International Journal of Innovation Engineering and Science Research
(ISSN: 2581-4591)

www.ijiesr.com

Application of Artificial Intelligence and Machine Learning in Finance Sector, especially for Stock Price Prediction

(A Study and a Prototype Using n8n Comparing OpenAI and DeepSeek)

Chuanyu Cao[1], Clotilde Rohleder[2] Indra Kusumah[3] John Davis[4]
[1,2,3]University of Applied Sciences Constance Constance, Germany
[4]Clemson University, USA

ABSTRACT

This paper addresses the application of artificial intelligence and machine learning in one domain of finance sector – stock price prediction. This study analyzes how sentiment analysis and financial market data can be used by artificial intelligence for stock price forecasting. To validate this approach, an automated workflow for stock price forecasting based on sentiment analysis and financial market data has been developed. Using the workflow tool n8n, sentiment scores are integrated into a probabilistic model. Two AI models, OpenAI and DeepSeek, are compared to assess their predictive capabilities. While DeepSeek proves more effective in capturing emotional nuances, both models show distinct strengths in forecasting market movements. The findings underscore the potential of combining sentiment-driven models with workflow automation. Moreover, n8n demonstrated a flexible platform that facilitates the integration of complex analytical processes. The study contributes an initial exploration of sentiment-based probabilistic modeling for financial market analysis on the n8n platform and enhances the state of the art.

Keywords—*stock price forecasting; sentiment analysis; financial market; n8n workflow; probabilistic model; AI models*

I. Introduction

Stock price prediction remains a major challenge in finance, especially in an environment where digital media continuously generate vast amounts of market information [1], [2]. While access to information is no longer scarce, filtering and interpreting relevant content has become increasingly difficult [3], [4], [5]. Traditional forecasting methods—based on statistical models and rational investor assumptions—struggle to account for or take advantage of rapid sentiment shifts in financial news and social media [1], which has gained substantial significance with increasing participation of users in digital media [6].

Recent advances in artificial intelligence and workflow technologies offer new opportunities [17], [18], [19]. Automated workflows such as n8n lower technical entry barriers [8], enabling the integration of sentiment extraction and predictive modeling into a single process. This study explores whether sentiment-based indicators, implemented in such workflows, can contribute to short-term stock price prediction [1], [3].

The Efficient Market Hypothesis (EMH) assumes that all available information is fully reflected in prices, leaving no room for systematic excess returns [9]. Yet, behavioral finance shows that investor decisions are strongly influenced by irrational emotions and herd behavior, which create predictable market anomalies [10], [11], [12], [13]. With digital platforms accelerating the spread of sentiment, traditional models appear less effective, highlighting the need for approaches that capture emotional dynamics [4], [14].

This paper investigates the potential of sentiment analysis for predicting stock price movements through an automated workflow. The study addresses three main questions:

1. Which data sources (e.g., social media, financial news) provide the most relevant input information for sentiment-based forecasting, and how do they differ with respect to availability, timeliness, and explanatory power?

2. How can n8n be configured to automate data integration, sentiment evaluation, and model comparison, and what practical challenges arise in implementation?

3. To what extent does sentiment analysis contribute to forecasting accuracy, and how do different AI models compare in terms of robustness and predictive quality?

Figure 1 provides an overview of the conceptual framework of this study. Information flows from financial news and social media are captured through sentiment analysis and serve as input for short-term stock price forecasting. The workflow integrates a probabilistic model implemented in n8n, where an AI agent generates and executes logistic regression code. Finally, the modeling performance for prediction with two AI models, OpenAI and DeepSeek, is compared within the workflow. This overview illustrates how the study integrates financial theory with workflow-based automation, demonstrating a prototype for sentiment-based probabilistic modeling.

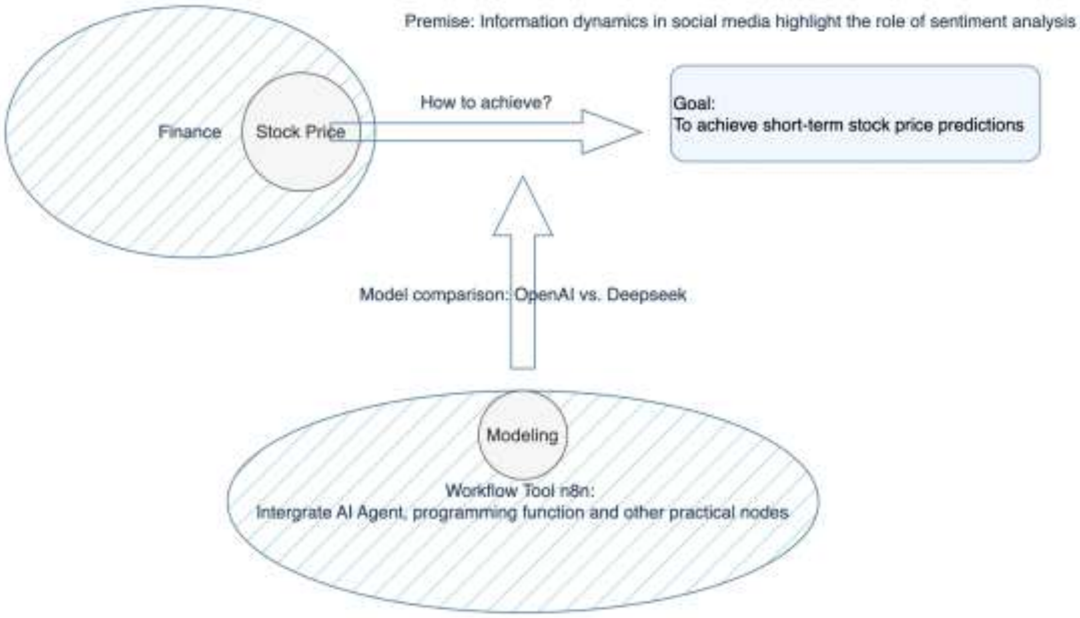

Fig. 1 Overview of research (Source: Own illustration using draw.io.)

The rest of this paper is structured as follows. Section 2 reviews the current state of research on the application of sentiment analysis in finance through artificial intelligence and machine learning, highlighting existing limitations. It further outlines the conceptual framework for this implementation within n8n and provides a detailed account of the operational processes. Section 3 presents the implementation of the workflow deployed on n8n. Section 4 compares the performance of two AI models and provides evaluations. And Section 5 concludes this work and offers suggestions for future research.

II. State of Art

Several studies have examined the application of sentiment analysis in finance using artificial intelligence (AI) and machine learning (ML).

a) For instance, Araci [7] developed the FinBERT model, which was trained specifically on financial news and achieved superior results in sentiment classification compared to generic language models. More recently, Multimodal Stock Price Prediction [20] explored textual data combining tweets, news articles and financial indicators through sentiment analysis using both FinBERT and ChatGP models to improve forecast accuracy.

b) Zhang et al. [15] combined Twitter data with historical stock prices, demonstrating significant relationships between sentiment and price movements. Recent work by Özöğür et al. [21] similarly integrates Twitter sentiment and deep learning models long short-term (LSTM) and one-dimensional convolutional neural network (1D CNN), for stock market prediction, confirming and extending Zhang's findings.

c) Nassirtoussi et al. [1] provided a comprehensive overview of text mining techniques in finance and emphasized the central role of the quality of the news interpretation in modeling market dynamics. Collectively, these studies underscore the potential of integrating sentiment information into financial prediction frameworks. More current reviews, such as Stock Market Prediction Using Machine Learning and Deep Learning Techniques: A Review [22], also discuss extended methods, datasets, and model issues in more depth.

Analysis of the state of the art:

Despite this progress, several shortcomings persist in the existing body of research:

- Lack of transferability and scalability: Many approaches remain at a theoretical level or are limited to historical backtesting, with little evidence of their robustness across diverse market conditions.
- Insufficient practical applicability: Existing models are often too complex or computationally demanding, hindering their adoption in real-world financial analysis.
- Fragmented methodological design: Few studies provide integrated end-to-end workflows that combine heterogeneous data sources, modeling steps, and automation technologies in a coherent and reproducible manner.
- Limited comparative analysis: Most prior work focuses on individual models, without systematically comparing alternative models, particularly in the context of emerging large language models.

III. ANALYSIS OF HOW AI CAN BE APPLIED TO STOCK PRICE PREDICTION

An independent empirical study could advance the state of the art. The aim of this paper is therefore to develop an automated, transparent, and reproducible workflow for sentiment-based modeling of stock price dynamics, using scored sentiment data from financial news. Within this framework, two advanced language models—OpenAI and DeepSeek—are systematically compared. The findings are expected not only to contribute to academic discourse but also to deliver practice-oriented implications for investors and analysts by highlighting the relative strengths and limitations of these models in real-world settings.

The methodological framework of this paper implemented through an automated workflow on n8n. As visualized in figure 2, the process integrates sentiment data and market data into a unified modeling environment, ensuring consistency and transparency of all steps. This modular workflow ensures the automation of data extraction, preprocessing, and predictive modeling, while maintaining flexibility for extension and comparison of different AI models. The operational logic follows three main stages, as shown in figure 3.

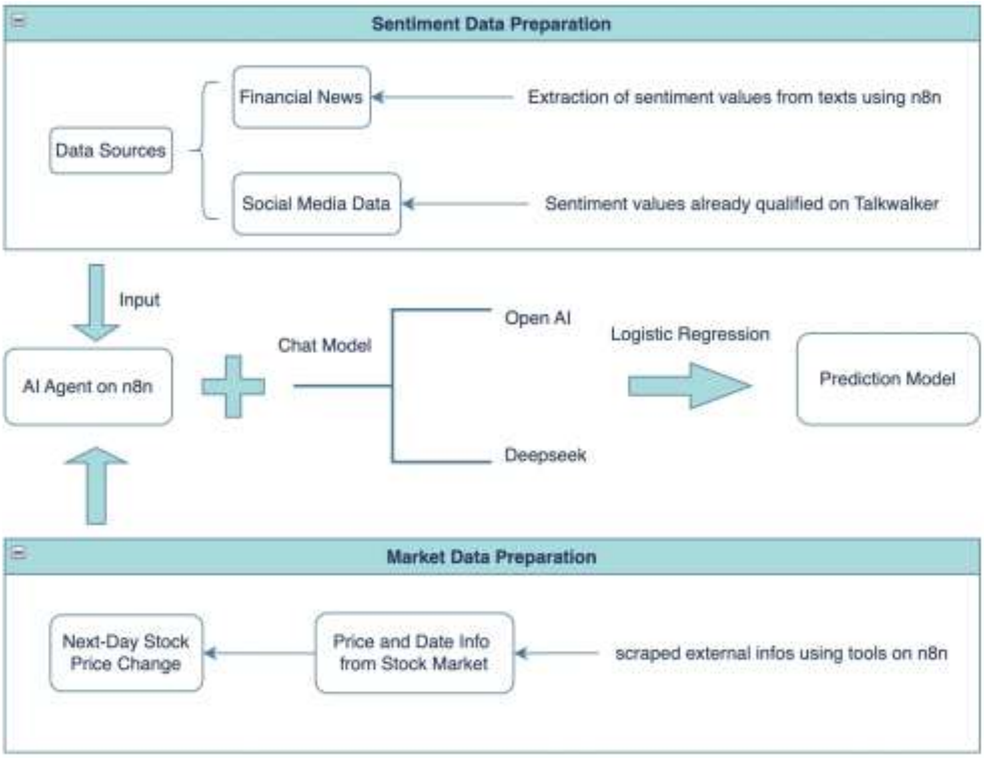

Fig. 2 Overview of the Operational Logic on n8n (Source: Own illustration using draw.io.)

a. Sentiment Data Preparation

- Financial News: Sentiment values are extracted from financial texts using AI models within the n8n workflow.
- Social Media Data: Pre-qualified sentiment values reflecting opinions are directly integrated using Talkwalker.

b. Market Data Preparation

- Historical price and date information from Alpha Vantage are scraped and structured on n8n.
- The next-day stock price change is derived as the binary target variable (1 = increase, 0 = stable/decrease).

c. Modeling and Prediction

- An AI Agent on n8n coordinates the integration of sentiment and market data.
- Chat-based large language models (OpenAI, DeepSeek) process sentiment inputs, which are then passed into a logistic regression model.
- The resulting prediction model estimates the likelihood of next-day price changes based on sentiment and market signals.

At the core of this empirical study is a modular workflow for the automated prediction modeling of stock price developments, based on modeling two different types of data. The process is divided into five steps:

1). Data Preparation

2). Feature Engineering

To forecast stock price movements, this study applies a logistic regression model, which is well suited for binary outcomes such as whether the price rises (1) or does not rise (0) on the following day [16]. The model estimates the probability $P(y=1|x)$ through the logistic function, which maps predictions to the interval [0,1] and thus reflects market dynamics more realistically than a linear relationship [16]. Econometrically, this relies on the logit transformation, where the logarithm of the odds is modeled. A positive coefficient indicates that higher sentiment scores increase the likelihood of a price rise. Model performance was evaluated based on training accuracy, and the final logic was exported as a Python function integrated into the n8n workflow, yielding a reusable and transparent forecasting system.

3). Model Development

Variant A (theoretically ideal): Combination of structured social media sentiment data from Talkwalker and sentiment-related information from financial news.

Variant B (practice-oriented): Structured Talkwalker sentiment data from social media only, without additional financial news.

To ensure data source diversity, this study discusses and attempts Variant a. However, in practice, Variant b was implemented as the basis of the workflow, since system barriers prevented the extraction of sentiment values from financial news within the workflow. Therefore, the focus of this study is on using the existing sentiment scores as workflow input.

4). Automated Implementation in n8n Workflow

5). Validation

To validate the model assumptions, a qualitative analysis is conducted based on selected economic theories and market mechanisms. Using visualizations such as correlation diagrams, the relationship between sentiment values and the probability of stock price movements is examined. The objective is to evaluate the practical significance of the applied AI models and data sources from a behavioral economics perspective.

IV. Validation of the Approach through AI-based Modeling on n8n Workflow

In the next step, the construction of a workflow for AI-supported analysis and forecasting is undertaken.

A. *Data Preparation*

The workflow integrates two data sources: (1) qualitative sentiment scores and (2) historical stock prices. Sentiment data, exported from Talkwalker and uploaded via Google Drive, was reduced to two essential columns—date and sentiment score—using a Set Node on n8n. As seen in figure 3, additional fields such as titles or text were excluded to avoid noise, as their informational content was already embedded in the pre-classified sentiment scores (–5, 0, 5).

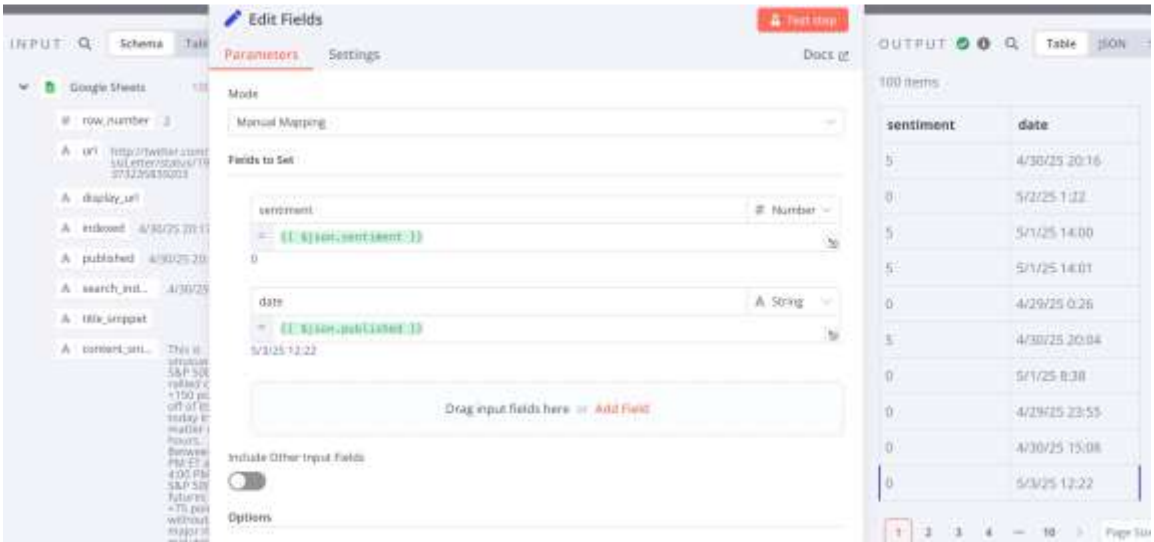

Fig. 3 Operational view in n8n: Filtering of the required columns for sentiment (Source: Own illustration using n8n.)

Near real-time historical stock price data was retrieved from the free Alpha Vantage API, which provides well-structured and high-quality outputs as shown in figure 4.

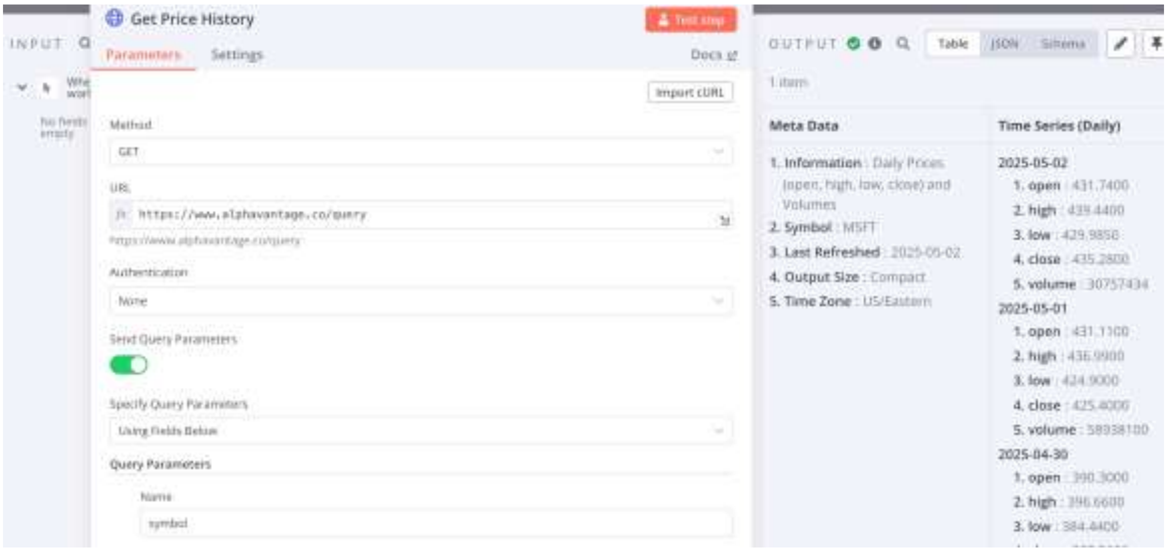

Fig. 4 Operational view in n8n: Retrieval of historical stock price data (Source: Own illustration using n8n.)

These should also be pre-processed within n8n using a Code Node and a Set Node to extract the relevant fields.

Since sentiment timestamps were recorded at minute-level precision while stock data was available only daily, timestamps were standardized to the date level to ensure consistency (see figure 5).

Fig. 5 Operational view in n8n: Standardization of sentiment data timestamps (Source: Own illustration using n8n.)

B. Feature Engineering

The final dataset comprised two variables:

- Independent variable: sentiment score (–5, 0, 5)
- Dependent variable: binary label indicating whether the stock price increased on the following day (1) or not (0).

Because this label, namely, dependent variable, was not included in the original data, it was calculated in n8n using a Code Node, where consecutive daily closing prices were used. For the last day in the series, no label was computed. The resulting dataset—date, close, and label—served as the second input for model training (see figure 6).

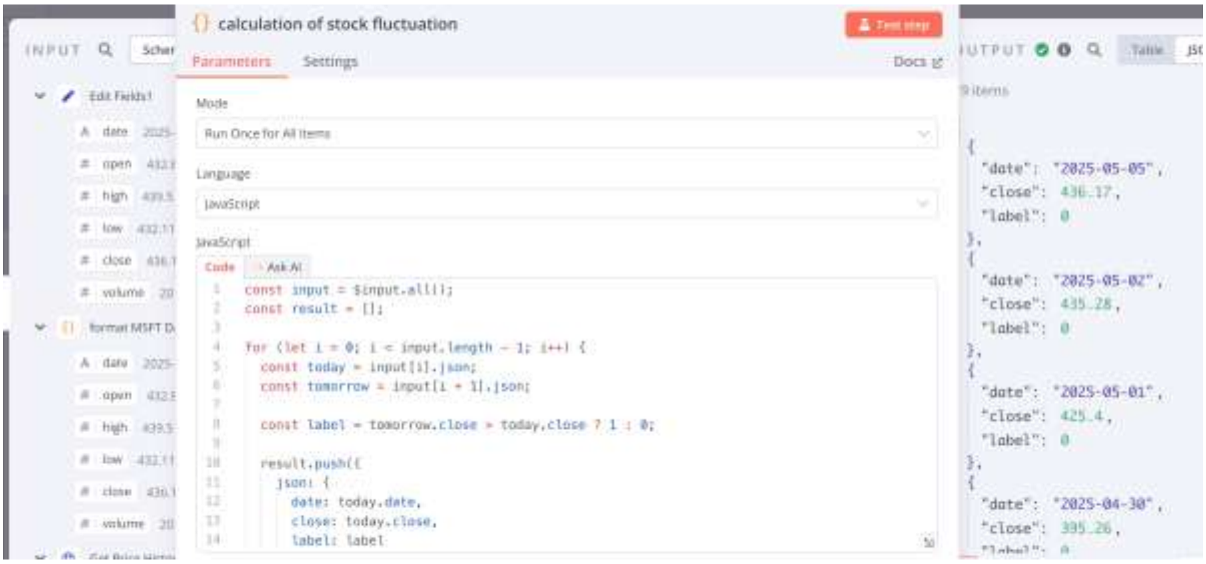

Fig. 6 Operational view in n8n: Calculation of stock price fluctuations (Source: Own illustration using n8n.)

C. Model Development

The two datasets were then merged. Figure 7illustrates the implementation of the workflow deployed on n8n.

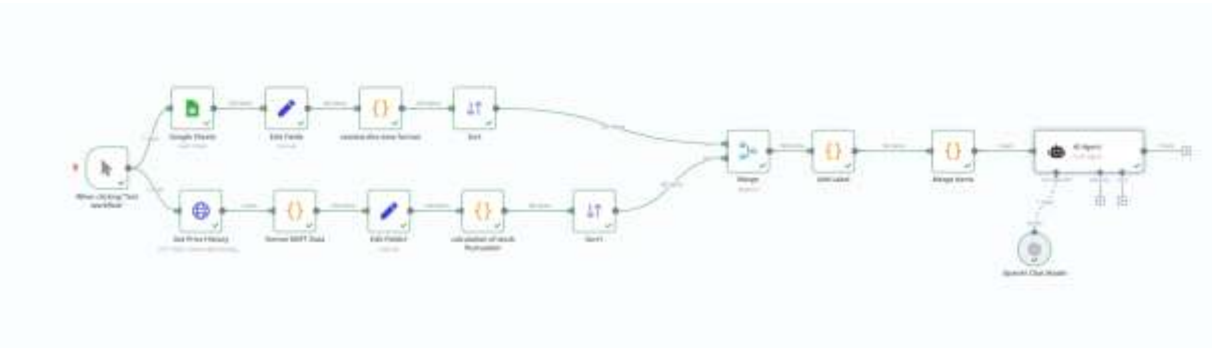

Fig. 7 Automated workflow in n8n: The complete process of model training (Source: Own illustration using n8n.)

A logistic regression model was then applied to learn the mapping:

- Sentiment score → Price movement (1/0)

The instructions for training the regression model were input into the AI agent as follows.

You are an expert machine learning assistant working in a low-resource workflow automation environment. Your job is to output clean and minimal Python code (compatible with Python 3.10), without external libraries like sklearn or numpy. Your code will be executed in a restricted environment, so do not use `input()`, file operations, or print statements. You must return only valid Python code or JSON, without explanation or natural language.

Train a logistic regression model using the following dataset:

[
{"x": -5, "y": 0},
{"x": -3, "y": 0},
{"x": -1, "y": 0},
{"x": 0, "y": 0},
{"x": 1, "y": 1},

{"x": 3, "y": 1},

{"x": 5, "y": 1}

]

Please return only a JSON object in the format:

{

"coef": <float>,

"intercept": <float>

}

This method falls under supervised learning, where historical closing prices provide the ground truth. The trained model outputs decision rules that can be directly embedded into the automated workflow. Although this functionality is not implemented in the current work, it represents a promising avenue for further research.

While the estimated parameters cannot be generalized beyond this dataset, the methodological implementation within n8n demonstrates practical feasibility. The workflow illustrates how low-code automation, combined with AI Agent capabilities, can support lightweight machine learning tasks, bridging data integration and modeling within a single process.

V. RESULT OF APPLYING THIS APPROACH ON A USE STUDY

Without any changes to the remaining input data, the language model integrated into the AI Agent was replaced by DeepSeek. The objective of this modification was to investigate possible differences in model performance and in the generated predictions. The resulting regression coefficients and the intercept are presented in figures 8 and 9.

Fig. 8 Operational view in n8n: Result of training a logistic regression model with OpenAI (Source: Own illustration using n8n.)

Fig. 9 Operational view in n8n: Result of training a logistic regression model with Deepseek (Source: Own illustration using n8n.)

To visualize the model differences, the regression curves estimated with OpenAI ($\beta_1 \approx 0.67$) and DeepSeek ($\beta_1 \approx 1.39$) were compared. The presentation is based on the standard form of the logistic regression function and illustrates how differently each model reacts to variations in sentiment values (see figure 10).

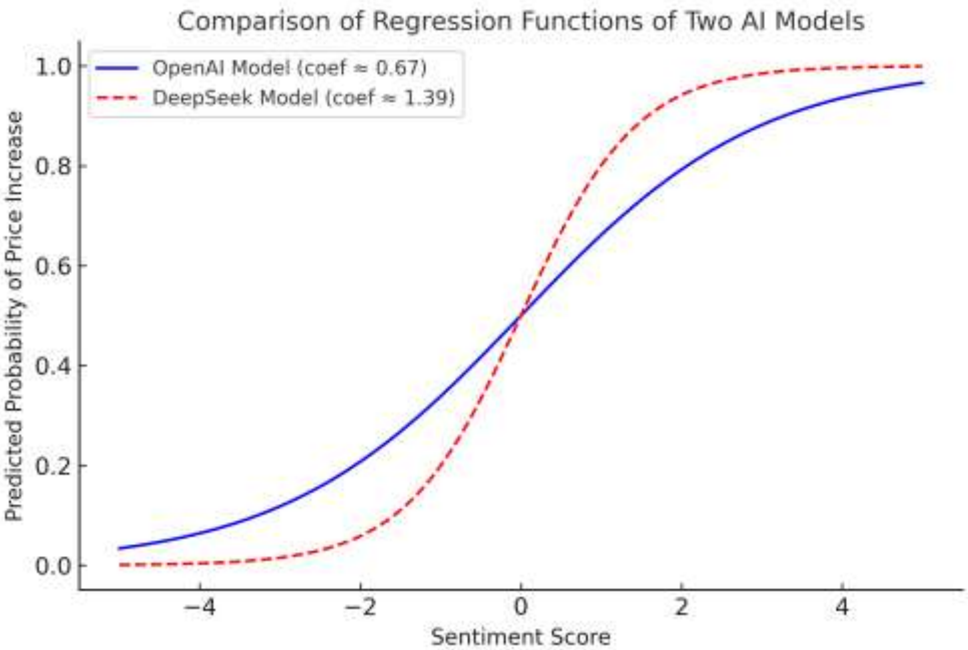

Fig. 10 Comparison of upward probability between the OpenAI and DeepSeek models (Source: Own illustration using n8n.)

From an economic perspective, the contrasting response patterns of OpenAI and DeepSeek represent two distinct approaches to sentiment-based forecasting. DeepSeek, with its steep probability curve, reacts strongly to moderate sentiment changes, resembling momentum-driven behavior and capturing short-term market impulses such as herd effects or overreactions—anticipating the continuation of existing trends and assigning greater weight to current market moods. This makes it particularly useful in dynamic contexts with high informational volatility, where real-time reactions are critical. By contrast, OpenAI only responds significantly under extreme sentiment signals, aligning more closely with a reversal logic that anticipates medium-term corrections of market exaggerations. In terms of risk preferences, OpenAI reflects a more robust and risk-averse stance, whereas DeepSeek corresponds to a risk-seeking, short-term oriented strategy.

Overall, DeepSeek's ability to respond sensitively to emotional peaks supports its use in scenarios where spontaneous reactions, real-time data, and short-term market dynamics are central. This aligns consistently with the theoretical framework of this study, which emphasizes the role of sentiment as a short-term yet market-relevant driver. OpenAI offers a slower but more stability-oriented alternative. The comparison highlights not only technical differences in regression sensitivity but also distinct economic logics underlying the two models.

VI. Conclusion and Future Work

In this paper, a practical and innovative approach for integrating sentiment analysis into financial forecasting has been presented and validated through an n8n-based workflow prototype. As a lightweight and modular platform, n8n enables transparent and flexible exploratory studies while lowering the technical barrier for AI-driven modeling. The study demonstrates that even simple regression models using AI agents can capture sentiment-driven market fluctuations, although the use of extreme sentiment values (–5, 0, 5) limits granularity and may affect predictions in intermediate ranges. This work highlights the potential of combining AI with workflow automation to address limitations of traditional forecasting methods. Future research could enhance prediction accuracy by adopting more complex, nonlinear models, integrating broader datasets, and leveraging advanced frameworks or scalable cloud environments. Ultimately, this architecture could serve as a prototype for a scalable, agent-based, AI-driven financial forecasting system.

Acknowledgment

We gratefully acknowledge that this study was completed as part of a graduation project at the University of Applied Sciences Constance. We sincerely appreciate the opportunity to carry out this research and are grateful for the academic environment, resources, and support that made this work possible.

References

[1] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, and D. P. Ngo, "Text mining for market prediction: A systematic review," Expert Systems with Applications, vol. 41, no. 16, pp. 7653–7670, 2014.

[2] F.-Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting: A survey," Artificial Intelligence Review, vol. 50, no. 1, pp. 49–73, 2018.

[3] Q. Li, T.-J. Wang, P. Li, L. Liu, Q.-X. Gong, and Y.-Z. Chen, "The effect of news and public mood on stock movements," Information Sciences, vol. 278, pp. 826–840, 2014.

[4] P. C. Tetlock, "Giving content to investor sentiment: The role of media in the stock market," Journal of Finance, vol. 62, no. 3, pp. 1139–1168, 2007.

[5] M.-Y. Day and C.-C. Lee, "Deep learning for financial sentiment analysis on finance news providers," in Proc. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 1127–1134, 2016.

[6] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," Journal of Computational Science, vol. 2, no. 1, pp. 1–8, 2011, doi: 10.1016/j.jocs.2010.12.007.

[7] D. Araci, "FinBERT: Financial sentiment analysis with pre-trained language models," arXiv preprint, arXiv:1908.10063, 2019.

[8] M. Schmidt, "Patterns for visual programming: With a focus on flow-based programming inspired systems," Proceedings of the 26th European Conference on Pattern Languages of Programs, 2021.

[9] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," Journal of Finance, vol. 25, no. 2, pp. 383–417, 1970.

[10] A. V. Banerjee, "A simple model of herd behavior," Quarterly Journal of Economics, vol. 107, no. 3, pp. 797–817, 1992.

[11] A. Devenow and I. Welch, "Rational herding in financial economics," European Economic Review, vol. 40, no. 3–5, pp. 603–615, 1996.

[12] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," Journal of Political Economy, vol. 100, no. 5, pp. 992–1026, 1992.

[13] W. F. M. De Bondt and R. Thaler, "Does the stock market overreact?," Journal of Finance, vol. 40, no. 3, pp. 793–805, 1985.

[14] A. W. Lo, "The adaptive markets hypothesis: Market efficiency from an evolutionary perspective," Journal of Portfolio Management, vol. 30, no. 5, pp. 15–29, 2004.

[15] X. Zhang, H. Fuehres, and P. A. Gloor, "Predicting stock market indicators through Twitter 'I hope it is not as bad as I fear'," Procedia - Social and Behavioral Sciences, vol. 26, pp. 55–62, 2011.

[16] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. New York: Springer, 2009.

[17] O. Odeyemi, et al., "Machine learning in financial forecasting: A US review: Exploring the advancements, challenges, and implications of AI-driven predictions in financial markets," World Journal of Advanced Research and Reviews, vol. 21, no. 2, pp. 1969–1984, 2024.

[18] K. J. Olowe, et al., "Review of predictive modeling and machine learning applications in financial service analysis," Computer Science & IT Research Journal, vol. 5, no. 11, pp. 2609–2626, 2024.

[19] M. R. Mahmud, et al., "Machine learning-powered financial forecasting in the US tourism industry: Predicting market trends and consumer spending with artificial intelligence," Journal of Computer Science and Technology Studies, vol. 7, no. 2, pp. 13–22, 2025.

[20] F. Karadaş, B. Eravcı, and A. M. Özbayoğlu, "Multimodal stock price prediction," arXiv preprint arXiv:2502.05186, 2025.

[21] S. Ö. Akyüz, P. K. Ataş, and A. Benkhaldoun, "Predicting stock market by sentiment analysis and deep learning," Operations Research and Decisions, vol. 34, no. 2, pp. 85–107, 2024.

[22] M. Saberironaghi, J. Ren, and A. Saberironaghi, "Stock market prediction using machine learning and deep learning techniques: A review," Predictive Analytics, vol. 5, no. 3, pp. 76, 2025.