DOI: 10.15514/ISPRAS-2025-37(3)-11

Enhanced Text Classification Using DistilBERT with Low-Rank Adaptation: A Comparative Study

B.D. Abodo Eloundou, ORCID: 0009-0009-1845-5867 <abodobrice@hotmail.com>
W. Quanyu, ORCID: 0009-0000-1059-7312 <272561@edu.itmo.ru>

AI Talent Hub, ITMO University,
49, Kronverksky pr., Saint Petersburg, 197101, Russia.

Abstract. In this article, we delve into the task of sentiment analysis applied to news articles covering sanctions against Russia, with a specific focus on secondary sanctions. With geopolitical tensions influencing global affairs, understanding the sentiment conveyed in news about sanctions is crucial for policymakers, analysts, and the public alike. We explore the challenges and nuances of sentiment analysis in this context, considering the linguistic complexities, geopolitical dynamics, and data biases inherent in news reporting. Leveraging natural language processing techniques and machine learning models, including Large Language Models (LLM), 1D Convolutional Layer (Conv1D), and Feed-Forward Networks (FFN), we aim to extract sentiment insights from news articles. Our analysis provides valuable perspectives on public opinion, market reactions, and geopolitical trends. Through our work, we seek to illuminate the sentiment landscape surrounding sanctions against Russia and their broader implications.

Keywords: text classification; sentiment analysis; sanctions; distilbert; lora; conv1d; feed-forward networks.

For citation: Abodo Eloundou B.D., Quanyu W. Enhanced Text Classification Using DistilBERT with Low-Rank Adaptation: A Comparative Study. Tr*udy ISP RAN/Proc. ISP RAS*, vol. 37, issue 3, 2025, pp. 159-170. DOI: 10.15514/ISPRAS-2025-37(3)-11.

Acknowledgements. This study was supported by the AI Talent Hub of ITMO University and the Russian Academy of Foreign Trade. The Russian Academy of Foreign Trade provided the dataset for the research.

Расширенная классификация текста с помощью модели DistilBERT с адаптацией низкого ранга LoRa: сравнительное исследование

Б.Д. Абодо Элунду, ORCID: 0009-0009-1845-5867 <abodobrice@hotmail.com>
В. Цюаньюй, ORCID: 0009-0000-1059-7312 <272561@edu.itmo.ru>

AI Talent Hub, Университет ИТМО,
Россия, 197101, г. Санкт-Петербург, Кронверкский пр., д. 49.

Аннотация. В данной статье мы рассматриваем задачу анализа тональности новостных статей, посвященных санкциям против России, с особым вниманием к вторичным санкциям. С учетом геополитической напряженности, влияющей на мировые события, понимание тональности новостей о санкциях имеет важное значение для политиков, аналитиков и широкой общественности. Мы изучаем вызовы и особенности анализа тональности в данном контексте, учитывая языковые сложности, геополитическую динамику и предвзятость данных в новостных материалах. Используя методы обработки естественного языка и модели машинного обучения, включая большие языковые модели (LLM), одномерные сверхточные слои (Conv1D) и полно связные нейросети (FFN), мы стремимся извлечь информацию о тональности из новостных статей. Наш анализ предоставляет ценные сведения об общественном мнении, реакции рынков и геополитических тенденциях. В рамках данной работы мы стремимся осветить тональный ландшафт, связанный с санкциями против России, и их более широкие последствия.

Ключевые слова: классификация текста; анализ настроений; санкции; модель distilbert; адаптация низкого ранга lora; одномерные сверхточные слои conv1d; полно связные нейросети feed-forward networks.

Для цитирования: Абодо Элунду Б.Д., Цюаньюй В. Расширенная классификация текста с помощью DistilBERT с адаптацией LoRa: сравнительное исследование. Труды ИСП РАН, том 37, вып. 3, 2025 г., стр. 159–170 (на английском языке). DOI: 10.15514/ISPRAS–2025–37(3)–11.

Благодарности. Это исследование было проведено при поддержке AI Talent Hub Университета ITMO и Российской академии внешней торговли. Российская академия внешней торговли предоставила набор данных для исследования.

1. Introduction

The ever-evolving geopolitical landscape presents policymakers, analysts, and the public with a constant need to understand the sentiment surrounding critical global events. In the case of sanctions against Russia, particularly secondary sanctions that impact various countries, gauging public perception is crucial. Sentiment analysis emerges as a powerful tool in this scenario, offering insights into the emotional undercurrents of news articles covering these sanctions.

Recent research by Kim et al. (2021) emphasizes the importance of understanding the political dynamics of sanctions [1]. Their comparative study of sanctions against Russia and North Korea highlights the varied effects and complexities involved. Sentiment analysis can contribute to this understanding by revealing public opinion towards these dynamics.

This article explores the application of sentiment analysis, specifically leveraging a Large Language Model (LLM), to analyze news articles focused on sanctions against Russia. We delve into the complexities of sentiment analysis within this context, acknowledging the challenges posed by linguistic nuances (e.g., sarcasm, ambiguity) [2], the fluid nature of geopolitics [3], and potential biases inherent in news reporting [5]. Through the application of natural language processing techniques and machine learning models, we aim to extract valuable sentiment-based insights from a vast corpus of news articles. This analysis will provide a deeper understanding of public opinion, market reactions, and broader geopolitical trends surrounding these sanctions. Ultimately, this article seeks to shed light on the sentiment landscape and its potential implications for the effectiveness and impact of sanctions against Russia.

2. Related Works

The field of text classification has seen significant advancements over the years, particularly with the advent of deep learning models. Traditional machine learning methods, such as Naive Bayes and Support Vector Machines (SVM), laid the foundation for text classification tasks [6]. However, with the rise of neural networks, models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have shown superior performance due to their ability to capture complex patterns in text data [7, 8].

2.1 Classification

Early works in text classification employed methods like Naive Bayes, SVM, and logistic regression, which relied heavily on feature engineering [9]. With the emergence of deep learning, CNNs and RNNs have been extensively used for text classification. These models automatically learn features from raw text, reducing the need for manual feature extraction [10, 11]. The introduction of Transformer models, particularly BERT and its variants, has revolutionized text classification tasks. These models use attention mechanisms to capture contextual relationships in text more effectively than traditional RNNs [12, 13]. DistilBERT, a lighter version of BERT, maintains its performance while being more efficient [14]. The use of Low-Rank Adaptation (LoRA) further enhances its adaptability and performance on specific tasks [15].

2.2 Challenges

Despite significant progress, there are several challenges associated with existing solutions. Transformer models, while powerful, are computationally expensive and require significant resources for training and inference [16]. Many models tend to overfit specific datasets, leading to poor generalization on unseen data [17]. Additionally, deep learning models, including transformers, often act as “black boxes”, making it difficult to interpret their decisions [18].

2.3 Distinction

This study aims to address some of the limitations of existing approaches by leveraging DistilBERT with Low-Rank Adaptation (LoRA) for text classification. By using DistilBERT, we achieve a balance between model performance and computational efficiency [14]. Incorporating LoRA allows for more efficient fine-tuning on specific tasks, improving performance without substantial increases in computational cost [15]. Furthermore, we provide a thorough comparison of DistilBERT with traditional Conv1D and Feedforward Neural Network (FFN) models, highlighting the strengths and weaknesses of each approach. In conclusion, our work builds upon the strengths of previous research while addressing some of their key limitations, providing a comprehensive evaluation of modern text classification techniques.

2.4 Economic Impact of Sanctions

The economic impact of sanctions has been extensively studied, with research highlighting both the intended and unintended consequences on the target nation’s economy. Hufbauer et al. (2007) provide a comprehensive analysis of economic sanctions, discussing their effectiveness and the various mechanisms through which they influence the targeted economy [19]. Their work illustrates how sanctions can lead to trade reductions, financial instability, and long-term economic decline. In the specific context of sanctions against Russia, Dreger et al. (2016) analyze the economic performance of Russia under sanctions and find significant impacts on trade patterns and financial markets, demonstrating the broad and profound economic consequences of such measures [20].

2.5 Political Dynamics and Framing Analysis

Understanding the political dynamics of sanctions is crucial for comprehending their overall effectiveness and reception. Kim et al. (2021) provide a comparative analysis of sanctions against Russia and North Korea, highlighting the complex political dynamics and varying effects of these sanctions [1]. Their research underscores the importance of considering the political context when evaluating the impact of sanctions. Additionally, the framing of sanctions in news media plays a vital role in shaping public perception and political discourse. Entman (2007) discusses the concept of framing bias and how media framing can influence public opinion and policymaking processes [21]. By analyzing the media narratives around sanctions, researchers can gain insights into how different framings affect public sentiment and political outcomes.

2.6 Low-Rank Adaptation (LoRA)

Low-Rank Adaptation (LoRA) has emerged as a promising technique in the field of machine learning, particularly for efficiently fine-tuning large language models. Hu et al. (2021) introduce LoRA as a method that reduces the number of trainable parameters by decomposing weight matrices into low-rank factors, thereby improving training efficiency without significantly compromising model performance [15]. This technique is particularly advantageous in scenarios with limited computational resources, enabling the practical application of large-scale models in tasks such as sentiment analysis. LoRA allows for the adaptation of large pre-trained models to specific tasks with minimal computational overhead, making it a valuable tool in the NLP toolkit.

2.7 Feed-Forward Networks (FFN)

Feed-Forward Networks (FFN) are a fundamental component of many neural network architectures and are widely used in various machine learning tasks due to their simplicity and effectiveness. Goodfellow et al. (2016) provide a detailed overview of FFNs, describing their architecture, training processes, and applications [4]. FFNs consist of multiple layers of interconnected neurons, where each layer transforms the input data through a series of weighted linear combinations followed by non-linear activations. In sentiment analysis, FFNs are often employed for their ability to model complex patterns in textual data, capturing the relationships between words and their contextual meanings to produce accurate sentiment predictions.

2.8 Convolutional Neural Networks (Conv1D)

Convolutional Neural Networks (Conv1D) are particularly effective in natural language processing tasks due to their ability to capture local patterns in sequential data. Kim (2014) demonstrates the application of Conv1D for sentence classification, showcasing its potential to extract meaningful features from text data [10]. Conv1D models apply convolutional filters to the input text, enabling the identification of n-gram features and local dependencies. This makes them well-suited for tasks like sentiment analysis, where understanding the context and relationships between words is crucial. The use of Conv1D in text classification has shown promising results, highlighting its role as a powerful tool in the NLP landscape.

3. Experiment Setup and Methodology

3.1 Dataset Description

To evaluate the effectiveness of DistilBERT with Low-Rank Adaptation for sentiment analysis in news articles about sanctions against Russia, we utilized a dataset compiled by the Russian Foreign Trade Academy. This dataset comprises text data related to various geopolitical and economic events concerning sanctions on Russia, and each article is labeled for sentiment analysis. Articles were sourced from reputable news outlets and encompass a range of languages, including Italian, German, Chinese, Japanese, Spanish, and Russian. Non-Russian articles were translated into Russian, creating a monolingual dataset for analysis.

To be included in the dataset, each article needed to explicitly address matters related to Russia post-February 2022, with a focus on politics, policies, and economic issues. The dataset comprises 583 articles divided into three sentiment categories: 218 marked as positive, 185 as negative, and 180 as neutral. Positive articles highlight ineffective or disruptive sanctions, new trade opportunities, or potential advantages for Russia; neutral articles are objective and lack value judgments; and negative articles discuss the adverse effects of sanctions on Russia or note new restrictive measures imposed by other countries.

To enhance the dataset, a back-translation augmentation technique was applied. This involved translating Russian texts into intermediate languages, such as English, and then back into Russian. This method helped diversify the dataset by generating additional textual variations, which were then labeled according to the original sentiment categories.

The augmented dataset provides a robust basis for training and evaluating sentiment analysis models, facilitating a comprehensive examination of model performance on real-world data related to sanctions and their geopolitical impact.

3.2 Preprocessing Steps

The raw text data underwent several preprocessing steps to ensure consistency and quality. These steps included:

- Converting text to lowercase.
- Removing HTML tags, email addresses, digits, punctuation, and newlines.
- Collapsing multiple spaces into a single space.
- Normalizing specific Russian characters.

3.3 Model Architecture and Training

We implemented three distinct models for comparison: DistilBERT with Low-Rank Adaptation (LoRA), Convolutional Neural Networks (Conv1D), and Feed-Forward Networks (FFN). Each model was meticulously designed to capture different aspects of textual data for effective sentiment classification.

DistilBERT with Low-Rank Adaptation (LoRA): DistilBERT, a distilled version of BERT, was fine-tuned using the Low-Rank Adaptation (LoRA) technique to enhance its adaptability for the sentiment analysis task. LoRA reduces the number of trainable parameters by decomposing weight matrices into low-rank factors, thereby improving training efficiency without significantly compromising performance.

Model Architecture: The DistilBERT model architecture utilized in this study is as follows:

- **Embeddings:**
  - **Word Embeddings:** Embedding layer with vocabulary size of 119,547 and embedding dimension of 768.
  - **Position Embeddings:** Embedding layer with a maximum sequence length of 512 and embedding dimension of 768.
  - **Layer Normalization:** Applied with ϵ = 1e-12.
  - **Dropout:** Applied with a probability of 0.1.
- **Transformer Layers:**

Six Transformer blocks, each comprising:

- **Multi-Head Self-Attention:** Includes dropout with p = 0.1, and linear layers for queries (q lin), keys (k lin), values (v lin), and output (out lin), each with input and output dimensions of 768.
- **Layer Normalization:** Applied before attention and feed-forward networks.
- **Feed-Forward Network (FFN)**: Consists of two linear layers with a GELU activation function between them, expanding from 768 to 3072 dimensions and back to 768.
- **Dropout:** Applied with a probability of 0.1 after each sub-layer.
- **Pre-Classifier:** Linear layer transforming 768 features to 768.
- **Classifier:** Linear layer transforming 768 features to 3 classes (positive, negative, neutral).
- **Dropout**: Applied with a probability of 0.2 before the classifier.

Hyperparameters: The DistilBERT model was trained with the following hyperparameters:

- **Learning Rate:** $4.836 \times 10{-}5$.
- **Batch Size:** 32.
- **Number of Training Epochs**: 7.
- **LoRA Parameters**:
  - **LoRA Rank (r):** 8.
  - **LoRA Alpha:** 32.
  - **LoRA Dropout:** 0.1.
- **Weight Decay:** 0.08.

LoRA Configuration: The LoRA configuration applied to DistilBERT is defined as follows:

- **PEFT Type:** LoRA.
- **Task Type:** Sequence Classification (SEQ CLS).
- **Target Modules:** attention.k_lin, attention.q_lin, attention.v_lin.
- **Fan In Fan Out:** False.
- **Bias:** None.
- **Use RSLora:** False.
- **Initialization:** LoRA weights initialized.
- **Runtime Configuration:** Ephemeral GPU offload disabled.

Training Procedure: The model was trained using the following procedure:

1. **Data Loading:** The preprocessed dataset was loaded and split into training and validation sets.
2. **Training:** The model was trained for 7 epochs with a batch size of 32.
3. **Evaluation:** After each epoch, the model's performance was evaluated on the validation set to monitor for overfitting.
4. **Checkpointing:** The best-performing model based on validation F1-score was saved for final evaluation.

Convolutional Neural Networks (Conv1D): Convolutional Neural Networks (Conv1D) have proven to be highly effective in natural language processing tasks by capturing local patterns and dependencies in sequential data. Kim (2014) demonstrated the application of Conv1D for sentence classification, highlighting its capability to extract meaningful n-gram features from text [10]. In our study, the Conv1D model architecture was designed with multiple convolutional layers, each employing a set of filters to capture different n-gram representations.

The model incorporates a max-pooling layer following each convolutional layer to reduce dimensionality and highlight the most salient features. Additionally, dropout layers were implemented to prevent overfitting and enhance generalization.

The final layers consist of a fully connected dense layer leading to the softmax activation function for classification.

Based on our experiments, the Conv1D model utilized the following hyperparameters:

- **Batch Size:** 32.
- **Number of Epochs:** 14.
- **Optimizer:** Adam.
- **Dropout Rate:** 0.5.
- **Filter Sizes:** 256, 128, 128.
- **The kernel size:** 7, 5, 3.
- **For activation:** relu.

These settings were chosen based on preliminary grid search experiments aimed at optimizing model performance while maintaining computational efficiency. The Conv1D model demonstrated robust performance in capturing local textual features, contributing significantly to the overall sentiment classification accuracy.

Feed-Forward Network (FFN): The Feed-Forward Network (FFN) model was developed to provide a baseline comparison against more complex architectures like DistilBERT and Conv1D.

Model Architecture: The FFN model architecture utilized in this study is as follows:

- **Input Layer:** Accepts input sequences of tokens with a maximum length of max_len.
- **Embedding Layer:**
- **Word Embeddings:** Embedding layer with a vocabulary size of vocab_size and an embedding dimension of 16.
- **Flatten Layer:** Flattens the embedded input to a one-dimensional vector.
- **Dense Layers:**
  - **Dense Layer 1:** 128 neurons with ReLU activation.
  - **Dense Layer 2:** 128 neurons with ReLU activation.
  - **Dense Layer 3:** 128 neurons with ReLU activation.
- **Dropout Layer:** Applied with a dropout rate of 0.5 to prevent overfitting.
- **Output Layer:** Dense layer with 3 neurons (corresponding to the sentiment classes) without activation (logits).

Hyperparameters: The FFN model was trained with the following hyperparameters:

- **Batch Size:** 64.
- **Number of Training Epochs:** 14.
- **Optimizer:** Adam.
- **Loss Function:** Sparse Categorical Crossentropy.
- **Metrics:** Sparse Categorical Accuracy.

Training Procedure: The FFN model was trained using the following procedure:

1. **Data Loading:** The preprocessed dataset was loaded and split into training and validation sets.
2. **Training:** The model was trained for 14 epochs with a batch size of 64.

3. **Evaluation:** After each epoch, the model's performance was evaluated on the validation set to monitor for overfitting.
4. **Checkpointing:** The model achieving the highest validation accuracy was saved for final evaluation.

To assess the performance of each model, we utilized standard evaluation metrics for text classification tasks, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of the models' effectiveness in correctly identifying the sentiment of news articles.

4. Results and Discussion

4.1 Performance Comparison

Our experimental results indicate that each model – FFN, Conv1D, and DistilBERT – demonstrates varying strengths across the different evaluation metrics. *Table 1* compares precision, recall, f1-score, and accuracy for each model, providing a detailed view of their performance on sentiment classification.

Table 1. Performance Metrics of FFN, Conv1D, and Distil-BERT Models.

Model (Weighted Avg)	Precision	Recall	F1-score	Accuracy
FFN	0.84	0.796	0.80	0.796
Conv1D	0.80	0.78	0.787	0.78
Distil-BERT	0.83	0.82	0.82	0.82

Weighted averages are calculated based on the support for each class.

4.2 Analysis

Feedforward Neural Network (FFN): The FFN model achieved a weighted average precision of 0.8404, a recall of 0.7969, and an f1-score of 0.8045. It performed particularly well on neutral sentiments, achieving high precision and f1-scores for both positive and neutral categories. This indicates that the FFN model has strong generalization capabilities, effectively distinguishing between sentiment categories with good balance.

Conv1D Model: The Conv1D model demonstrated a weighted average precision of 0.8043, recall of 0.7812, and an f1-score of 0.7873. It showed robust performance in recognizing the positive sentiment category but faced challenges with neutral sentiments. This suggests that Conv1D may struggle to fully capture nuances between neutral and other sentiments, potentially due to its reliance on local text patterns.

DistilBERT Model: The DistilBERT model achieved the highest accuracy among the three models at 82%, with a weighted average precision of 0.83 and f1-score of 0.82. Notably, it performed best on the negative category with a precision of 0.89 and an f1-score of 0.88, indicating it is particularly effective at capturing this sentiment. While DistilBERT exhibited strong recall across all categories, it showed a slightly lower precision for positive sentiments (0.70), suggesting room for further fine-tuning in this category.

Comparison and Insights: Across all three models, FFN and DistilBERT exhibited higher overall precision and f1-scores compared to Conv1D. While FFN achieved strong performance on all sentiment categories, DistilBERT excelled in classifying negative sentiments, which may make it particularly useful for tasks where identifying negative sentiment is critical. Conv1D, while competitive, demonstrated lower performance on neutral sentiments, which highlights the importance of selecting models based on the specific sentiment needs of a task.

The confusion matrices indicate that FFN and DistilBERT were more effective in distinguishing between all sentiment classes, with FFN showing a slight advantage in generalization across categories. DistilBERT, on the other hand, demonstrated superior precision in identifying negative sentiment, which suggests its suitability for contexts where distinguishing between negative and other sentiments is particularly valuable.

4.3 Comparison with Baseline Models

To assess the effectiveness of our models, the performance of the FFN, Conv1D, and DistilBERT models was compared to a simple baseline model. Both the FFN and DistilBERT models outperformed the baseline in all evaluation metrics, demonstrating the advantages of using more advanced architectures for sentiment analysis tasks. DistilBERT, in particular, showed competitive performance with less computational overhead due to its reduced model size.

4.4 Discussion

Alignment with Ground Truth Data: Our model's performance was evaluated using standard metrics, including accuracy, precision, recall, and F1-score, which provide quantitative measures of how well the model's predictions align with the annotated sentiment labels in the dataset, representing the "ground truth." Confusion matrices for each model (FFN, Conv1D, and DistilBERT) were analyzed to gain insight into areas of strong performance and potential challenges. For example, the DistilBERT model exhibited high precision and recall for negative sentiment, indicating a strong alignment with the ground truth for this category.

Additionally, misclassification instances were examined to identify recurring error patterns, such as neutral articles being incorrectly classified as positive or negative. These misalignments suggest areas for improvement in fine-tuning or feature extraction, as they often reflect challenges in handling nuanced linguistic cues.

To ensure the dataset's annotations remained consistent, each article was labeled by students of the Russian Foreign Trade Academy based on specific guidelines for categorizing sentiments as positive, negative, or neutral. Cross-validation was employed to confirm that the models generalized effectively across different data subsets, minimizing overfitting to specific patterns or biases present in the training set.

Real-World Scenario Performance: In practical applications, our model could be deployed on live news streams or a constantly updated corpus of articles, enabling near real-time sentiment analysis for geopolitical events. Given that it was trained on diverse topics related to sanctions against Russia, it is expected to perform reliably in detecting sentiment in similar contexts. Specifically, the DistilBERT model's high accuracy for negative sentiment suggests it would be particularly valuable in scenarios that require identifying critical developments, such as new sanctions, making it useful for policymakers, financial analysts, and media organizations.

However, real-world texts often contain linguistic nuances such as sarcasm, idiomatic expressions, and culturally specific references that may pose challenges for model accuracy. Although the model shows strong alignment with the ground truth in the dataset, certain cases may affect its performance due to these nuances. To enhance real-world accuracy, future work could integrate sentiment-shifting indicators (e.g., sarcasm markers) or develop a sentiment lexicon tailored to news. Additionally, since the training data includes both translated and native Russian texts, subtle differences from translation might impact real-world Russian news analysis. Expanding the dataset with more native language sources could help capture these linguistic intricacies.

Adaptation to Evolving Contexts: Given that geopolitical events and news content are dynamic, the model's performance may vary over time as the nature of discussions around sanctions changes. Regular retraining with updated data is essential for maintaining accuracy and alignment with the current political and economic context.

To further enhance adaptability, implementing continual learning or online learning methods could allow the model to update itself with new information in near real-time. This would improve its relevance and utility in ongoing geopolitical analysis, ensuring the model remains a reliable tool for sentiment analysis as global events continue to evolve.

5. Conclusion

In this study, we conducted a comprehensive evaluation of three models—Feed-Forward Networks (FFN), Convolutional Neural Networks (Conv1D), and DistilBERT with Low-Rank Adaptation (LoRA)—for sentiment analysis on a dataset of 583 news articles concerning sanctions against Russia. Each model exhibited distinct strengths:

- **FFN:** Achieved the best overall balance in performance metrics, demonstrating strong generalization across all sentiment categories.
- **Conv1D:** Showed particular competency in capturing positive sentiments, effectively identifying n-gram features pertinent to this category.
- **DistilBERT with LoRA:** Excelled in identifying negative sentiments, leveraging its deep contextual understanding to accurately classify adverse sentiments.

The detailed hyperparameter optimization and inclusion of representative textual examples enhanced the reproducibility and interpretability of our findings. Additionally, the back- translation data augmentation technique contributed to the robustness of the models by introducing diverse linguistic variations.

Future Work:

- **Model Refinement:** Further refine the DistilBERT with LoRA model to enhance its classification capabilities for positive and neutral sentiments, potentially integrating additional linguistic features or advanced fine-tuning techniques.
- **Architectural Enhancements:** Explore hybrid models that combine the strengths of Conv1D and Transformer-based architectures to capture both local and global textual patterns.
- **Expanded Data Augmentation:** Implement more sophisticated data augmentation strategies, such as synonym replacement and contextual augmentation, to further diversify the training dataset.
- **Cross-Domain Application:** Apply the developed models to other geopolitical contexts and languages to evaluate their generalizability and adaptability.
- **Real-Time Analysis:** Develop a pipeline for real-time sentiment analysis of live news streams, facilitating timely insights for policymakers and analysts.

Our findings provide valuable insights for sentiment analysis applications within geopolitical and economic domains, highlighting the importance of model selection and hyperparameter optimization in accurately capturing public sentiment. Understanding these sentiments can significantly inform policy decisions, market strategies, and broader geopolitical analyses.

References

[1]. Kim, J., Weiss, J., and Wilensky, M.S.: The Political Dynamics of Sanctions: A Comparative Study, Journal of International Affairs, vol. 75, no. 2, pp. 123-145, 2021.
[2]. Tausczik, Y., and Pennebaker, J.: The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology, vol. 29, no. 1, pp. 24-54, 2017.
[3]. Mearsheimer, J.: The Tragedy of Great Power Politics, W.W. Norton and Company, New York, 2001.
[4]. Goodfellow, I., Bengio, Y., and Courville, A.: *Deep Learning*. MIT Press, Cambridge, 2016.
[5]. Waisbord, S.: Watchdog Journalism in South America: News, Accountability, and Democracy. Columbia University Press, New York, 2000.

[6]. McCallum, A., and Nigam, K.: A Comparison of Event Models for Naïve Bayes Text Classification. AAAI-98 Workshop on Learning for Text Categorization, 1998.
[7]. LeCun, Y., Bengio, Y., and Hinton, G.: Deep Learning. Nature, vol. 521, pp. 436-444, 2015.
[8]. Hochreiter, S., and Schmidhuber, J.: Long Short-Term Memory. Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[9]. Joachims, T.: Text Categorization with Support Vector Machines: Learning with Many Relevant Features. European Conference on Machine Learning, pp. 137-142, 1998.
[10]. Kim, Y.: Convolutional Neural Networks for Sentence Classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746-1751, 2014.
[11]. Lai, S., Xu, L., Liu, K., and Zhao, J.: Recurrent Convolutional Neural Networks for Text Classification. *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pp. 2267-273, 2015.
[12]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I.: Attention is All You Need. *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998- 6008, 2017.
[13]. Devlin, J., Chang, M-W., Lee, K., and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, pp. 4171-4186, 2019.
[14]. Sanh, V., Debattista, L., Gozdz, W., Sanh, A., Chaumond, T., Lhoest, Q., Launay, J., Rush, A., and Ott, M.: DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108, 2019.
[15]. Hu, Z., Shen, Y., Liu, Z., and Sun, M.: Low-Rank Adaptation for Efficient Text Classification. Proceedings of ACL-IJCNLP 2021, pp. 2692-2703, 2021.
[16]. Strubell, E., Ganesh, A., and McCallum, A.: Energy and Policy Considerations for Deep Learning in NLP. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3645-3650, 2019.
[17]. Zhang, C., Bengio, S., Hardt, M., Recht, B., and Vinyals, O.: Understanding Deep Learning Requires Rethinking Generalization. Communications of the ACM, vol. 64, no. 3, pp. 107-115, 2021.
[18]. Doshi-Velez, F., and Kim, B.: Towards a Rigorous Science of Interpretable Machine Learning. arXiv preprint arXiv:1702.08608, 2017.
[19]. Hufbauer, G.C., Schott, J.J., Elliott, K.A., and Oegg, B.: Economic Sanctions Reconsidered. 3rd ed., Peterson Institute for International Economics, Washington D.C., 2007.
[20]. Dreger, C., Gros, K., Kooths, K., and Ulbricht, D.: The Impact of Sanctions and Oil Prices on the Russian Economy. Journal of Comparative Economics, vol. 44, no. 3, pp. 598-615, 2016.
[21]. Entman, R.: Framing Bias: Media in the Distribution of Power. Journal of Communication, vol. 57, no. 1, pp. 163-173, 2007.

Информация об авторах / Information about authors

Brice Donald ABODO ELOUNDOU – Master student at AI Talent Hub, ITMO University. Research interests: natural language processing, machine learning, AI in medicine and geocoding.

Брис Дональд АБОДО ЭЛУНДУ – Магистрант AI Talent Hub Университета ИТМО. Исследовательские интересы: обработка естественного языка, машинное обучение, ИИ в медицине и геокодирование.

Wang QUANYU – Master student at AI Talent Hub, ITMO University. Research interests: deep learning, text classification, and geopolitical data analysis.

Ван ЦЮАНЬЮЙ – Магистрант AI Talent Hub Университета ИТМО. Научные интересы: глубокое обучение, классификация текстов, анализ геополитических данных.