Articles related to **Predicting Stock Market Trends baed on Historical Data and News Sentiment Analysis (Enhancing Stock Market Predictions Through Historical Data Including Both Prices and Patterns, Multilingual News Sentiment Analysis and Financial Indicators Status Using Limited Labeled Data)**

1. <https://arxiv.org/abs/2409.05698>

**MANA-Net: Mitigating Aggregated Sentiment Homogenization with News Weighting for Enhanced Market Prediction**

This passage highligths an important issue in financial sentiment analysis described as Aggregated Sentiment Homogenization, where existing methods simplify sentiment aggregation, ledaing to the losss of valuable information. To address this limitation, it introduces the Market Attention-weighted News Aggregation Network (MANA-Net.

**Chapters**: Introduction, Related Works (Financial news sentiment analysis, News-based Market Predicition, Attention Mechanism), Preliminaries (Problem Formulation, Data, ), Sentiment Aggregation Challenges, Methodology (Weighted News Aggregation, Market Prediciton and Training), Experiment Settings (Evaluation Metrics, Baseline Methods, Training and Validation), Result and Discussion (Prediction model Selection, Prediction Results and Analysis, Analysis of News Weights), Conclusion

**Introduction**:

This chaptr highlights the limitations of static aggregation in financial news sentiment analysis, which fails to account for the dynamic impact of news on market predictions, so the authors made a research on MANA-Net approach.

**Related Works:**

This section reviews that MANA-Net introduces a dynamic attention mechanism that assigns weights to news items based on their market impact, improving the accuracy of market predictions.

**Preliminaries**:

The "Preliminaries" chapter outlines the problem of predicting daily market trends using price and sentiment data and highlights the importance of effectively integrating these elements for accurate market predictions.

**Sentiment Aggregation Challenges:**

The "Sentiment Aggregation Challenges" chapter highlights the analysis done, using TRNA and FinBERT datasets.

**Methodology**:

MANA-Net addresses sentiment homogenization by integrating a trainable market-news attention mechanism, dynamically weighting news items based on their relevance to market price changes. This system allows adaptive sentiment aggregation and improves market prediction by optimizing both news and price inputs.

**Experiment Settings:**

The researchers evaluated MANA-Net using specific GPUs, achieving strong performance on Accuracy, Profit & Loss (PnL) and proving superior market prediction capabilities.

**Result and discussion:**

Multi-Layer Perceptrons outperform Transformers in MANA-Net's market prediction tasks due to its simplicity and effectiveness, despite Transformer’s complexity making difficult achieving overall performance in this context.

**Conclusion:**

MANA-Net is based on Aggregated Sentiment Homogenization in financial news by dynamically weighting news items based on market impact, improving sentiment aggregation and enhancing market prediction performance.

66 references and no citations

1. [[2007.02238] News Sentiment Analysis (arxiv.org)](https://arxiv.org/abs/2007.02238)

**News Sentiment Analysis**

This research paper shows that lexicon-based approach works analysing news headlines and no data labeling was required, so this way it could work both supervised and unsupervised learning.

**Chapters**: introduction, Related work, Research Methodologies, Result and Discussion, Conclusion

**Introduction:**

Introduction reiterates that this research focuses on lexicon-based sentiment analysis, which uses a dictionary to automatically determine the sentiment in news articles.

**Related work:**

More researchers have advanced news sentiment analysis using diverse methodologies, with studies concluding that machine learning approaches yield more accurate results than traditional lexicon-based methods.

**Research Methodologies:**

This chhapter describes a Lexicon-based approach for sentiment analysis of news artciles, utilizing predefined sentiment dictionaries to determine the polarity of words. The methodology consists of data, text preprocessing, calculating sentiment classifiction with a lexical dictionary, aggregating total sentiment scores for each document, and classifying the news articles.

**Result and Discussion:**

The study faced limitations, such as focusing only on English articles from a single source and challenges in word coverage and sentiment classification accuracy across diverse datasets.

**Conclusion:**

This paper examined sentiment analysis of BBC news articles finding that business and sports had more positive sentiments, while entertainment and tech were predominantly negative and future work will involve applying machine learning for this topic.

22 references and 24 citations

1. [[2410.07143] SARF: Enhancing Stock Market Prediction with Sentiment-Augmented Random Forest (arxiv.org)](https://arxiv.org/abs/2410.07143)

**SARF: Enhancing Stock Market Prediction with Sentiment-Augmented Random Forest**

This research introduces a new technique, SARF, for optimizing the accuracy of stock market predictions by combining sentiment analysis with FinGPT and optimized Random Forest model. The promising results indicate the potential of this approach for any future research of this specific topic.

**Chapter**: Introduction, Related Work, Proposed Method, Experimental Evaluation, Experimental Result, Conclusion

**Introduction:**

Predicting stock trends is challenging due to the complex and dynamic nature of the market, but machine learning models offer superior performance and robustness in forecasting by learning from historical data.

**Related work:**

Various studies utilize technologies like statistics and data mining, notably random forest models, to enhance stock value predictions, demonstrating improved accuracy and reduced errors through effective feature selection and integrated modeling.

**Proposed Method:**

This section presents the Sentiment-Augmented Random Forest (SARF) model, which integrates technical indicators with sentiment analysis from FinGPT to enhance stock trend predictions by capturing market sentiment and improving model robustness against overfitting.

**Experimental Evaluation:**

The Random Forest model was trained on daily time series with exponential smoothing applied to enhance predictive performance.

**Experimentl results:**

Preliminary results indicate that the Sentiment-Augmented Random Forest (SARF) model outperforms conventional Random Forest and LSTM models in predicting stock market movements, especially in volatile conditions, with optimal performance achieved at a 60-day time window.

**Conclusion:**

The promising results of this approach indicate the potential of this approach for real-world applications in financial forecasting.

25 references and 7 citations

1. <https://arxiv.org/abs/2206.09591>

**Domain-Adaptive Text Classification with Structured Knowledge from Unlabeled Data**

This study is relevant to my paper due to the fact that stock market news can cover different sectors and financial contexts. Using techniques like Domain Adaptive Text Classification (as proposed in this article - DASK) allows a NLP (Natural Language Processing) model to adjust and learn more accurately across various sectors or domains.

**Chapters**: Introduction, Related Work, Method, Experiments, Conclusion

**Introduction**:

The introducition presents DASK, a novel pivot-based domain adaptation method for text classification used for enhancing the model's ability to generalize across domains.

**Related Work:**

It details the construction of knowledge graphs (KGs), focusing on unsupervised approaches like MAMA that utilize high attention scores from pre-trained language models like BERT.

**Method**:

DASK facilitates domain adaptation by constructing a knowledge graph (KG) to capture relationships between pivots and non-pivots, enhancing feature learning through a two-step approach: extracting and filtering facts; injecting these facts into source domain texts.

**Experiments**:

The experimental results demonstrated that DASK outperforms baseline models across various domain pairs. Hyperparameter tuning is critical for optimizing performance, with specific settings for different parameters.

**Conclusion**:

This paper introduces Domain-Adaptive text classification with Structured Knowledge (DASK), which integrates pivot-based methods, knowledge graphs, and pretrained language models to enhance sentiment classification.

27 references and 56 citations

1. <https://arxiv.org/abs/2410.00024>

**Cross-Lingual News Event Correlation for Stock Market Trend Prediction**

The research introduces a cross-lingual Natural Language-based Financial Forecasting (NLFF) pipeline, which is importan for my serearch papre due to its innovative approach in financial analysis

**Introduciton**:

This paper discusses how FinTech and AI natural language processing (NLP) are transforming the finance industry, particularly in stock market analysis.

**Related Work:**

It highlights the significance of sentiment analysis in understanding market sentiments and trends, noting advancements in Urdu sentiment analysis, such as the development of the Urdu Corpus for Sentiment Analysis

**Methodology**:

This researhc focuses on extracting, mapping, and visualizing a timeline of news events affecting the Pakistani stock market, starting with the collection of financial news articles in Urdu.

**Results and discussion:**

The analysis revealed a strong focus on the economic sector within the news, indicating a preference for English sources among the Pakistani business community due to perceived quality

**Conclusion and future work:**

This study presssents a cross-lingual NLFF pipeline utilizing semantic similarity, sentiment analysis, and named entity recognition to create a financial events timeline from Urdu and English news sources, successfully correlating stock price movements with news sentiments. Future research can focus on identifying relevant news events impacting stock prices, improving Urdu sentiment analysis.

Chapters: Introduction, Related Work, Methodology, Results and discussion, Conclusion and future work

64 references and no citations

**Bibliography:**

Wang, M., & Ma, T. (2023). MANA-Net: Mitigating aggregated sentiment homogenization with news weighting for enhanced market prediction. *Proceedings of the 2023 International Conference on Artificial Intelligence and Big Data (ICAIBD)*

Samuels, A., & McGonigal, J. (2023). News sentiment analysis. *Proceedings of the 2023 International Conference on Data Science and Advanced Analytics (DSAA)*

Talazadeh, S., & Peraković, D. (2023). SARF: Enhancing stock market prediction with sentiment-augmented random forest. *Proceedings of the 2023 IEEE International Conference on Data Mining Workshops (ICDMW)*

Li, T., Chen, X., Dong, Z., Yu, W., Yan, Y., Keutzer, K., & Zhang, S. (2023). Domain-adaptive text classification with structured knowledge from unlabeled data. *Proceedings of the 2023 IEEE International Conference on Data Mining (ICDM)*

Arshad, S., Azhar, N., Sajid, S., Latif, S., & Latif, R. (2023). Cross-lingual news event correlation for stock market trend prediction. *Proceedings of the 2023 International Conference on Computer and Communication Systems (ICCCS)*