DOI: https://doi.org/10.15276/hait.07.2024.23
UDC 004.89:336.76:330.43

Development of a software service for stock price forecasting based on sentiment analysis and autoregressive models

Vitaliy M. Kobets[1]
ORCID: https://orcid.org/0000-0002-4386-4103; vkobets@kse.org.ua. Scopus Author ID: 56006224700
Nikita D. Stang[1]
ORCID: https://orcid.org/ 0009-0007-1382-3187; nstang2002@gmail.com
[1] Kherson State University, 27, Universitetska Str. Kherson, 73003, Ukraine

ABSTRACT

This paper addresses the critical need for efficient market analysis tools in the era of big data and artificial intelligence. We present a novel software service that integrates real-time news sentiment analysis with stock market prediction, enhancing the accuracy and speed of trading decisions. The system employs APIs for data collection, FinBERT for sentiment analysis, and MongoDB for data storage, overcoming limitations of existing platforms like Investing.com and MarketWatch. Our methodology combines sentiment analysis with autoregressive models to forecast stock prices for 11 major companies. The experiment utilized 141 observations, applying multiple regression and binary outcome models. Results demonstrate that investor sentiment significantly affects stock prices for 2 out of 11 companies, with Meta showing a 70 % determination coefficient in price direction changes based on sentiment. The study reveals that incorporating both quantitative (previous stock prices) and qualitative (sentiment) data improves forecast accuracy for certain stocks. This research contributes to the field of financial analytics by providing a more comprehensive approach to stock price prediction, integrating ML models and data analytics to support informed decision-making in dynamic financial markets.

Keywords: Stock price forecasting; sentiment analysis; financial analytics; real-time data processing; machine learning in finance

***For citation*: Kobets V. M., Stang N. D. "Development of a software service for stock price forecasting based on sentiment analysis and autoregressive models". *Herald of Advanced Information Technology*. 2024; Vol. 7 No. 3: 321–329. DOI: https://doi.org/10.15276/hait.07.2024.23**

INTRODUCTION

Market efficiency and informed decisions are critical components of successful trading and investing. Financial markets are highly sensitive to various news events, such as economic reports, political developments, and corporate announcements, which can cause significant price fluctuations. By understanding the potential impact of news, traders and investors can make more informed decisions, reducing the risk of unexpected losses. This knowledge enables them to anticipate market movements and react accordingly.

Furthermore, identifying market trends and sentiment through news analysis is invaluable. It helps traders understand the underlying forces driving market movements, allowing them to spot emerging trends early. This insight can be leveraged to align trading strategies with prevailing market conditions. Additionally, sentiment analysis of news can reveal the overall mood of the market, whether bullish or bearish, aiding in the timing of entry and exit points for trades.

However, analyzing news by hand can be exhausting and unproductive, as it requires sifting through vast amounts of information, often leading to delays in decision-making and missed opportunities. Automated systems, powered by artificial intelligence and machine learning, can efficiently process and analyze news in real-time, providing actionable insights instantly. These systems can detect patterns, sentiments, and correlations that may be difficult for humans to identify, thereby enhancing the speed of trading decisions. This automation not only saves time but also ensures that traders and investors can focus on strategic planning rather than getting bogged down by the tedious task of manual analysis.

The relationship between news sentiment and stock market behavior has been a focal point in financial research, with advancements in AI and big data analytics driving new methodologies. Traditional systems, such as those offered by Investing.com and MarketWatch, provide real-time financial news linked to stock prices but often lack the capability to accurately measure the sentiment's impact on market movements.

The problem that arises in investor decision-making is the need to quickly rebalance the investment portfolio and make decisions on buying and selling stocks in the event of important news that will affect the stock price. Existing services only visualize and/or track stock price movements after news releases, without investigating the causal relationship between the sentiment of the news and the stock price. This creates a gap between the interpretation of financial news content on the stock price and investors' decision-making, which can lead to significant financial losses in the short-run.

Our research presents a novel system that integrates real-time news sentiment analysis with stock market prediction service, addressing the limitations of existing tools. This system automates the collection and analysis of data using APIs, ensuring that traders and investors receive timely and relevant insights. The inclusion of a MongoDB database supports efficient data storage and retrieval, allowing for more comprehensive long-term analysis.

By enhancing the speed and accuracy of trading decisions, this system marks a significant improvement over manual analysis methods. However, it also highlights areas for future development, such as refining news selection algorithms and expanding the range of financial instruments monitored. These improvements could further empower investors to make better-informed decisions in the financial markets dynamic.

The purpose of the paper is to develop a service for predicting the direction of change in the share price, taking into account investor sentiment and historical data on the share price.

The structure of the paper is as follows: section 2 includes literature review, section 3 describes architecture of stock price prediction system, section 4 introduces experimental application of service for prediction. Last section concludes.

RELATED WORKS

The study of news sentiment and its influence on stock returns has seen significant advancement, incorporating a range of analytical techniques to enhance forecasting precision. The body of literature on this topic presents a diverse array of findings, emphasizing the evolving nature of how sentiment analysis is utilized in financial markets. The main themes emerging from the literature review include the following key ideas:

Real-time sentiment analysis has emerged as a crucial tool in predicting high-frequency stock returns. By employing a mixed-frequency-rolling decomposition forecasting method, research has demonstrated the value of incorporating timely sentiment data into financial models. This approach allows for the capturing of the immediate and short-term impact of investor sentiment on stock prices, especially in environments characterized by high-frequency trading. The ability to integrate real-time sentiment analysis into prediction models has shown to enhance the accuracy and responsiveness of stock forecasts, providing investors with a more dynamic and real-time tool for navigating market fluctuations [1].

The dispersion of news sentiment, or the variation in sentiment among different news sources, has been identified as a significant factor influencing market events such as mergers and acquisitions (M&A). Studies have explored how differences in sentiment across various media outlets can impact the perception and outcome of M&A transactions. The findings suggest that greater dispersion in sentiment tends to increase uncertainty in the market, which in turn affects the decision-making processes of both investors and firms. This insight highlights the nuanced and often complex nature of interpreting news sentiment, where not only the overall sentiment but also the consistency or variability of that sentiment can have profound effects on market behavior and investor confidence [2].

Incorporating news sentiment analysis into the construction of investment portfolios has been shown to significantly enhance investment strategies. By leveraging sentiment analysis, investors can make more informed decisions, leading to improved portfolio performance. Research in this area has explored how the integration of sentiment data with traditional financial indicators can create more intelligent portfolio construction methodologies. This approach allows for a more holistic analysis of market conditions, enabling investors to anticipate market trends and adjust their strategies accordingly. As a result, portfolios that incorporate sentiment analysis are often better positioned to capitalize on market opportunities and mitigate potential risks [3, 4].

The application of machine learning techniques, such as Particle Swarm Optimization (PSO) hypertuned neural networks, has further refined the accuracy of stock price predictions. By combining sentiment analysis with advanced machine learning models, researchers have been able to optimize prediction performance, leading to more reliable and robust stock forecasts [5]. Additionally, deep learning-based models that incorporate both local

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)

and global event sentiment have been developed to efficiently predict movements in stock exchanges. These models demonstrate the potential of integrating artificial intelligence with sentiment analysis, providing a powerful tool for financial forecasting that can adapt to the complex and dynamic nature of global markets [6].

In recent years, the application of deep learning and sentiment analysis techniques for stock price forecasting has gained significant attention. Research has shown that deep learning models, particularly neural networks, excel in predicting stock prices due to their ability to capture non-linear patterns in financial data. Studies comparing these models highlight their effectiveness in addressing the complexity of financial market behavior.

Hybrid models that integrate deep learning with investor sentiment analysis have also been explored. These models incorporate both market data and investor sentiment, demonstrating improvements in predictive accuracy by considering the psychological factors driving market trends. This approach has proven valuable in enhancing the performance of traditional forecasting methods. [7]

The evolution of data-driven approaches for financial market prediction has been the subject of various studies, emphasizing the growing importance of big data and advanced computational techniques. These models leverage large datasets and sophisticated algorithms to enhance forecasting precision, offering new insights into stock price movements. [8]

The informativeness of technical indicators has also been studied extensively, showing that combining technical analysis with advanced forecasting models can lead to more accurate predictions. By incorporating multiple data sources, such as technical indicators, the predictive power of these models is significantly improved. [9, 10]

Sentiment analysis plays a crucial role in enhancing stock price prediction models. Research has demonstrated that analyzing news and market sentiment can provide valuable insights into market reactions, influencing stock returns. Deep learning techniques have been applied to sentiment analysis, utilizing large-scale data to refine sentiment-driven predictions. Additionally, the dynamic relationship between news sentiment and stock return volatility has been explored through regime-switching models, illustrating how sentiment can impact market fluctuations. [11, 12]

Further advancements in sentiment analysis using deep learning methods have shown the importance of capturing market behavior from unstructured data sources, such as news and social media. These studies collectively emphasize the critical role of integrating deep learning, technical analysis, and sentiment analysis to improve the accuracy and reliability of stock price forecasting models [13, 14, 15, 16].

The Investing.com news widget (Fig. 1) is a comprehensive service that provides up-to-date financial news on a wide range of markets, and correlates it to current financial instrument price, however it does not indicate how much of an impact the news article caused, and whether the impact is positive or not.

Fig. 1. **Investing.com news widget**
Source: **compiled by the authors**

MarketWatch news articles have a ‘referenced symbols feature’ that can help track the impact of a news article, however it is hard to determine whether it was this specific article that caused the market change (Fig. 2).

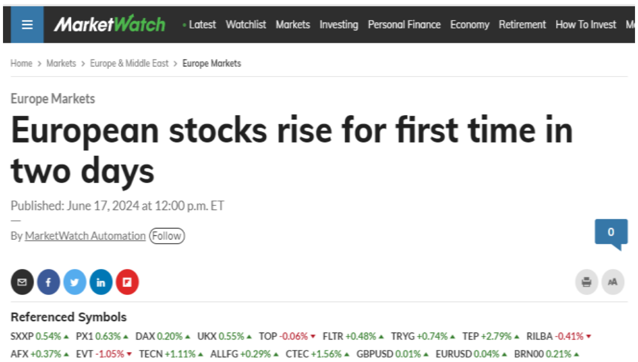

Fig. 2. **Example of a MarketWatch news article**
Source: **compiled by the authors**

Both of these cases have significant drawbacks: in the first example, the statistical significance and direction of the news impact on the stock price are uncertain, and in the second case, the correlation and causal relationship between the news and the share price are not verified. If a software module were created that could overcome these shortcomings, the forecast of a stock prices would become more reliable and would contribute to more informed decision-making by investors. We propose a software module that can overcome these challenges in next section.

ARCHITECTURE OF STOCK PRICE PREDICTION SYSTEM

Let us consider system description for gathering of relevant news about stocks and their transformation in sentiment of investors using FinBert service.

1. The system fetches data from three APIs: Stock Price, News, and Sentiment Analysis.
2. The Data Controller aggregates this data and passes it to the Data Service.
3. The Data Service processes the data and sends it to the Data Repository or sends back an error to Data Controller
4. The Data Repository manages the data before it is finally stored in the MongoDB database, also it can send back an error to the Data Service

This description is visualized in diagram below (Fig. 3)

The collection of data is described in Fig. 3, which shows the gathering of information on the two predictors and the dependent variable. The dependent variable is the share price at time t, and one of the predictors is the share price at time t-1. The data is processed in real time. At the same time from all dataset, the closing price for the stock price for the respective day is selected, because only one stock price per day is used in our experimental model.

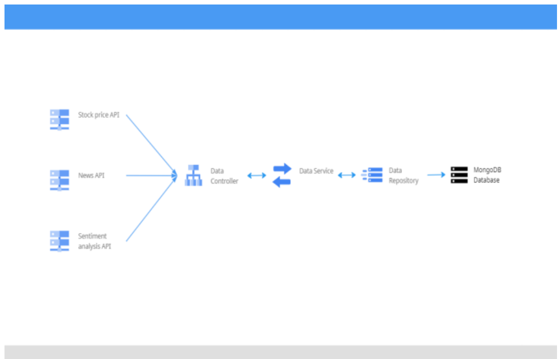

Fig. 3. ***High-level system architecture diagram***
Source: prepared by the authors

Simultaneously the highest-rated news from a financial website about the stocks of the companies in the sample of selected list of stocks is identified, and then processed using FinBERT to determine the sentiment of the news (positive, negative or neutral).

Structure of system is described in Table 1.

Table 1. **Structure of system**

System components	Modules (if any)	Description
APIs	Stock Price API	Provides real-time stock price information for selected stock.
	News API	Delivers news articles from google news, configured to select the most popular article by the end of the day (UTC +0) for selected stock
	Sentiment Analysis API	Analyzes the sentiment of the news title to provide a sentiment score (e.g., positive, negative, neutral)
Data Controller		This component collects data from Stock Price and News API's and then fetches the sentiment of a news header from Analysis API. It acts as an aggregator that ensures data from different sources is unified and formatted correctly for further processing.
Data Service		The Data Service is responsible for the business logic. It processes the collected data, which may include filtering, transforming, and enriching the data before it is stored.
		It may interact with the Data Controller to fetch data and with the Data Repository for storing and retrieving processed data.
Data Repository		This is an intermediate storage layer where processed data is stored temporarily. It serves as a buffer between the Data Service and the MongoDB database.
		The Data Repository ensures data consistency and provides efficient access to data for both storing and retrieving.
MongoDB Database		The final storage destination for the data. The MongoDB database is a NoSQL database known for its scalability and flexibility in handling different types of data.
		It stores the processed and structured data for long-term persistence and future retrieval.

Source: **prepared by the authors**

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)

Advantages and limitations of selected technologies for system architecture are prepared in Table 2 and Table 3. The .NET Framework is a proprietary software framework developed by Microsoft that runs primarily on Microsoft Windows. MongoDB is a source-avaliable, cross-platform, document-oriented database program. Classified as a NoSQL database product, MongoDB utilizes JSON-like documents with optional schemas. MongoDB is developed by MongoDB Inc. and current versions are licensed under the Server Side Public License (SSPL).

Table 2. **Characteristics of .NET Framework**

Aspects of technology	Description of .NET Framework aspects
Advantages	**Cross-Platform Support**: It is possible to develop and deploy applications using NET Core on Windows, Linux, and macOS. This flexibility is beneficial for diverse deployment environments. **Performance** Just-In-Time (JIT) compiler, along with other optimizations, ensures that applications run smoothly and efficiently **Strong Typing and Compile-Time Checking** C# (as primary language for .NET development) is a statically typed language, which means many errors can be caught at compile time rather than at runtime. This leads to more reliable and maintainable code. **Asynchronous Programming** .NET provides robust support for asynchronous programming with async and await keywords, which is crucial for handling I/O-bound operations, such as API calls, without blocking the main thread.
Limitations	**Resource Consumption** .NET applications can sometimes be more resource-intensive compared to some other languages and frameworks, particularly for smaller-scale applications or those with limited server resources.

Source: **prepared by the authors**

Table 3. **Characteristics of MongoDB**

Aspects of technology	Description of MongoDB aspects
Advantages	**Flexibility** MongoDB is a NoSQL database that stores data in a flexible, JSON-like format (BSON). This schema-less design allows for easy modifications to the data structure without the need for complex migrations. **Performance** MongoDB offers high performance for both read and write operations. Its design allows for fast querying, indexing, and data retrieval, which is essential for real-time data applications.
Limitations	**Consistency Model** MongoDB uses an eventually consistent model, which might not be suitable for applications requiring strong consistency and immediate data accuracy across all nodes. **Complex Transactions** While MongoDB has improved its transaction support, it may still be less robust compared to traditional relational databases, especially for applications requiring complex multi-document transactions.

Source: **prepared by the authors**

The key deciding factor in favor of MongoDB was its flexibility, allowing shifting focus from developing database schemas to other areas of the system, and thus saving time.

The system was manually tested using QA practices and seeded with mock data prior to the start of the experiment on a test system environment. After successful testing a new production environment was created to ensure that system changes did not affect the existing experiment. Both environments are hosted locally and run using Internet Information Services (IIS).

EXPERIMENT

The system collects results from APIs every weekday, selecting top news from Google News (if it can find one) for 10 companies and related stocks: Netflix, Microsoft, Shell, Totalenergies, Nvidia, Meta, Amazon, AMD, Google and Intel and writes results in the database for future analysis. Fragment of extracted values of prices and sentiments are shown in Table 4. Total number of observations is 141.

After preliminary data cleaning and preprocessing we can get following Table 5, where $p_t =$ ValueAfterHour, p_{t-1}= Value, s_{t-1}= Sentiment of investors.

In the first model AR(1), as autoregressive model of first order, can be described as ARIMA(1,0,0) model with two predictors that are used to forecast the share price in the form of a multiple regression. First predictor is a share price before the news (Value) and second predictor is an investor sentiment (Sentiment).

$$p_t = a_0 + a_1 \cdot p_{t-1} + a_2 \cdot s_{t-1} + u_t, \quad (1)$$

where p_t is dependent variable or value of price after hour of sentiment impact, p_{t-1} is explanatory variable or value of price before sentiment impact; s_{t-1} is explanatory variable or value of sentiment impact (positive, neutral or negative); u_t is error term.

The second model is constructed as follows:

$$dp_t = b_0 + b_1 \cdot s_{t-1} + w_t \quad (2)$$

where dp_t is dependent variable that equals 1 if $p_t > p_{t-1}$, -1 if $p_t < p_{t-1}$ and equals 0 if $p_t = p_{t-1}$, s_{t-1} is explanatory variable or value of sentiment impact (positive, neutral or negative); u_t is error term.

After estimating both models using the OLS method and applying the t-criterion to check the statistical significance of the parameters, we obtain the following results

Table 4. **Gathered results through API**

Stock	Value	Headline	Value After Hour	Sentiment	Date
MSFT	422.92	Xbox Cloud Gaming finally supports keyboard and mouse inputs on web browsers	422.965	0.28985899686813354	$2024-05-15T16:52:15Z
SHEL	73.3	Tamagotchi collectors rejoice: Bandai is finally rereleasing a beloved model from 2004	73.41	- 0.0826990008354187	$2024-05-15T16:46:45Z
AMD	159.199	10 Things You Should Be Doing Daily to Protect Your Eye Health - CNET	158.942	0.41071946918964386	$2024-05-15T14:05:41Z
INTC	31.195	Intel's New Thunderbolt Share Provides File and Screen Sharing Without Hurting Network Performance	31.21	0.0814552903175354	$2024-05-15T16:42:00Z
NKE	91.6	I have spent entirely too much time thinking about Mark Zuckerberg's outfit at his birthday party	91.571	- 0.1332276463508606	$2024-05-15T17:58:47Z
NFLX	638.33	Netflix releases first look at new Witcher after Henry Cavill left for Warhammer 40K	638.33	0.33275464177131653	$2024-05-22T20:14:13Z
NVDA	1007.0	Nvidia just keeps hitting it out of the park	1007.0	- 0.1962529420852661	$2024-05-22T22:33:16Z
GOOG	177.42	Snap brings its AR lenses to Chrome through an extension	177.59	0.1168382465839386	$2024-05-22T19:28:05Z
TTE	70.5	Patrick PouyannΓ© affirme que TotalEnergies restera en France et rassure Macron	70.5	0.30006177723407745	$2024-05-23T19:32:13Z

Source: **prepared by the authors**

ISSN 2663-0176 (Print)
ISSN 2663-7731 (Online)

Table 5. **Values of stock prices before and after sentiments of investors**

Stock	Value	ValueAfterHour	Sentiment	Date	Time
MSFT	422.92	422.97	0.289858997	15.05.2024	16:52:15
SHEL	73.3	73.41	-0.082699001	15.05.2024	16:46:45
TTE	73.5	73.48	0.25498122	15.05.2024	18:00:02
AMD	159.199	158.94	0.410719469	15.05.2024	14:05:41
INTC	31.195	31.21	0.08145529	15.05.2024	16:42:00
NKE	91.6	91.57	-0.133227646	15.05.2024	17:58:47
NKE	92.44	92.44	0.353558853	22.05.2024	22:05:10
NFLX	638.33	638.33	0.332754642	22.05.2024	20:14:13
NVDA	1007	1007	-0.196252942	22.05.2024	22:33:16
GOOG	177.42	177.59	0.116838247	22.05.2024	19:28:05
NKE	92.44	92.44	0.353558853	22.05.2024	22:05:10
NKE	91.39	91.47	0.416316167	23.05.2024	18:22:14
NFLX	634.45	634.45	0.028437734	23.05.2024	22:00:33
SHEL	70.33	70.19	0.407320112	23.05.2024	13:58:25
TTE	70,5	70,5	0,300061777	23.05.2024	19:32:13
NVDA	1056.4	1037.9	0.397178553	23.05.2024	13:38:21

Source: prepared by the authors

Table 6. **Parameters of price prediction models**

Stock	a_1	a_2	R_1^2	b_1	R_2^2
TTE	0.9986*	0.2199	99.58%	0.8870	5.69%
SHEL	1.0432*	0.1254	99.21%	0.0407	0.01%
NVDA	0.9986*	-5.4765	99.99%	1.1046	9.00%
NKE	0.9990*	0.0726	99.97%	0.8422	3.92%
NFLX	0.9979*	**-3.8008***	99.95%	-1.8080	20.60%
MSFT	1.0046*	0.2195	99.98%	2.6682	26.06%
META	0.9956*	**1.4641***	99.99%	**3.4737***	**70.00%**
INTC	0.9104*	-0.2202	90.04%	-0.5167	3.04%
GOOG	1.0097*	-0.5636	99.88%	-1.5804	11.04%
AMZN	0.8740*	-1.4521	99.84%	3.0305	23.83%
AMD	1.0369*	-0.1624	99.95%	1.0345	11.52%

* means statistical significance of parameters

Source: prepared by the authors

The results of the calculations for 11 companies showed that only two of them had an impact on their share price. The impact of sentiment analysis on share price was negative for Nvidia (- 3.8008) and positive for META (1.4641). For META, investor sentiment was statistically significant and had a positive impact in the second model (3.4737), as well as a coefficient of determination of $R^2 = 70$ %, which demonstrates that a change in investor sentiment determines the direction of change in META's share price by 70%.

The models also demonstrated that for all 11 companies, the previous stock price had a statistically significant impact on the current share price, as described by the ARIMA model. For two of the 11 companies, the predicted price value depended not only on the previous share price, but also on investor sentiment, which demonstrates the increased accuracy of the forecast price when using quantitative (share price) and qualitative (sentiment) information than using only a quantitative predictor (previous price). Our models outperforms existing services in Fig. 1 and Fig. 2, because our models identify causal rather than correlational relationships and convert investor sentiment from news into a quantitative value, allowing for more accurate stock price forecasting based on both historical data and investor sentiment.

CONCLUSIONS

Sentiment analysis of news can reveal the overall mood of the market, whether bullish or bearish, helping to time entry and exit points for trades. Financial IT services can identify patterns, sentiment and correlations that may be difficult for humans to detect, thereby increasing the speed of trading decisions. This automation not only saves time, but also ensures that traders and investors can focus on strategic planning rather than getting bogged down in the tedious task of manual analysis.

Our research presented a novel system that integrates real-time news sentiment analysis with a stock market forecasting service, overcoming the limitations of existing tools. This system automates the collection and analysis of data using APIs, ensuring that traders and investors receive timely and relevant insights.

Our models outperform existing services because they identify causal rather than correlational relationships and convert investor sentiment from news into a quantitative value, allowing for more accurate stock price predictions based on both historical data and investor sentiment.

REFERENCES

1. Cai, Y., Tang, Z. & Chen, Y. “Can real-time investor sentiment help predict the high-frequency stock returns? Evidence from a mixed-frequency-rolling decomposition forecasting method”. *The North American Journal of Economics and Finance* 2024; 72: 102147. DOI: https://doi.org/10.1016/j.najef.2024.102147.

2. Chen, Y., Lu, J., Ma, W., Kumar, S. & Shahab, Y. “Dispersion in news sentiment and M&As Outcomes”. *Research in International Business and Finance* 2024; 71: 102415. DOI: https://doi.org/10.1016/j.ribaf.2024.102415.

3. Hung, M. C., Hsia, P. H., Kuang, X. J. & Lin, S. K. “Intelligent portfolio construction via news sentiment analysis”. *International Review of Economics & Finance* 2024; 89: 605–617. DOI: https://doi.org/10.1016/j.iref.2023.07.103.

4. Shahi, T. B., Shrestha, A., Neupane, A. & Guo, W. “Stock Price Forecasting with Deep Learning: A Comparative Study”. *Mathematics.* 2020; 8 (9): 1441. DOI: https://doi.org/10.3390/math8091441.

5. Chauhan, A., Shivaprakash, S.J., Sabireen, H., Abdul Q. Md. & Venkataraman, N. “Stock price forecasting using PSO hypertuned neural nets and ensembling”. *Applied Soft Computing.* 2023; 147: 110835. DOI: https://doi.org/10.1016/j.asoc.2023.110835.

6. Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., Selim, M. M. & Muhammad, K. “A local and global event sentiment based efficient stock exchange forecasting using deep learning”. *International Journal of Information Management.* 2020; 50: 432–451. DOI: https://doi.org/10.1016/j.ijinfomgt.2019.07.011.

7. Jing, N., Wu, Z. & Wang, H. “A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction”. *Expert Systems with Applications.* 2021; 178: 115019. DOI: https://doi.org/10.1016/j.eswa.2021.115019.

8. Dai, T., Hu, J. & Ding, X. “Progress and prospects of data-driven forecasting models for financial markets”. *International Journal of Cognitive Computing in Engineering.* 2023; 19 (2): 145-168. DOI: https://doi.org/10.1016/j.ijcin.2023.102415.

9. Shi, Y., Ho, K.-Y. & Liu, W.-M. “Public information arrival and stock return volatility: Evidence from news sentiment and a Markov Regime-Switching Approach”. *International Review of Economics & Finance.* 2016; 42: 291–312. DOI: https://doi.org/10.1016/j.iref.2015.10.011.

10. Zhou, Y., Tang, J., & Wang, Q. “Investigating the in formativeness of technical indicators: A comprehensive study”. *Knowledge-Based Systems.* 2022; 236: 107492. DOI: https://doi.org/10.1016/j.knosys.2022.107492.

11. Li, X., Xie, H., Chen, L., Wang, J. & Deng, X. “News impact on stock price return via sentiment analysis”. *Knowledge-Based Systems.* 2014; 69: 14–23. DOI: https://doi.org/10.1016/j.knosys.2014.07.011.

12. Souma, W., Vodenska, I. & Aoyama, H. “Enhanced news sentiment analysis using deep learning methods”. *Journal of Computational Social Science.* 2019; 2 (1): 33–46. DOI: https://doi.org/10.1007/s42001-019-00040-6.

13. Shahi, T. B., Shrestha, A., Neupane, A. & Guo, W. “Stock price forecasting with deep learning: A comparative study”. *Mathematics.* 2020; 8 (9): 1441. DOI: https://doi.org/10.3390/math8091441.

14. Sohangir, S., Wang, D., Pomeranets, A. & Khoshgoftaar, T. M. “Big Data: Deep learning for financial sentiment analysis”. *Journal of Big Data.* 2018; 5(1): 1–25. DOI: https://doi.org/10.1186/s40537-017-0111-6.

15. Ivanov, O. & Kobets, V. "Data analysis for predicting stock prices using financial indicators based on business reports". *Proceedings of the Information and Communication Technologies in Education, Research, and Industrial Applications*. *Springer.* Cham. 2023; 1980: 227–239. DOI: https://doi.org/10.1007/978-3-031-48325-7_17.

16. Kobets, V., Yatsenko, V., Mazur, A. & Zubrii, M. "Data analysis of personalized investment decision making using robo-advisers". *Science and Innovation.* 2020; 16 (2): 80–93. DOI: https://doi.org/10.15407/scine16.02.080.

Conflicts of Interest: the authors declare no conflict of interest

Received 26.07.2024
Received after revision 06.09.2024
Accepted 20.09.2024

DOI: https://doi.org/10.15276/hait.07.2024.23
УДК 004.89:336.76:330.43

Розробка програмного сервісу для прогнозування цін акцій на основі аналізу настроїв та авторегресійних моделей

Кобець Віталій Миколайович[1)]
ORCID: https://orcid.org/0000-0002-4386-4103; vkobets@kse.org.ua. Scopus Author ID: 56006224700
Штанг Нікіта Денисович [1)]
ORCID: https://orcid.org/0009-0007-1382-3187; nstang2002@gmail.com
[1)] Херсонський державний університет, вул. Університетська, 27. Херсон, 73003, Україна

АНОТАЦІЯ

У статті розглянута нагальна потреба в ефективних інструментах аналізу ринку в еру великих даних і штучного інтелекту. Ми представляємо новий програмний сервіс, який інтегрує аналіз новинних настроїв у реальному часі з прогнозуванням фондового ринку, підвищуючи точність і швидкість прийняття торгових рішень. Система використовує API для збору даних, FinBERT для аналізу настроїв та MongoDB для зберігання даних, долаючи обмеження існуючих платформ, таких як Investing.com та MarketWatch. Наша методологія поєднує аналіз настроїв з авторегресійними моделями для прогнозування цін на акції 11 найбільших компаній. В експерименті було використано 141 спостереження із застосуванням множинної регресії та моделі з бінарними результатами. Результати демонструють, що настрої інвесторів суттєво впливають на ціни акцій 2 з 11 компаній, причому Meta демонструє коефіцієнт детермінації 70%, що пояснює напрямки зміни цін акцій на основі настроїв інвесторів. Дослідження показує, що врахування як кількісних (попередні ціни акцій), так і якісних (настрої інвесторів) даних підвищує точність прогнозу для акцій певних компаній. Стаття робить внесок у сферу фінансової аналітики, пропонуючи більш комплексний підхід до прогнозування цін на акції, інтегруючи моделі машинного навчання та аналізу даних для підтримки прийняття обґрунтованих рішень на динамічних фінансових ринках.

Ключові слова: прогнозування цін акцій; аналіз настроїв; фінансова аналітика; обробка даних у реальному часі; машинне навчання у фінансах

ABOUT THE AUTHORS

Vitaliy M. Kobets - Doctor of Economic Science, Professor, Professor, Department of Computer Science and Software Engineering. Kherson State University. 27, Universitetska St. Kherson, 73003, Ukraine
ORCID: https://orcid.org/0000-0002-4386-4103; vkobets@kse.org.ua. Scopus Author ID: 56006224700
***Research field*:** Data Science in Economics; Evolutionary Microeconomics; Robo-Advisor

Кобець Віталій Миколайович - доктор економічних наук, професор, професор кафедри Комп'ютерних наук та програмної інженерії. Херсонський державний ун-т, вул. Університетська, 27. Херсон, 73003, Україна

Nikita Stang - student of the Master's Degree Program on Information Systems and Technology.
Kherson State University. 27, Universitetska St. Kherson, 73003, Ukraine
ORCID: https://orcid.org/0009-0007-1382-3187; nstang2002@gmail.com
***Research field*:** Business analytics; data science; machine learning

Штанг Нікіта Денисович - студент магістерської освітньої програми Інформаційні системи та технології кафедри Комп'ютерних наук та програмної інженерії. Херсонський державний ун-т, вул. Університетська, 27.
Херсон, 73003, Україна