UDC: 004.8

NEURAL NETWORK AND HYBRID MECHANISMS IN DATA PROCESSING FOR FINANCIAL ANALYTICS

G. S. Ticona Perales, E. A. Kubrakova

Moscow Institute of Physics and Technology (National Research University)

In this work, we explore hybrid mechanisms that combine sentiment analysis and traditional technical indicators to enhance predictive modeling in financial analytics. The project focuses on comparing and fine-tune neural network-based natural language processing models for sentiment classification of investor-generated content to improve the performance of the hybrid model. A core hypothesis of this research is that news sentiment—when processed with domain-specific models—can serve as a significant and explainable feature for financial forecasting models, particularly when integrated alongside classical stock market indicators.

Our main data source comprises user-generated financial publications from a social media platform called Pulse, covering the period 2020–2024. We filtered the data based on a list of tickers: 'SBER', 'MOEX', 'GMKN', 'MTSS', 'AFLT', and 'LKOH'. Previous experiments showed that aggregated daily sentiment scores (based on PCA with 5 components) explained up to 90% of the variance in closing prices when combined with technical indicators, confirming their complementary role in predictive modeling.

To achieve this, our hybrid architecture (Fig. 1) combines Technical Analysis (TA) features—derived from price and volume indicators—with Sentiment Analysis (SA) outputs obtained from publication streams. The conceptual workflow follows the hybrid model framework proposed in [Kubrakova, 2024], in which machine learning features are drawn from both structured time-series and unstructured text data.

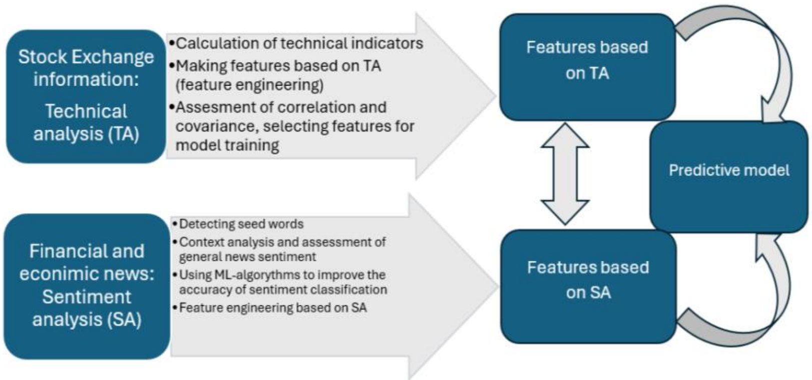

Figure 1: The structure of a hybrid predictive model combining features derived from Technical Analysis (TA) and Sentiment Analysis (SA). Adapted from: Kubrakova E.A., *"Hybrid approach to sentiment analysis and its parametrization in machine learning models"*, XI International Conference "Modern Econometric Tools and Applications" (META-2024), Nizhniy Novgorod, Higher School of Economics (2024).

The datasets employed include:

- Telegram Financial Sentiment Dataset
- Small self-labeled Pulse dataset
- RuFinancial News (2014–2024)

After surveying available neural network architectures suited for financial sentiment tasks [1][3], we selected two primary models for experimentation:

1. **RuBERT-base-cased-sentence**: A pre-trained transformer model fine-tuned on a corpus google-translated to Russian for learning natural language inference. The model benefits from strong general language understanding and performs well on short, noisy texts typical of social media.
2. **FISHQA (Financial Sentiment Analysis with Hierarchical Query-Driven Attention)** [2]: This model is designed for interpretable document-level sentiment analysis. It employs a hierarchical structure (processing documents at the word and sentence levels) and integrates a query-driven attention mechanism, allowing it to emphasize content relevant to user-defined aspects (e.g., litigation, personnel changes, or financial risk). The architecture of FISHQA is outlined below (Fig. 2):

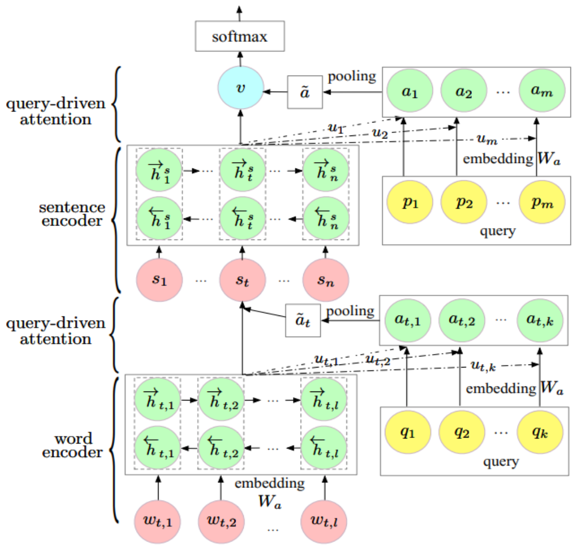

Figure 2: The network structure of FISHQA. Source: Yang, B., Cardie, C., & Teng, F. (2018). *Beyond polarity: Interpretable financial sentiment analysis with hierarchical query-driven attention*. Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18), 4461–4467

In our experiments, we compare both models on the task of classifying financial sentiment for the selected tickers over the given period. The evaluation includes performance metrics such as **accuracy, F1-score**. We assume that the model that performs better will consequently perform better in the hybrid model (Reducing MSE). In particular, we assess whether FISHQA's explainability and attention-weighted sentence selection outperform a standard transformer baseline.

References

[1] Qiu, J., Wu, Q., Ding, G., Xu, Y., Feng, S., & Shi, Y. (2016). A survey of machine learning for big data processing. *Journal of Big Data, 3*(1), 1–17. https://doi.org/10.1186/s40537-017-0111-6

[2] Yang, B., Cardie, C., & Teng, F. (2018). Beyond polarity: Interpretable financial sentiment analysis with hierarchical query-driven attention. *IJCAI-18 Proceedings*, 5, 4461–4467. https://www.ijcai.org/proceedings/2018/590

[3] Jin, D., Zhang, Y., Zhu, H., He, X., Li, Y., & Wu, J. (2024). Financial sentiment analysis: A survey. *ACM Computing Surveys, 56*(3), 1–38. https://doi.org/10.1145/3649451

[4] Kubrakova E.A. (2024). Hybrid approach to sentiment analysis and its parametrization in machine learning models. *XI International Conference "Modern Econometric Tools and Applications" (META-2024), HSE, Nizhniy Novgorod.*