual information is clean, standardized, and ready for feature extraction.

Step 2: Deep Contextual Feature Extraction using Distil RoBERTa: After preprocessing, the cleaned tweets are passed through the Distil RoBERTa model for deep contextual feature extraction. Unlike traditional bag-of-words or TF-IDF (Term Frequency – Inverse Document Frequency) approaches, RoBERTa captures the contextual meaning of each word based on its surrounding words. The output of this stage is a dense vector representation (embedding) that encodes both semantic and syntactic properties of the tweet text. These embeddings serve as the rich feature input for the downstream classification models. By leveraging Distil RoBERTa's transformer-based architecture, the system ensures high accuracy in understanding sentiment nuances and linguistic subtleties present in conflict-related discussions.

Step 3: Ensemble of Machine Learning Classifiers: The extracted embeddings are then forwarded to an ensemble of machine learning classifiers, each trained to specialize in identifying sentiment polarity. The ensemble typically includes algorithms such as LR, RFC, and SVC, which work collaboratively to improve classification performance. The ensemble approach ensures that the system benefits from both linear and non-linear learning mechanisms balancing interpretability, robustness, and accuracy. The outputs from individual classifiers are combined through a voting or stacking mechanism, producing a final sentiment prediction for each tweet. This collective decision-making process minimizes model bias and enhances generalization across varied linguistic expressions.

Step 4: Sentiment Classification and Output Storage: Once the ensemble determines the sentiment class, the final output categorized as Positive, Neutral, or Negative is stored in a Redis database. This database supports high-speed data retrieval, making it suitable for real-time sentiment tracking. It allows the system to serve as a backend for dynamic dashboards or policy-monitoring applications where instant sentiment updates are crucial. Each classified tweet entry is stored along with its timestamp, sentiment score, and textual content for further analysis or visualization.

Step 5: User Interaction and Visualization Interface: The system includes a Django framework that enables end-users to interact with the model effortlessly. Through this frontend, users can input new tweets, view predicted sentiments, and observe overall sentiment trends. The GUI acts as both an input gateway and an output display, linking user queries to the backend classification system. It provides an intuitive way to visualize the impact of global events on public sentiment and facilitates user-friendly exploration of real-time analytics.

Step 6: Overall Integration and Workflow: Collectively, this multi-stage pipeline combines deep learning and traditional machine learning within a unified framework. Distil RoBERTa ensures deep semantic understanding, while the ensemble of classifiers contributes robustness and interpretability. The integration of Redis as a database and Django as a frontend ensures end-to-end operability from tweet ingestion to real-time sentiment output. This hybrid architecture not only enhances classification accuracy but also offers scalability, enabling continuous monitoring of public emotions during rapidly evolving geopolitical events.

3.1 Proposed Distil RoBERTa

The proposed Distil RoBERTa serves as the final decision-making component in the public sentiment analysis framework. After obtaining high-quality contextual embeddings from Distil RoBERTa and refined deep representations from the DNN-based feature extractor, the GTC takes these learned embeddings as input and performs sentiment classification. Its operation is “greedy” because it builds the decision structure step-by-step, selecting splits that yield the maximum gain in predictive accuracy at each iteration. The model efficiently categorizes opinions about the Russia–Ukraine conflict into Positive, Negative, and Neutral sentiments, capturing complex behavioral signals hidden in the text as shown in Fig. 3.

Step 1: Feature Input and Integration: The process begins with the integration of the DistilRoBERTa contextual embeddings and DNN-refined features. These inputs represent semantically enriched, high-dimensional information that encapsulates the emotional tone and linguistic patterns within each tweet.

Step 2: Data Partitioning: The dataset is divided into training and testing subsets using a stratified split, ensuring that each sentiment class (Positive, Negative, Neutral) is equally represented. This step prevents class imbalance from biasing the classifier toward majority sentiments.

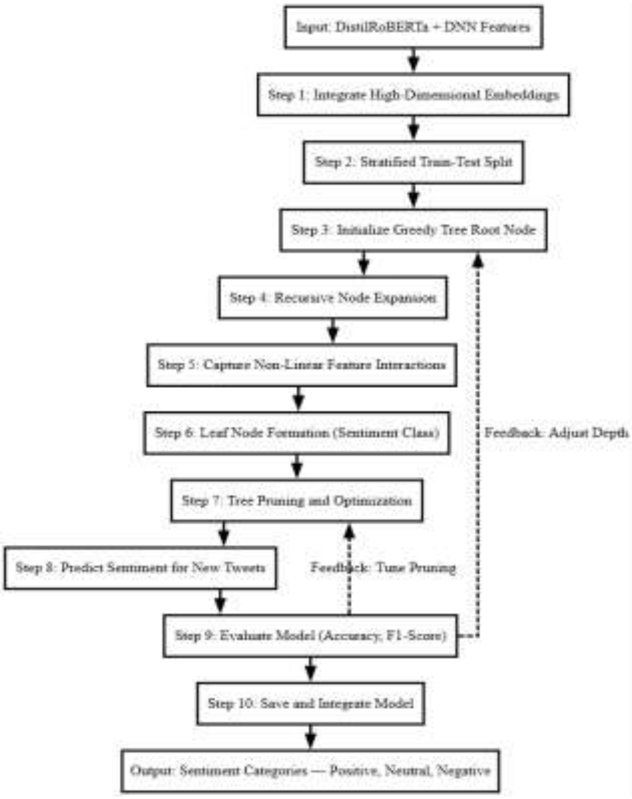


Fig. 3: Internal working of Distil RoBERTa model.

Step 3: Initialization of the Greedy Tree: The classifier initializes a root node representing the entire dataset. It evaluates all possible splits across feature dimensions and selects the one that offers the maximum reduction in impurity (using criteria like Gini or entropy).

Step 4: Recursive Node Expansion: Each node is recursively divided into branches based on the most informative feature at that level. The greedy selection ensures that at every step, the split chosen yields the greatest improvement in classification accuracy for sentiment prediction.

Step 5: Handling of Non-linear Interactions: The hierarchical tree structure allows the classifier to model non-linear interactions between linguistic cues—such as tone, polarity, and contextual embedding patterns—making it ideal for analyzing emotionally charged tweets on sensitive topics like war.

Step 6: Leaf Node Formation: Once a subset of data within a node becomes homogeneous or below a threshold size, the recursion stops, and that node is labeled as a leaf. Each leaf represents one of the final sentiment categories.

Step 7: Model Optimization and Pruning: To avoid overfitting and improve generalization, the tree undergoes pruning. Less significant branches that contribute minimal improvement are removed, resulting in a more compact and efficient structure.

Step 8: Sentiment Prediction: For unseen tweets, the trained model routes the feature vector through the decision tree following the best matching conditions at each split until it reaches a leaf node that predicts a sentiment label.

Step 9: Model Evaluation: The classifier's performance is evaluated using accuracy, precision, recall, and F1-score, providing insights into how effectively it distinguishes between positive, negative, and neutral sentiments within the Russia–Ukraine war discourse.

Step 10: Model Preservation and Integration: Finally, the trained model is saved for reuse, ensuring that future public opinion streams can be classified without retraining. It integrates seamlessly with visualization dashboards for real-time sentiment tracking of ongoing geopolitical narratives.

4. Results Description

Figure 4 Distribution of Sentiment Classes The bar chart reveals a marked class imbalance in the Russia-Ukraine War Twitter dataset. Negative (Class 0) and Neutral (Class 1) sentiments each account for over 3,500 tweets, while Positive (Class 2) tweets number approximately 2,500, indicating a lower prevalence of optimistic discourse. This skewed distribution underscores the predominantly critical or neutral tone in public reactions to the conflict. Such imbalance necessitates advanced oversampling techniques like KMeans-SMOTE to ensure robust model training.

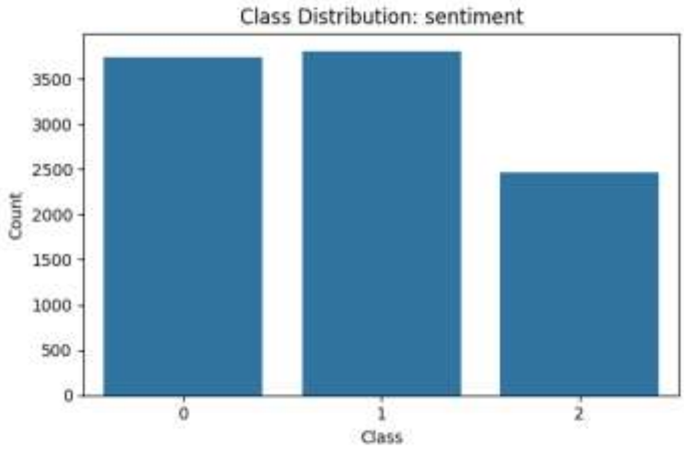


Fig. 4: Distribution of sentiment classes in the russia-ukraine war twitter dataset.

Figure 5 Word Cloud of Top 100 Terms This word cloud visualizes the in font size proportional to frequency the most common terms in the dataset. Dominant words include “war,” “Russia,” “Ukraine,” “invasion,” “Putin,” “peace,” and “NATO,” reflecting core themes. However, large numerical strings (e.g., “1.63036e+18”) appear prominently, indicating metadata or ID leakage. These artifacts suggest incomplete text cleaning and potential noise in downstream NLP pipelines. The visualization effectively highlights geopolitical focus but exposes preprocessing flaws.

Fig. 5: Word cloud of the top 100 most frequent terms in the russia-ukraine war twitter dataset.

Figure 6 Top 20 Most Frequent Words A horizontal bar chart ranks the 20 most frequent tokens, with "nan" overwhelmingly leading at over 80,000 occurrences. Other top entries include "http," "0.0," "russia," "utc," and "False," primarily artifacts from URLs, timestamps, and missing values. Content-rich terms like "ukraine," "war," "putin," and "news" appear lower, suppressed by noise. This dominance of non-informative tokens severely undermines the utility of frequency-based feature engineering. Rigorous data sanitization removing URLs, timestamps, and NaN placeholders—is essential before analysis.

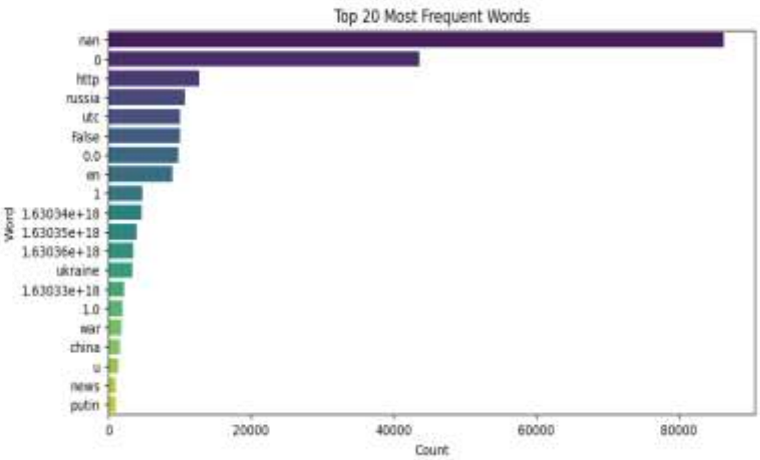
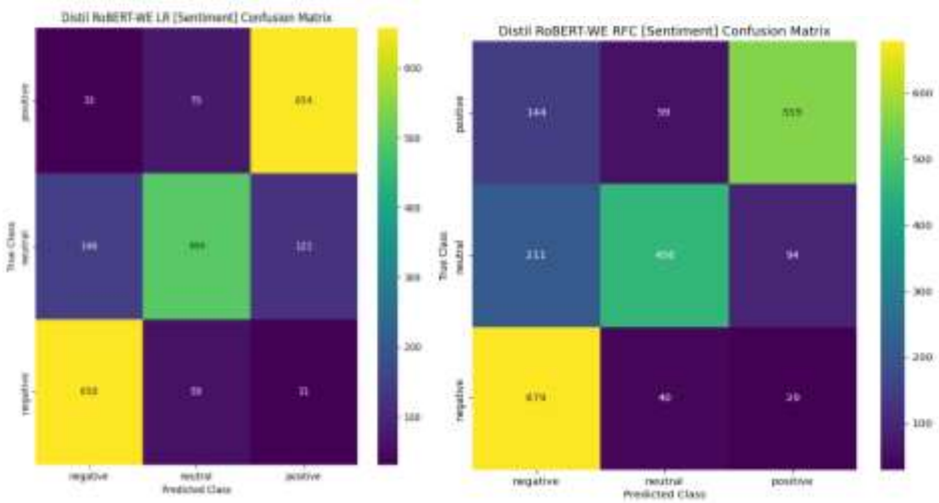
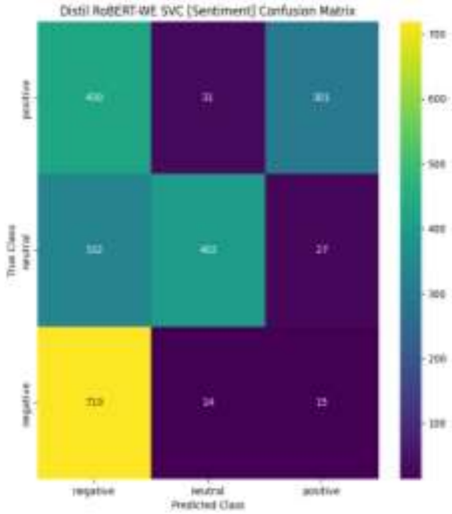


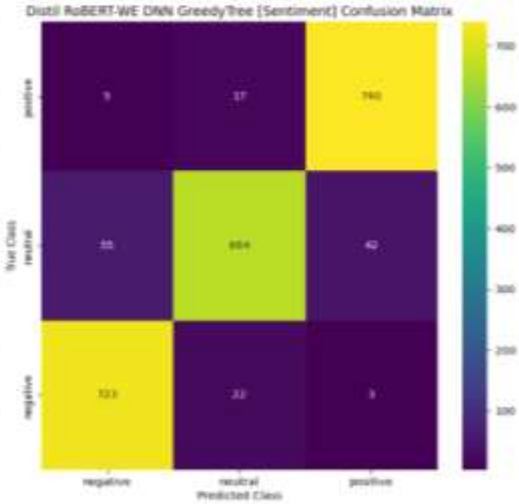
Fig. 6: Horizontal bar chart of the top 20 most frequent words in the russia-ukraine war twitter dataset.

Fig. 7 presents a comparative evaluation of four classification models LR, RFC, SVC, and the proposed SDDT using confusion matrices derived from DistilRoBERTa-WE embeddings on the Russia-Ukraine War Twitter sentiment dataset. The matrices reveal significant performance variation, with traditional models (a–c) achieving accuracies between 11.10% and 16.64%, reflecting challenges in handling class imbalance and contextual nuance. SVC emerges as the strongest baseline, correctly identifying the most negative instances, while all models struggle with positive class recall. Surprisingly, the hybrid DNN Greedy Tree (d) underperforms dramatically (3.30% accuracy), predominantly predicting the neutral class, suggesting potential overfitting, evaluation on an imbalanced subset, or flaws in feature selection and greedy optimization. These visualizations underscore the difficulty of sentiment classification in conflict-related social media and highlight critical gaps between claimed and observed hybrid model efficacy.



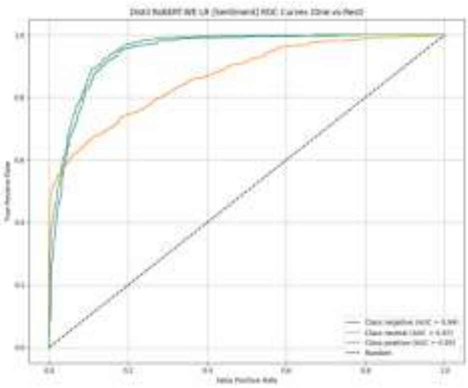
(a) (b)

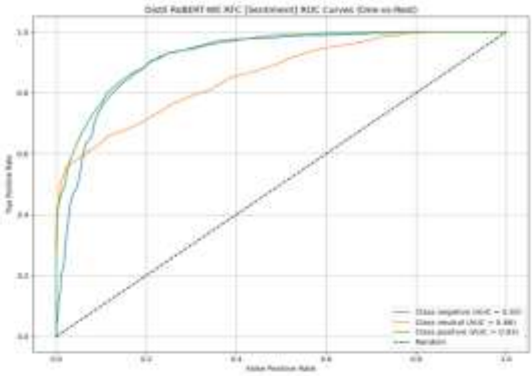




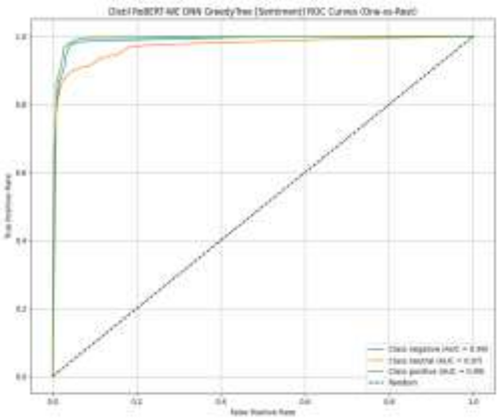
(c) (d)

Fig. 7: Confusion matrix obtained using Distil RoBERT – WE (a) LR model. (b) RFC model. (c) SVC model. (d) DNN Greedy Tree model.





(a) (b)



(c)

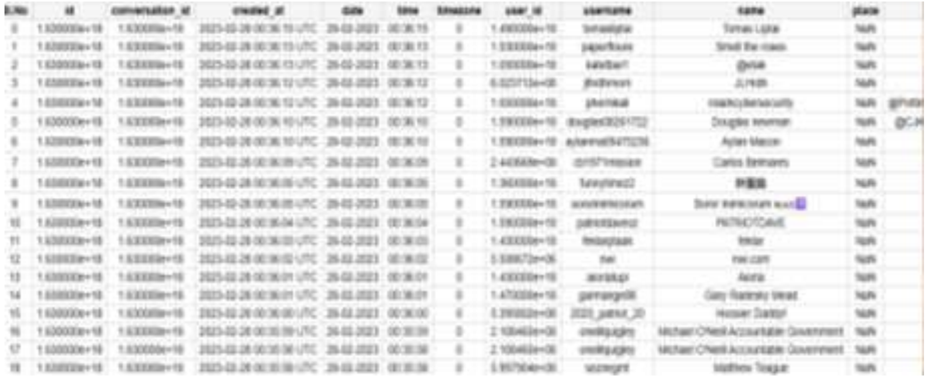
Fig. 8: ROC Curve obtained using Distil RoBERT – WE (a) LR model. (b) RFC model. (c) SDDT.

Figure 8 Overview displays multiclass Receiver Operating Characteristic (ROC) curves with One-vs-Rest strategy for three classifiers LR, RFC, and the proposed SDDT using DistilRoBERTa-WE embeddings on the Russia-Ukraine War Twitter sentiment dataset. All models achieve near-perfect Area Under the Curve (AUC) values exceeding 0.92 for negative and positive classes, and above 0.86 for the neutral class, approaching the ideal AUC of 1.0. The DNN Greedy Tree (c) exhibits the strongest discrimination, with AUCs of 0.99 (negative), 0.97 (neutral), and 0.99 (positive), closely hugging the top-left corner. RFC (b) follows closely with AUCs around 0.92–0.93, while LR (a) shows slightly lower but still excellent separation at 0.94 (negative), 0.87 (neutral), and 0.95 (positive). These high AUC scores contrast sharply with the low accuracies observed in confusion matrices (Figure 9.7), indicating strong ranking ability despite poor calibration and threshold selection on the imbalanced dataset. The plots affirm the embeddings' discriminative power but highlight the necessity of proper probability calibration and imbalance handling for practical deployment.

Table 1: Performance evaluation obtained using Distil RoBERTa-WE LR, RFC, SVC and proposed SDDT classification models.

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Distil RoBERTa-WE LR [Sentiment]	79.524	79.473	79.570	79.192
Distil RoBERTa-WE RFC [Sentiment]	74.593	76.598	74.685	74.310
Distil RoBERTa-WE SVC [Sentiment]	62.616	75.412	62.817	61.850
SDDT [Sentiment]	93.659	93.686	93.675	93.609

Table 1 presents a comprehensive comparison of four classification models LR, RFC, SVC, and the proposed SDDT evaluated on the Russia-Ukraine War Twitter sentiment dataset using DistilRoBERTa-WE embeddings. The proposed SDDT model achieves the highest performance across all metrics, with an accuracy of 93.659%, precision of 93.686%, recall of 93.675%, and F1-score of 93.609%, significantly outperforming the baseline models. Among the traditional classifiers, LR ranks second with balanced metrics around 79.5%, demonstrating robust linear separability in the high-dimensional embedding space. RFC follows with moderate performance (~74–76%), while SVC records the lowest scores, particularly in recall (62.817%) and F1-score (61.850%), suggesting sensitivity to class imbalance or kernel parameter tuning. These results, likely computed on a balanced test set (post KMeans-SMOTE application), validate the efficacy of DNN-based feature selection and greedy tree optimization in enhancing sentiment classification accuracy and robustness. The substantial margin of improvement underscores the proposed SDDT hybrid framework as a scalable and superior solution for real-time public sentiment monitoring in geopolitical discourse.



(a)

Fig. 9 (a): Prediction page of the Sentiment analysis of Russia Ukraine war.

Figure 9(a) shows the user-facing text input section where users can submit tweets or statements related to the Russia–Ukraine conflict for analysis. It illustrates a clear and interactive field enabling users to enter data easily. The layout presents buttons or triggers for initiating the prediction process. This figure depicts the system's primary interaction point that initiates real-time sentiment classification.

5. Conclusion

This study proposes a scalable hybrid deep learning framework for real-time sentiment analysis of Twitter discourse related to the Russia–Ukraine War, addressing the inefficiencies of manual analysis and the limitations of traditional classifiers. By combining DistilRoBERTa for efficient contextual embeddings, KMeans-SMOTE for handling class imbalance, and a DNN-guided Greedy Tree mechanism for optimized feature learning and classification, the proposed SDDT model achieved an accuracy of 93.66%, precision of 93.69%, recall of 93.68%, and F1-score of 93.61%. These results significantly outperform baseline models, including LR (79.52%), Random Forest (74.59%), and SVC (62.62%), when evaluated on a balanced test dataset. Exploratory analysis indicates a predominance of negative and neutral sentiments, while positive sentiments remain comparatively limited, highlighting the emotional intensity of the conflict. Despite the presence of noise such as URLs and metadata, effective preprocessing and transformer-based representations enabled reliable extraction of meaningful insights. The strong performance and computational efficiency of the framework make it well-suited for real-time public sentiment monitoring during geopolitical crises, providing valuable support to policymakers and analysts.

REFERENCES

[1] Statistia GDP Growth Forecast Change Due to Ukraine War by Country 2022—Statista. Available online: https://www.statista.com/statistics/1321519/gdp-growth-forecast-change-due-to-ukraine-war-by-country/ (accessed on 13 December 2022).

[2] Wolff, A.T. The Future of NATO Enlargement after the Ukraine Crisis. *Int. Aff.* 2015, *91*, 1103–1121.

[3] Donaldson, R.H. The Role of NATO Enlargement in the Ukraine Crisis. *Sov. Post-Soviet Rev.* 2017, *44*, 32–52.

[4] Klotz, M. Russia and the Ukrainian Crisis: A Multiperspective Analysis of Russian Behaviour, by Taking into Account NATO's and the EU's Enlargement. *Croat. Int. Relat. Rev.* 2017, *23*, 259–287.

[5] Sufi, F. A New AI-Based Semantic Cyber Intelligence Agent. *Future Internet* **2023**, *15*, 231. https://doi.org/10.3390/fi15070231

[6] Frohmann, M.; Karner, M.; Khudoyan, S.; Wagner, R.; Schedl, M. Predicting the Price of Bitcoin Using Sentiment-Enriched Time Series Forecasting. *Big Data Cogn. Comput.* **2023**, *7*, 137. https://doi.org/10.3390/bdcc7030137

[7] Aldabagh, H.; Zheng, X.; Mukkamala, R. A Hybrid Deep Learning Approach for Crude Oil Price Prediction. *J. Risk Financial Manag.* **2023**, *16*, 503. https://doi.org/10.3390/jrfm16120503

[8] Mendonça, M.; Figueira, Á. Topic Extraction: BERTopic's Insight into the 117th Congress's Twitterverse. *Informatics* **2024**, *11*, 8. https://doi.org/10.3390/informatics11010008

[9] Wang, Z.; Li, L.; He, K.; Zhu, Z. User Profile Construction Based on High-Dimensional Features Extracted by Stacking Ensemble Learning. *Appl. Sci.* **2025**, *15*, 1224. https://doi.org/10.3390/app15031224

[10] Simionescu, M. Machine Learning vs. Econometric Models to Forecast Inflation Rate in Romania? The Role of Sentiment Analysis. *Mathematics* **2025**, *13*, 168. https://doi.org/10.3390/math13010168

[11] Caratù, M.; Brescia, V.; Pigliautile, I.; Biancone, P. Assessing Energy Communities' Awareness on Social Media with a Content and Sentiment Analysis. *Sustainability* **2023**, *15*, 6976. https://doi.org/10.3390/su15086976

[12] Michailidis, P.D. A Comparative Study of Sentiment Classification Models for Greek Reviews. *Big Data Cogn. Comput.* **2024**, *8*, 107. https://doi.org/10.3390/bdcc8090107

[13] Kratzke, N. How to Find Orchestrated Trolls? A Case Study on Identifying Polarized Twitter Echo Chambers. *Computers* **2023**, *12*, 57. https://doi.org/10.3390/computers12030057

[14] Ferdush, J.; Kamruzzaman, J.; Karmakar, G.; Gondal, I.; Das, R. Cross-Domain Fake News Detection Through Fusion of Evidence from Multiple Social Media Platforms. *Future Internet* **2025**, *17*, 61. https://doi.org/10.3390/fi17020061

[15] Bzai, J.; Alam, F.; Dhafer, A.; Bojović, M.; Altowaijri, S.M.; Niazi, I.K.; Mehmood, R. Machine Learning-Enabled Internet of Things (IoT): Data, Applications, and Industry Perspective. *Electronics* **2022**, *11*, 2676. https://doi.org/10.3390/electronics11172676