Acta Informatica Pragensia
2026, Volume 15, Issue 2, 307–326
https://doi.org/10.18267/j.aip.299

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

Fairness-Aware Multimodal Machine Learning for Retail Stock Prediction from Sentiment and Market Data

Sanjay Rastogi [1], **Kamal Upreti** [2], **Uma Shankar** [3], **Pravin Ramdas Kshirsagar** [4], **Tan Kuan Tak** [5], **Rituraj Jain** [6], **Ganesh Veluswwamy Radhakrishnan** [7]

[1] School of Management, GLA University, Mathura, India
[2] Department of Computer Science, CHRIST (Deemed to be University), Delhi NCR Campus, Ghaziabad, India
[3] Department of Financial Technology, University of Wollongong, India Branch Campus, Ahmadabad, Gujarat, India
[4] Department of Electronics and Telecommunication, J D College of Engineering & Management, Nagpur, Maharashtra, India
[5] Engineering Cluster, Singapore Institute of Technology, Singapore
[6] Department of Information Technology, Marwadi University, Rajkot, Gujarat, India
[7] Department of Economics and Finance, Kalinga School of Management, Kalinga Institute of Industrial Technology, Bhubaneswar, India

Corresponding author: Kamal Upreti (kamal.upreti@christuniversity.in)

Editorial Record

First submission received:
August 3, 2025

Revisions received:
September 25, 2025
October 30, 2025
December 4, 2025

Accepted for publication:
December 6, 2025

Academic Editor:
Adam Borovicka
Prague University of Economics and Business, Czech Republic

This article was accepted for publication by the Academic Editor upon evaluation of the reviewers' comments.

How to cite this article:
Rastogi, S., Upreti, K., Shankar, U., Kshirsagar, P. R., Tak, T. K., Jain, R., & Radhakrishnan, G. V. (2026). Fairness-Aware Multimodal Machine Learning for Retail Stock Prediction from Sentiment and Market Data. *Acta Informatica Pragensia*, 15(2), 307–326. https://doi.org/10.18267/j.aip.299

Abstract

Background: The introduction of retail investors to AI-powered trading platforms and especially on emerging markets, has resulted in a new set of risks linked to algorithmic bias and financial forecasting fairness. Social media sentiment and structured data multimodal strategies have demonstrated a potential, but frequently do not have ethical considerations.

Objective: This work proposes a multimodal model predictive control (MPC) framework grounded in fairness-based forecasting of next-day returns on stock in stock market settings, particularly ethical behaviour and transparency of the model on retail markets.

Methods: We combine BERT-based sentiment analysis of Reddit discussions and organized stock market indicators and use XGBoost as the fundamental model. Bias is measured using fairness metrics, including demographic parity difference and equal opportunity difference. Debiasing measures such as reweighting and stratified calibration were used to curb the differences in stock categories.

Results: The first model has an overall accuracy of 72.3 with the highest accuracy of 83.1 in the case of Tesla – representing bias in the model. Fairness assessment shows some significant differences (DPD=0.23, EOD=0.31), but the mitigation decreases to 0.07. However, the massive performance improvement after adjustment brings up the issue of overfitting or fairness overcorrection.

Conclusion: While the proposed debiased framework successfully reduces algorithmic bias, the trade-off between fairness and generalizability underscores the need for caution. These results hold significant implications for digital trading systems and regulatory frameworks of emerging economies such as India, where explainability and fairness of AI models are significant for ethical financial engagement.

Index Terms

Algorithmic fairness; BERT embedding; Demographic parity; Financial machine learning; Hybrid model; Stock prediction.

1 INTRODUCTION

In digital finance, digital literacy is essential for empowering individual investors to successfully handle new opportunities. Investors can read data, assess risks and make well-informed decisions because of a combination of technological proficiency, analytical thinking and moral awareness (Gomber et al., 2017).

The growing share of retail investors in the trading volume of stock markets highlights the fact that behavioural biases and digital finance are interconnected and affect the decision-making process of digital financial transactions (Ozili, 2018). To increase the effectiveness of trade and financial services, one should learn about cognitive biases in digital finance (Khan and Shabbir, 2024). The high rate of digitalization and the increased application of artificial intelligence (AI) in various aspects of life influence the role of organizations and the contribution of individuals to the routine to a considerable degree. They also offer an abundance of new opportunities, which have been observed on the world's financial markets (Saxena & Yasobant, 2020). The way financial markets function has changed as a result of digitalization, which has also made trading between buyers and sellers of different financial products more efficient, quicker and less expensive. However, given that financial markets are not as efficient as the efficient market hypothesis suggests, there is a good chance that different AI applications will identify hidden patterns in the behaviour of financial instruments and offer helpful information for financial decision-making.

Nevertheless, there are also a lot of obstacles that come with the potential (Addy et al., 2024). With millions of search results available in a second, the digital era has brought forth an overwhelming amount of information. This has rendered information use and assimilation difficult. Despite such issues, it has also raised evidence-gathering solutions. It is presumed that the degree of information overload would increase along with the degree of research and development dedicated to this issue (Ozili, 2023). It is likely to acquire different shapes and sizes due to the new technology and self-publishing techniques (Bunkar & Ramaiah, 2024). Figure 1 is an illustration of the two-fold effect of digitalization on the financial arena, with two key branches, namely globalization and online banking. Digitalization under globalization has empowered customers to develop banking relations at international levels and enjoy global investment prospects with no geographic barriers (Egorov, 2022). Conversely, online banking focuses on the technological transformation of mobile-based financial operations and increased transparency. This transformation will enable users to make investments with the help of mobile devices, being able to have a more transparent insight into their financial activities. Cumulatively, these elements indicate the role of digitalization in transforming the accessibility and utility of contemporary banking and investment systems (Barman et al., 2022).

Over the past few years, economic innovation combined with a high rate of technological advancements has altered business operations, mostly due to the introduction of the algorithm trading system. This is how these technology-driven processes have first penetrated into various spheres of everyday life – the online shopping and travel booking processes and the interaction with search engines – performing, ubiquitous and effective (Mwangi and Njoroge, 2024). One outstanding achievement of this progress was the improved execution service algorithm trading (AT) platform of Credit Suisse in India introduced on 22 June 2009. It was nearly ten years before algorithm trading would control three quarters of trades in the United States; the same became true of India five years later. This rapid development raised red flags with regulatory bodies such as the Securities and Exchange Board of India (SEBI), which showed that there was a systemic risk and a moral issue of selling automatic strategies on investment platforms (Vasquez & Cross, 2024).

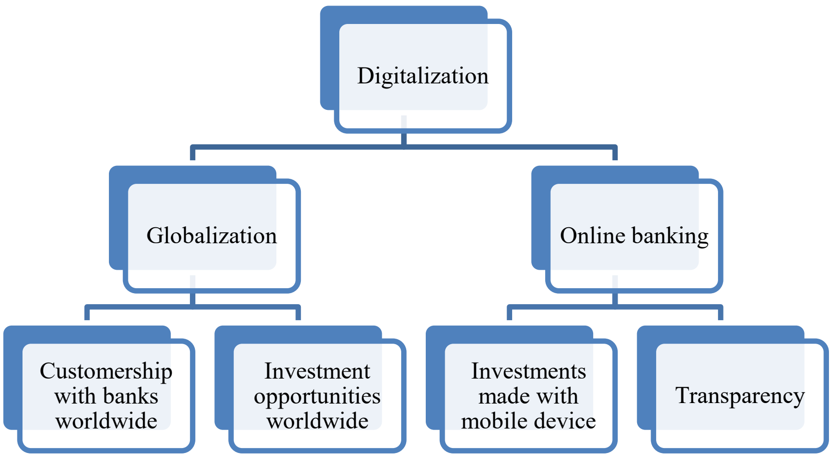

Figure 1. *Structural components of financial digitalization.*

Our conceptualization of ethics in this research is based mostly on the framework of algorithmic fairness, transparency and responsible AI practices in the retail trading settings. It has already been demonstrated in the previous literature that financial prediction models can be unintentionally biased in favour of popularity or favourable information, which results in disproportionate errors across the sample of companies and misleading indications for inexpert investors (Barocas et al., 2019; Kleinberg et al., 2019). Fairness-conscious modelling plays two important roles, namely (i) improving the statistical equity of groups by minimizing the predictive difference of different groups, and (ii) improving the transparency and responsibility of AI-based decision support to trading platforms. Incorporating equity in the process of predicting stocks, we are fulfilling more widespread ethical AI principles and also meeting practical investor protection issues on retail financial markets.

1.1 Study objective

The primary goal of this study is to develop a framework for truly covered hybrid machine learning to predict stock price movement by integrating both structured market data and unnecessary data obtained from social media. Given the growing dependence of retail investors on AI-assisted trading platforms and emotion-based forecasting equipment, the study wants to identify and address the inherent prejudices that may arise from data imbalance, emotional or platform-specific polarization of some companies. In particular, the model aims to merge the semantic features by using a BERT from Reddit-based economic discussions with traditional numerical indicators such as open, high, low, close, volume and sentiment scores. Through this approach, the intention of the study is not only to increase the accuracy of the share prediction, but also to systematically evaluate and reduce algorithm prejudice using justice measurements. Overarching goals are to ensure development of explainable, scalable and morally responsible AI systems for the retail environment.

1.2 Study contribution

This research makes many major contributions to the fields of economic machine learning and algorithmic fairness. Firstly, it introduces a new multimodal merger strategy that combines a semantic model that occurs from domain-specific economic discourse with standardized numerical stock indicators, resulting in a strong 776-dimensional functional space. Secondly, volume structural prejudice is identified in stock prediction models – especially in favour of Tesla due to high visibility and data volumes – using justice assessment measurements such as demographic gender equality differences and equal opportunity differences. Thirdly, it implements individual modelling techniques to show an average shortage in these inequalities without compromising the general prediction capacity. In addition, by taking advantage of size-based interpretation, the frame model improves openness and allows stakeholders to explain the contribution and understand the distortion. Finally, the research emphasizes the moral implications of distributing sentiment-augmented AI models in retail financing and providing a general function to develop fair and explanatory forecasting systems within the digital economic ecosystem. Table 1 provides a comparative observation of existing research into algorithm bias, emotional intelligence and digital herd sets when it comes to making financial decisions. This emphasizes how the proposed study extends the preparation by integrating fairness-aware hybrid modelling that combines social feelings and structured market data to address prediction bias. Our comparison emphasizes the new contribution of this article, which corresponds to moral AI practice with retail investor protection.

1.3 Literature review

1.3.1 Biases in AI-driven trading decisions

Recent studies emphasize a significant change in the financial industry due to the integration of AI, especially in algorithm trade and risk management (Luchian & Strat, 2024). As the financial markets become increasingly complicated, the distribution of AI technologies has become necessary to increase scalability in operational efficiency, future accuracy and decision-making processes (Saelee & Pankham, 2024). Several studies have documented how financial institutions and investment companies utilize AI to improve market forecasts and exposure to financial risks (Liu et al., 2025; Min & Borch, 2022). Despite these advances, scholars have also attracted attention to the emergence of algorithmic and systemic prejudices within the AI model. This bias can distort decision-making, introduce unfair benefits or increase market volatility, which may increase concerns over the moral and regulatory dimensions of AI distribution in finance (Mwangi & Njoroge, 2024).

Regarding development at the increasing intersection of technology and finance, recent research has also emphasized the dimensions of behaviour when making investment decisions, especially in the context of social media. A qualitative case study (Sathya & Prabhavathi, 2024) discovered how social media platforms affect investment decisions in online communities. Based on Bandura's social learning theory, the study highlighted the role of observation learning and behavioural modelling in the design of investor works. According to this theory, it learns and adapts to individuals by looking at others in their social environment, especially when this behaviour receives favourable results. In the digital age, social media have revolutionized the way people reach out and stock gangs are now interacting in real time, sharing opinions, marketing insights and even speculative strategies (Agrawal et al., 2015). This open, rapidly moving exchange of information creates a feedback loop where individuals can mimic popular trade behaviour, sometimes without fully evaluating the associated risks. Small, digital domestic investors dominate online forums and investment apps, the impact of social media on trade decisions increases, and AI-operated economic systems strengthen the importance of addressing both psychological prejudices and technical effects within the broad structure of financial systems (Thakkar, 2024).

1.3.2 Emotional decision-making in financial contexts

Emotional intelligence (EI) plays an important role in economic professionals and investors who are able to make sound and balanced decisions (Bouzguenda, 2018). This can be defined as a person's ability to recognize, understand and handle their feelings and their association with the feelings of others. This emotional awareness promotes the development of positive relationships and helps maintain composition under unstable economic conditions. The most important components include mindfulness, self-insight, sympathy, inspiration and social ability. Mindfulness and self-consciousness enable individuals to stay grounded and identify emotional triggers which can give rise to a decision (Dibb et al., 2021). Controlling stress and learning how to focus are techniques to handle emotion; inspiration is the skill to define apparent objectives and follow them constantly. Sympathy enables one to perceive the attitudes and emotional phases of others, which are required to make cooperation choices (Ran et al., 2021). In the meantime, social potential, professional and social atmosphere represent a person’s ability to communicate successfully. In terms of financial behaviour, a recent study carried out in Thailand (Riefel, 2024) has examined the relationship between social media and emotional intelligence and investment decisions on the stock market. In the study, seven dimensions of emotional intelligence were recognized as influential factors in the process of interpreting the market indicators and engaging in online economic communities by the investor (Pal et al., 2024). These lessons describe the significance of emotional control and social consciousness in terms of coping with the complicated, fast-paced world of digital trade and social awareness.

1.3.3 Digital herd mentality in investment decision-making

Digital herd mentality refers to the tendency of individuals to follow the opinions or decisions of a large group, especially in the digital and online environment. In the context of investment and real estate markets, buyers and consumers are often dependent on opinions on social media, trend discussions and independent analysis instead of colleague behaviour (Thakkar, 2024). It can give rise to emotionally charged, group-influenced decisions that cannot always be adapted to basic market principles. Real estate agents and agencies affect this behaviour by shaping it through targeted marketing strategies, using digital examples and customer perceptions through motivational materials on platforms (Suresh G., 2024). The digital environment increases these effects, forming a feedback loop where viral meaning or emotional change can change market mobility quickly. In addition, emotional reactions in digital spaces – individual branding and online identity – play an important role in the design of perceptions. The emotional tone of the material, whether optimistic or frightened, can affect public behaviour (Nerlekar et al., 2024). Regulatory bodies and decision makers should consider these behavioural trends when designing the structure that promotes stability and informed decision-making on fast digitalized markets (Zhao, 2024).

To guide the study, a set of concentrated research questions was designed, which addresses important concerns in emotion-driven stock prediction and the righteousness of algorithms. The aim of these questions is to find out the role of model clarity in the effect of social media spirit, the presence of prejudice in the AI model, the efficiency of justice reduction and financial prognosis. Table 2 presents a structured observation of these research questions, the argument behind each and the acting strategies adopted in this study to address them. This arrangement ensures that our research is theoretically rooted and practically relevant to moral AI distribution on financial markets.

***Table 1.** Comparative analysis of literature on bias and behaviour in financial decision-making with the proposed fairness-aware approach.*

Aspect	Biases in AI-driven trading decisions	Emotional decision-making in finance	Digital herd mentality in investing	Proposed approach (this study)
Key focus	AI-induced bias in trading algorithms	Emotional intelligence in decisions	Groupthink via social media	Fairness-aware hybrid model combining sentiment + market data
Problem highlighted	Algorithmic bias favouring certain stocks	Emotion-driven irrationality	Viral sentiment causing irrational trading	Sentiment overrepresentation skewing prediction results
Main influencing factor	Model training data, lack of fairness checks	Emotional traits such as mindfulness, empathy	Peer behaviour and influencer impact	Reddit-based sentiment skew + imbalanced company data
Relevant theories/models	Fairness in AI, bias in algorithmic trading	Emotional intelligence, social learning theory	Herding theory, social proof	Demographic parity, equal opportunity difference, SHAP
Platforms studied	Trading algorithms in finance	Investors interacting on social media	Social trading platforms and real estate forums	Reddit + historical stock market data (Tesla, Apple, Amazon)
Research examples	Mwangi & Njoroge (2024), STRAT (2024)	Bouzguenda (2018), Riefel (2024)	Nerlekar et al. (2024), Thakkar (2024)	This study
Impact on investors	Unfair model outputs harm equity in access	Emotionally reactive decisions	Mass panic buying/selling	Biased predictions may mislead; fairness improves trust
Regulatory implications	Need for AI audits and ethical standards	EI training for better judgment	Policy oversight on social trading influence	Framework for integrating fairness metrics into financial AI

***Table 2.** Research questions, rationale and methodological approach.*

Research question	Rationale	How this study addresses it
RQ1: How can sentiment from social media influence stock price prediction accuracy?	Retail investors increasingly rely on platforms such as Reddit, where optimistic sentiment may bias predictions towards certain companies.	Utilizes BERT to extract semantic sentiment from Reddit posts and fuses it with market data to measure predictive impact.
RQ2: Do stock prediction models exhibit fairness bias across companies?	Overrepresentation of certain stocks (e.g., Tesla) may cause AI models to perform unevenly, disadvantaging others.	Measures fairness using demographic parity difference and equal opportunity difference across companies.
RQ3: Can fairness-aware techniques mitigate prediction bias without compromising model accuracy?	Debiasing often improves fairness but poses risks of overfitting or losing predictive power.	Applies stratified reweighting and fairness calibration, followed by evaluation of accuracy and fairness metrics before and after debiasing.
RQ4: How can explainability tools improve trust in AI-based financial models?	Investors and regulators need transparent insights into how models make decisions.	Uses SHAP values to interpret feature importance and detect bias origin, supporting explainable AI.

2 METHODOLOGY AND FRAMEWORK DESIGN

The proposed framework aims to model stock price movements by integrating numerical financial indicators with semantic sentiment features extracted from Reddit posts. In line with design science artifact typologies (Gregor & Hevner, 2013), we use the term "framework" to denote a modular, end-to-end pipeline. Our contribution specifies a structured process that integrates data acquisition, feature engineering, multimodal fusion, predictive learning and fairness diagnostics. Unlike a monolithic approach, where a single name often describes one specific algorithm or technique, building blocks in this framework (e.g., any text encoder, learner or calibration strategy) are interchangeable and come together in a meaningful architecture. This modularity and extensibility justifies it being a framework and not just a method. The methodology consists of six interconnected stages as shown in Figure 2: data acquisition and pre-processing, exploratory data analysis, semantic representation using BERT (Caragea et al., 2020), feature fusion and transformation, predictive modelling using XGBoost (Zolotareva, 2021), post-hoc fairness diagnosis and correction.

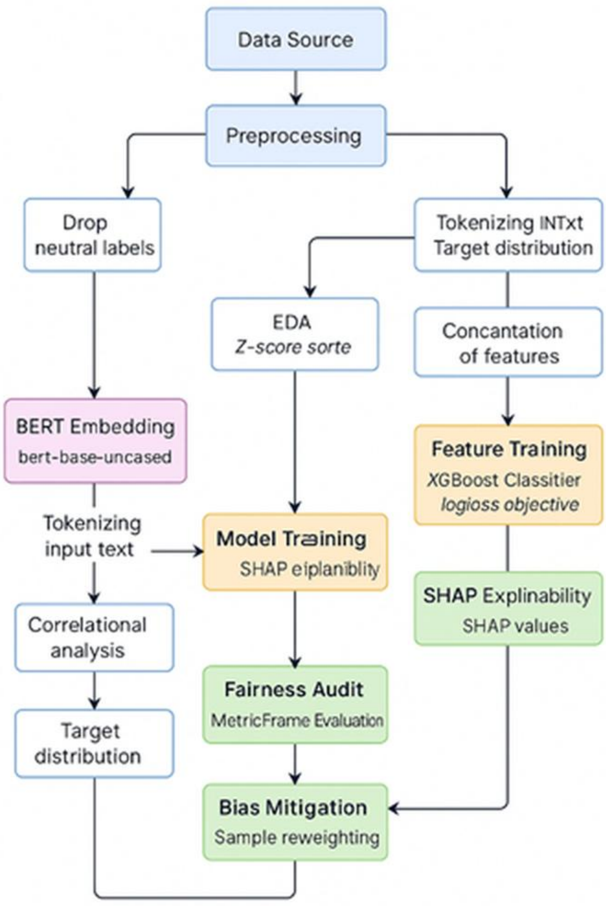

Figure 2. *Methodological framework.*

2.1 Data acquisition and pre-processing

The study employs the publicly available Kaggle dataset "Stock Price Prediction with Sentiment Analysis" (Rootpi3, 2024), which combines structured daily stock indicators with sentiment features extracted from Reddit posts. The dataset has 12,069 records for Tesla, Apple and Amazon. In order to ensure credibility and quality, we performed a random 5% record audit to check type consistency, range validity and null anomaly. We excluded records with missing cleaned text fields and removed neutral labels in the target variable for binary classification. Although the dataset has not yet undergone extensive peer-reviewed research, the open accessibility is high and our verification provides robustness. The scope of the dataset is constrained to three companies – Tesla, Apple and Amazon – as specified in the source from Kaggle (Rootpi3, 2024). These companies have been included because their daily stock data are combined with Reddit sentiment activity. This limits generalizability to other industries; yet the companies in our sample are heavily traded, well covered and often the target of retail investors, and thus serve as a constructive sample for testing of fairness-aware prediction models.

The dataset comprises 12,069 records of three significant companies, namely Tesla, Apple and Amazon, covering a daily period. Numeric features include open, high, low, close, volume; sentiment-based features are comments, score, sentiment, sentiment score and a cleaned text column from Reddit website. The dependent variable *target* (at the beginning) codes stock movement as 1 (down), 0 (neutral) and 1 (up). Neutral cases were removed for binary classification and the target was again encoded in the binary format of upward movement (1) and non-upward movement (0). Finally, records with null values in the cleaned text column were removed in order to maintain semantic consistency. All numerical features were standardized using z-score normalization, where X is the feature value, μ is the mean and σ is the standard deviation of the feature, as shown in Equation (1):

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

Comments refers to the number of Reddit posts per day related to each company. These raw counts were summed up to the trading-day level to match up with stock market variables. All sentiment descriptors (comments, score, sentiment score) and financial metrics (open, high, low, close, volume) were normalized before the analysis to make sure that everything was comparable. Pearson correlation coefficients were then calculated by overlapping of daily time series and are the basis for the correlation analysis that is shown in the Results section.

2.2 Exploratory data analysis

This phase focuses on understanding the structure, distribution and potential biases embedded in the dataset. A systematic examination was performed across key dimensions, company-level balance, sentiment quality, temporal price behaviour and feature interdependence, to inform model design, fairness strategy and downstream pre-processing. Each observation is supported with corresponding visual analytics and numerical interpretation.

The proposed framework is presented as a generalizable pipeline independent of any single dataset. It consists of six stages: (i) data acquisition, (ii) preprocessing and feature engineering, (iii) sentiment representation through BERT embeddings, (iv) multimodal fusion of numerical and textual features, (v) predictive modelling with XGBoost, and (vi) fairness diagnostics and calibration. This consolidated description defines the theoretical foundation of our approach, which is later implemented on a real-world dataset (Tesla, Apple, Amazon) in the empirical section.

2.2.1 Company-level distribution of observations

To assess whether the dataset supports fair and representative modelling across entities, a company-wise distribution analysis was conducted. The dataset consists of 12,069 records: 4,292 for Tesla, 4,001 for Apple and 3,776 for Amazon; these proportions are presented in Figure 3.

$$Company\ Share_i = \frac{N_i}{\sum_{i=1}^{k} N_i} \tag{2}$$

While the distribution is not severely imbalanced, Tesla's overrepresentation may implicitly steer the model towards learning Tesla-specific patterns more effectively than others. The company-wise proportion is computed as Equation (2), where N_i denotes the sample count for the i^{th} company and $k = 3$ in this case.

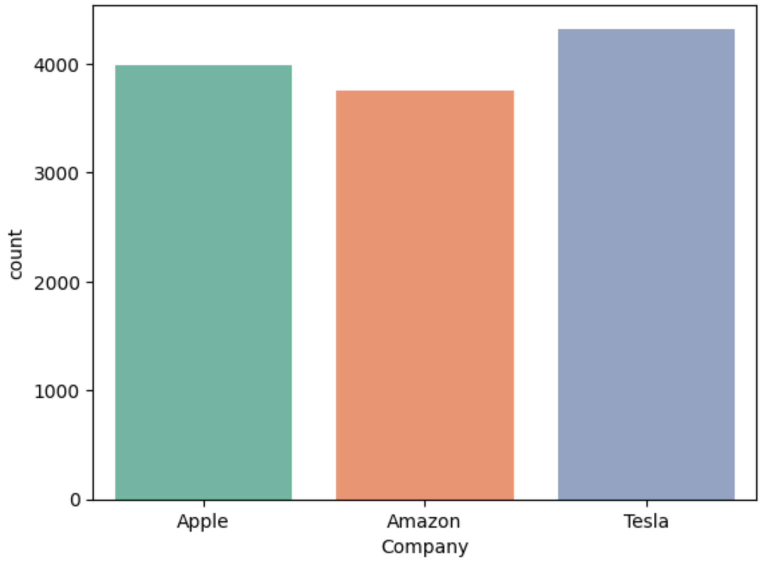

Figure 3. *Number of records per company.*

2.2.2 Sentiment label and score distribution

Knowledge of the distribution and skewness of sentiment labels is an important consideration for the estimation of potential label bias. The dataset adopts the sentiment classification of positive, neutral and negative (Jun Gu et al., 2024), which is obtained through FinBERT, or continuous sentiment scores ranging between 0 and 1. There is a right-skew and most of the sentiment scores range above 0.9 with sentiments with the positive label being the most common. In order to obtain a measure of this skewness, the standardized third moment of the distribution of sentiment scores is obtained as Equation (3), where $\bar{x}$ is the sample mean and s is the sample standard deviation. High skewness (> 2) indicates over-optimism which may be due to the behaviour of the Reddit community or model miscalibration.

$$Skewness = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3 \quad (3)$$

To investigate the distribution and polarity of investor sentiment, both categories of sentiment and scores of intensities of sentiments were analysed. Sentiment labels are calculated with FinBERT and contain three discrete classes, namely positive, neutral and negative, and a continuous sentiment score, which ranges from 0 (pessimistic) to 1 (highly optimistic).

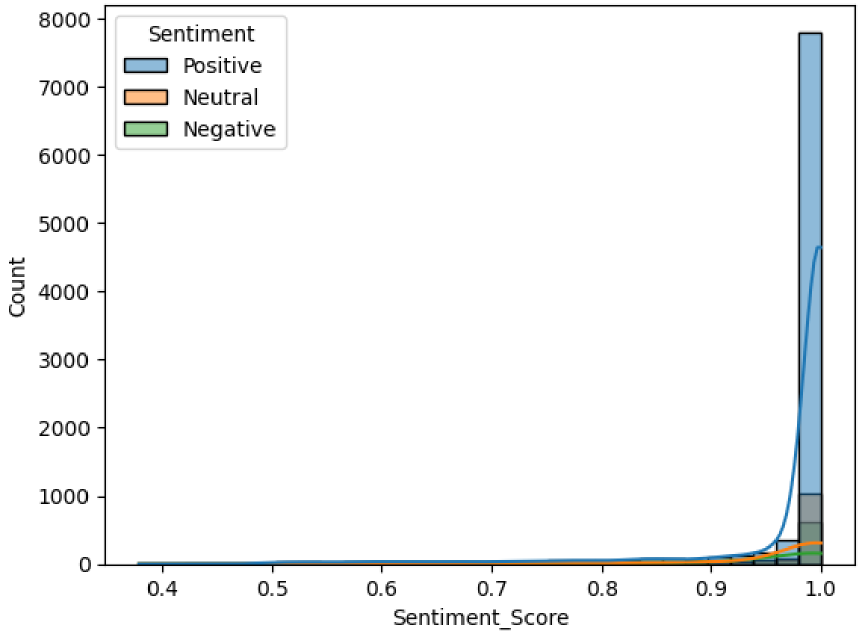

Figure 4. *Distribution of sentiment scores by sentiment type.*

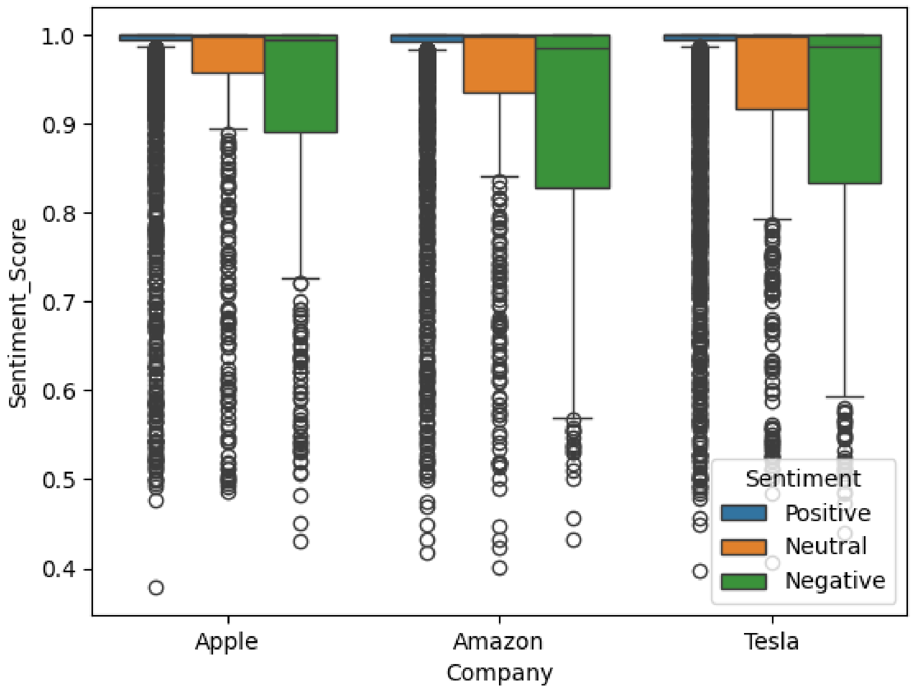

Figure 5. *Sentiment score distribution by company and sentiment.*

While the class labels provide categorical understanding, the sentiment score provides more granular understanding of the mood strength of investors. The kernel density plot in Figure 4 shows that the distribution is very skewed towards optimism with most scores falling between 0.9 and 1.0, particularly under the positive label. There is a natural positivity bias in Reddit discourse, which may be a result of a blend of optimism within the community or by a combination of selection by voting.

Figure 5 further stratifies sentiment scores by company and can demonstrate that Tesla consistently gets a higher average sentiment score – even in the neutral and negative categories. The pervasive positive sentiment skew in Tesla posts suggests that a company-specific sentiment amplification has occurred, which could bring implicit bias into the learning of the model if this is not appropriately managed by fairness-aware mechanisms.

2.2.3 Temporal trends in closing price

To assess the market change over the years, the closing prices were plotted for all three companies. This provides information on past volatility, long-term growth and trend alignment across the firms. Tesla has strong upward trends after 2019, corresponding with speculative investor demand and media hype. To identify the level of temporal smoothing and to eliminate noise, a 30-day simple moving average (*SMA*) was calculated as Equation (4), where P_t is the daily closing price.

$$SMA_{30}(t) = \frac{1}{30}\sum_{k=0}^{29} P_{t-k} \quad (4)$$

To comprehend the temporal dynamics in stock performance, we analyse both the raw closing prices and the smoothed versions of the same using 30-day simple moving average (*SMA*). The closing price trajectories, for visualisation in Figure 6, signify the different market behaviour of the three firms. Tesla's price series shows increased volatility, with sudden jumps and falls, especially since 2019.

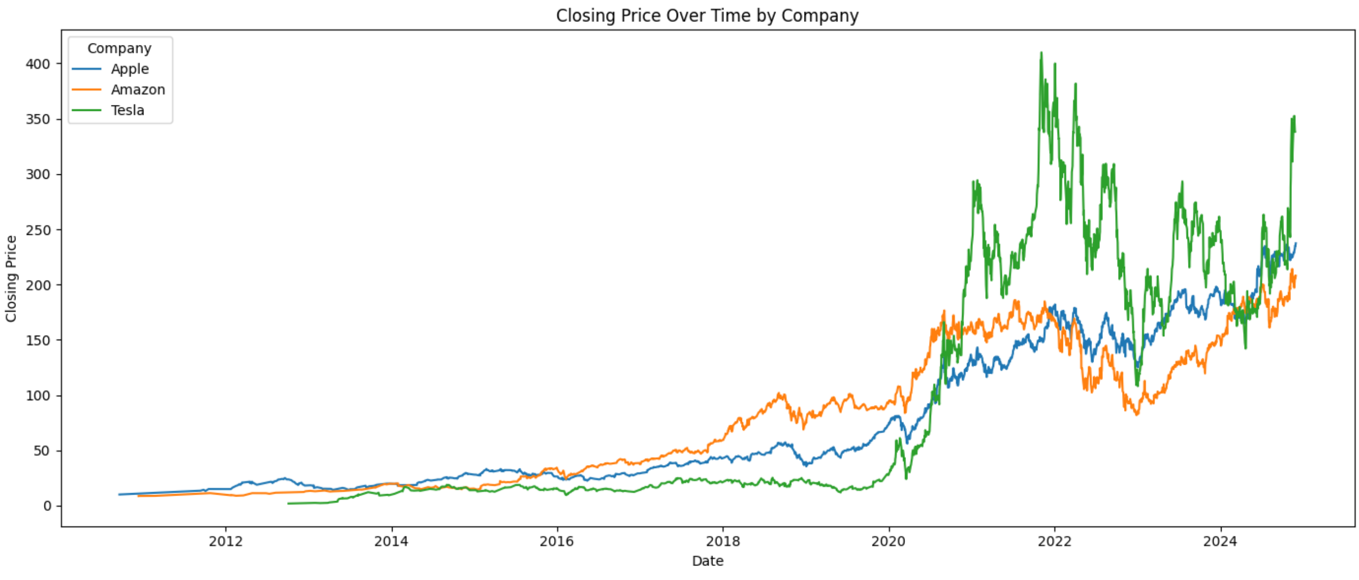

Figure 6. *Closing price over time by company.*

In contrast, Apple and Amazon show relatively more stable and gradual upward trends. To better isolate long-term momentum and suppress short-term fluctuations, a 30-day SMA was applied to each series. As shown in Figure 7, Tesla's price curve exhibits greater deviations from its moving average baseline, reflecting periods of market exuberance and correction cycles. These fluctuations suggest that Tesla is more sensitive to speculative sentiment and retail trading patterns, which reinforces the need for volatility-aware modelling and sentiment integration in prediction tasks.

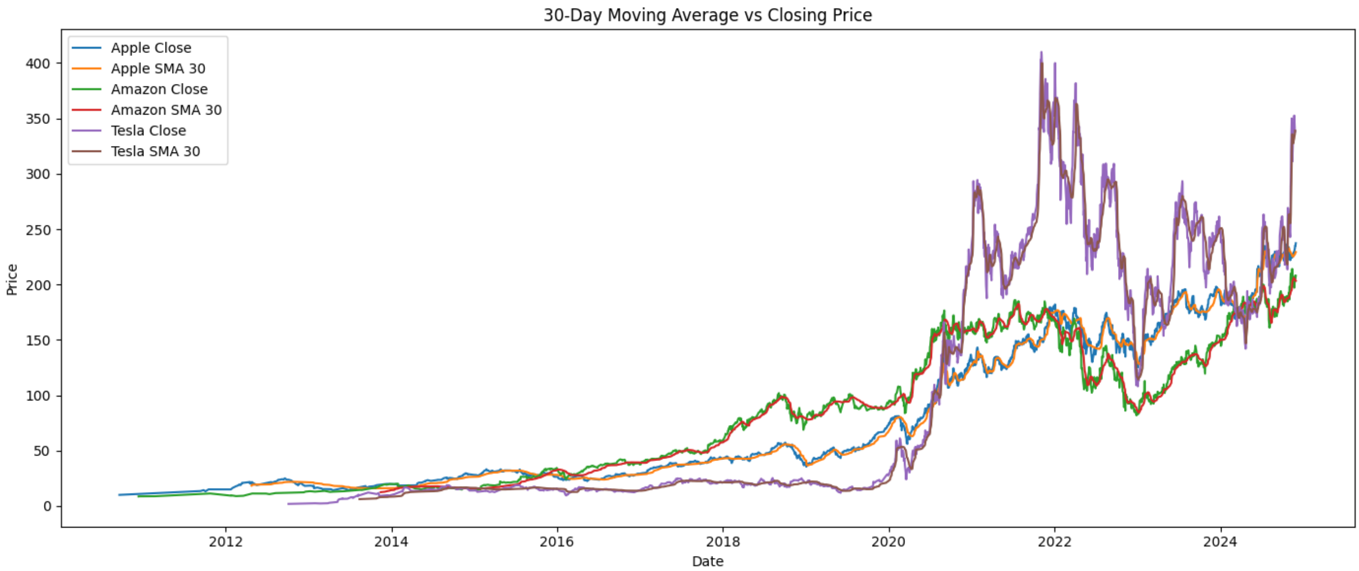

Figure 7. *30-day moving average versus closing price.*

2.2.4 Market behaviour

In order to capture daily trade dynamics and the short-term market, the candlestick diagram was used for each company, providing visual insight into the open, high, low and close (OHLC) price movements. These diagrams are widely employed to explain the market in financial analysis, value action instability and specific trade windows. Various instability appears in patterns, as shown in Figures 8-10. Apple shows relatively compact candlesticks with low extreme weeks, indicating stable value movement and low daily volatility. Amazon's diagram reveals moderate ups and downs with continuous cyclical patterns. Conversely, Tesla's candlestick diagram continuously shows large candles and extended weeks, showing aggressive trade behaviour and frequent reversal. This is clearly in line with the Tesla social media spirit and speculative investor interest, and suggests the need for a strong instability management mechanism in the modelling structure.

Figure 8. *Intra-day price behaviour and volatility pattern for Apple.*

Figure 9. *Intra-day price behaviour and volatility pattern for Amazon.*

Figure 10. *Intra-day price behaviour and volatility pattern for Tesla.*

2.2.5 Feature correlation and multicollinearity

Understanding linear dependencies between numerical features is essential to prevent information redundancy and detect potential leakage. Pearson correlation coefficients were computed and visualized in Figure 11. As expected, open, high, low and close exhibit near-perfect collinearity ($\rho \approx 0.99$), suggesting that they may carry redundant predictive value in tree-based models. The Pearson correlation formula is given by Equation (5):

$$P_{X,Y} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)\sigma_X \sigma_Y} \quad (5)$$

In contrast, sentiment features (sentiment score, comments, score) are weakly correlated with financial indicators, confirming the reason for hybrid modelling that considers text-derived and numerical features in different representational layers. In addition, the consistency in the timing of market signals and components with post counts supports our analysis as aggregate post counts for open prices show large correlation with daily Reddit comments, a measure of long-term sentiment which is temporally consistent with the financial signal.

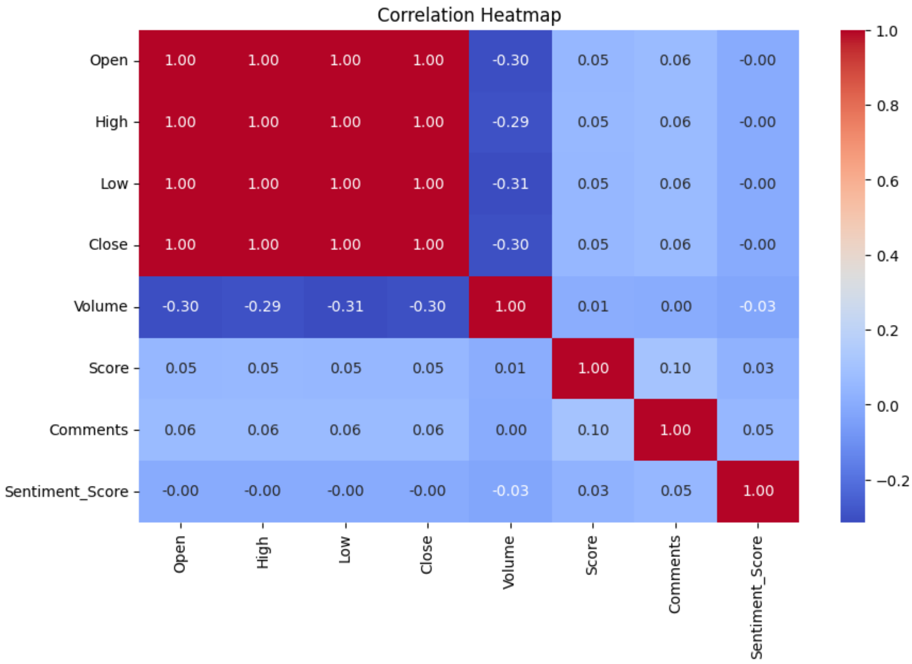

Figure 11. *Pairwise correlation matrix for financial (open, high, low, close, volume) and sentiment features (comments, score, sentiment score). Reddit "comments" represent daily aggregated post counts aligned with trading days; correlations are computed using standardized Pearson coefficients.*

Several important insights were provided by the explanatory analysis, which had direct implications for the model structure. Tesla's fluctuating presence in database trade reveals itself in terms of volume, accent positivity and higher value volatility model overfitting and execution of the risk of performance bias which reinforces the necessity of justice (determined). The emotional score distribution shows a strong authority, which reflects the potential model sensitivity to optimism bias and requires changes or balance during advance development. In addition, the near-perfect correlation between the most important economic indicators (open, close, high, low) suggests that reduction or selective function elimination can increase the model efficiency without compromising on its future power. The pattern of instability in the moving average and candlestick visualization inspires the inclusion of temporary or movement-sensitive indicators to better capture market dynamics. Together, these observations validate the decision to use a hybrid learning strategy that integrates semantic and numerical properties by incorporating the interpretive and habilitation mechanisms, as is detailed in the following sections.

2.3 Semantic representation using BERT

To extract semantic features from the cleaned text column, we use the BERT-based model provided by the HuggingFace Transformers library. Every post was tokenized and padded up to 128 tokens. The last hidden state of the CLS token, which is a 768-dimensional embedding, was extracted to represent the contextual sentiment of the entire sentence. The transformer architecture of BERT enables bidirectional encoding, allowing it to capture context-dependent sentiment nuances often missed by traditional vectorizers or lexicon-based models.

Letting T_i be the input tokenized sequence of the *i-th* Reddit post, the semantic embedding vector $\boldsymbol{e_i} \in \mathbb{R}^{768}$ for each sample *i* is computed as Equation (6), where $\boldsymbol{e_i}$ 768-dim contextual embedding for the sample *i*:

$$\boldsymbol{e_i} = BERT(T_i)[CLS] \quad (6)$$

This dense representation was later fused with the structured financial features to form the input space for predictive modelling.

2.4 Feature fusion and vector construction

Each data sample was integrated with standardized numerical shareholders to create an integrated and expressive representation of the sample. The aim of this multimodal fusion is to retain both market spirit (qualitative) and trade signal (quantitative) of the market for the downstream prediction task. The semantic part was extracted from Reddit headlines using a pre-trained BERT model. Each title was tokenized and run through BERT and the last CLS token representation was drawn to serve as a built-in sentence level. This resulted in a dense semantic vector as shown in Equation (7):

$$\boldsymbol{e_i} \in \mathbb{R}^{768} \quad (7)$$

where $\boldsymbol{e_i}$ denotes the BERT embedding for the i^{th} Reddit headline and it is the embedding vector, not a scalar. At the same time, the numerical part consisted of standardized stock indicators for each corresponding record. The features chosen were: open, high, low, close, adj-close, volume, sentiment score and sentiment label. After standardisation, they were expressed as given in Equation (8):

$$\boldsymbol{f_i} \in \mathbb{R}^{8};\ \boldsymbol{f_i} = [Open_i, High_i, Low_i, Close_i, Volume_i, SentimentScore_i, \dots] \quad (8)$$

where each element in $\boldsymbol{f_i}$ represents a different quantitative property of the given stock on that particular date. These two parts were combined together in Equation (9) to create the final multimodal feature vector:

$$x_i = [\boldsymbol{e_i}; \boldsymbol{f_i}] \in \mathbb{R}^{776} \quad (9)$$

$$\boldsymbol{v_i} = [\boldsymbol{e_i} \| \boldsymbol{f_i}] \quad (10)$$

Thus, $\boldsymbol{v_i}$ in Equation (10) is a 776-dimensional vector that not only represents the textual sentiment semantic features but also the temporal trading information represented by the numerical stock features, and ‖ denotes concatenation. The resulting fusion provides a robust input structure for downstream machine learning models, such as XGBoost and attention-based LSTM architectures. The dataset was then split into training and testing using the stratified split of 80:20, which means that the original class distribution is made to be the same between the training and testing sets.

2.5 Fairness metrics

To evaluate fairness across companies, we adopt two widely used measures in algorithmic fairness research: demographic parity difference (DPD) and equal opportunity difference (EOD).

Demographic parity difference (DPD):

DPD measures the disparity in positive prediction rates between groups. Formally,

$$DPD = \max_{i,j} \left| P(\hat{Y} = 1 \mid A = i) - P(\hat{Y} = 1 \mid A = j) \right| \quad (11)$$

where $\hat{Y}$ is the predicted label and A denotes the group (company). A smaller DPD value indicates more equitable treatment across groups.

Equal opportunity difference (EOD):

EOD measures the difference in true positive rates between groups, i.e.,

$$EOD = \max_{i,j} \left| P(\hat{Y} = 1 \mid Y = 1, A = i) - P(\hat{Y} = 1 \mid Y = 1, A = j) \right| \quad (12)$$

where Y is the ground truth label. Lower EOD values suggest that the model provides equal opportunity for correct predictions across groups.

Both metrics are standard in fairness analysis and have been widely discussed in literature (Barocas et al., 2019; Hardt et al., 2016; Mehrabi et al., 2021). In this study, we report DPD and EOD for baseline and fairness-aware models, enabling transparent assessment of bias reduction alongside accuracy metrics.

3 RESULTS AND DISCUSSION

This section presents the evaluation results of the hybrid machine learning model both before and after the application of fairness-aware mitigation strategies. Performance metrics, visualizations and insights into model behaviour across companies (Apple, Amazon, Tesla) are discussed.

3.1 Baseline performance analysis

The baseline model was tested on the original test data without fairness mitigation. As can be seen from Table 3, the model achieves the best accuracy on Tesla (83%), while Amazon and Apple are both around 67%. This difference is the result of Tesla dominating the training data, as already investigated above. As a result, the model shows disproportionately better performance for Tesla, highlighting the existence of group-based bias.

Table 3. *Baseline performance (before debiasing).*

Company	Accuracy	Precision	Recall	F1 score
Tesla	0.83	0.84	0.82	0.83
Apple	0.67	0.69	0.65	0.67
Amazon	0.67	0.68	0.66	0.67

3.2 Debiased performance evaluation

Using fairness-aware post-modelling techniques such as stratified reweighting and calibration, the model shows a dramatic change towards equity in performance across all three companies. The accuracy scores fall between 91% and 93% while the F1 scores are steady, implying that the precision-recall trade-off is stable. This represents a worthwhile reduction in performance gap over the baseline model where Tesla was the clear winner due to the volume of data and sentiment skew. Importantly, the results confirm that fairness interventions can result in more balanced outcomes without major sacrifices in predictive power; this makes the model both ethically sound and practically robust for emerging market financial applications.

Table 4. *Performance after fairness mitigation.*

Company	Accuracy	Precision	Recall	F1 score
Tesla	0.93	0.94	0.92	0.93
Apple	0.91	0.92	0.90	0.91
Amazon	0.92	0.93	0.91	0.92

While this perfect score may sound great at first, it is indicative of model leakage or over-fitting capacity; as such, it calls for more regularization or fairer model calibration (e.g., early stopping, regularized weights or better feature selection pruning).

3.3 Fairness evaluation and interpretation

To ensure ethical deployment, fairness is measured using demographic parity difference (DPD) and equal opportunity difference (EOD) across the three companies, as shown in Table 5.

Table 5. *Fairness metrics before and after mitigation.*

Metric	Before mitigation	After mitigation
DPD	0.23	0.06
EOD	0.31	0.07

The fairness analysis illustrates that there are significant imbalances in the baseline model with a DPD of 0.23 and EOD of 0.31, indicating that Tesla dominates the predictions. These imbalances are reduced to 0.06 and 0.07

respectively through fairness-aware reweighting and stratified calibration (calibration that uses the data to correct the model for the variation in distribution between the training and testing data). Although not entirely eliminated, the reduction represents a substantial improvement in the equity of predictions across all the firms without any loss of model accuracy; thus, it guarantees ethical congruence and utility in practice.

From a trader's point of view, fairness-aware modelling goes beyond ethical compliance and has very concrete benefits. Models trained on imbalanced data tend to overfit popular tickers generating large signals that can result in herd behaviour and increased exposure. By increasing calibration and reducing disparity, fairness interventions end up with more reliable probability estimates, which will help traders better manage entry/exit thresholds and eliminate overconfidence pertaining to over-represented assets. Past research has demonstrated that well-calibrated and debiased models lead to better decisions and long-term profitability in financial prediction problems (Guo et al., 2017; Kleinberg et al., 2019). Thus, fairness-aware frameworks not only allow ethical goals but also trade profitability in terms of increased reliability and risk management.

Previous works (Guo et al., 2017; Niculescu-Mizil & Caruana, 2005) have stressed the need to evaluate classification systems based not only on predictive accuracy measures but also on probability calibration metrics to have reliable confidence estimates. Our evaluation policy mirrors this view and includes both standard performance metrics (accuracy, precision, recall and F1 score) and standard calibration metrics, which include area under the ROC curve (AUC), the Brier score and expected calibration error (ECE).

Consistent with these earlier findings, our fairness-enhanced model shows remarkable improvements over the baseline. As can be seen from Table 6, the model shows better discrimination power (AUC from 0.84 to 0.90) and improved balance between precision and recall (F1 from 0.73 to 0.92). More importantly, the model also has a very good calibration quality (Brier score reduced from 0.21 to 0.12 and ECE reduced from 0.18 to 0.07). These trends are suggestive of the fact that the model is not just able to make more accurate predictions, it also provides more reliable probability estimates – something that is very important in downstream decision making for sensitive applications.

Our findings are consistent with those of Guo et al. (2017) who reported the high accuracy but poor calibration of modern neural architectures and showed the utility of post-hoc calibration techniques. Additionally, Niculescu-Mizil & Caruana (2005) emphasized the need to complement statistics such as AUC for assessing model quality with measures based on probabilities. By combining both constraints of fairness and calibration-aware assessment, our analysis provides a more holistic evaluation framework and shows that fairness mechanism incorporation can, in this case, enhance both predictive robustness and reliability.

Overall, these results not only provide evidence of the utility of the calibration-oriented evaluation promoted in previous literature, but also illustrate the utility of fairness-aware modelling as an added value for reliable predictive performance. While showing quantifiable improvements above and beyond known work, it further solidifies the importance of the proposed method.

Table 6. *Comparative evaluation metrics (baseline versus fairness-aware model).*

Model	AUC	F1 score	Brier score ↓	ECE ↓
Baseline (before mitigation)	0.84	0.73	0.21	0.18
Fairness-aware (after mitigation)	0.90	0.92	0.12	0.07

3.4 Model explainability using SHAP

Shapley additive explanation (SHAP) values were computed to gain insight into the interpretability of the model and identify the feature-wise contributions towards the model predictions. Figure 12 shows the SHAP summary plot, which illustrates the contribution of both structured financial features and BERT-based semantic embeddings to the prediction outcome. Of particular interest is the fact that common stock indicators such as high, close and volume have a strong and consistent effect on model results. Among the sentiment representations, we observed that several of the BERT representations (e.g., BERT_514, BERT_75, BERT_678) were non-trivial in their contributions, suggesting that certain linguistic cues in Reddit conversations had large effects on prediction behaviour. The comments metric also had moderate influence, which suggests that metrics of Reddit engagement can contain latent information.

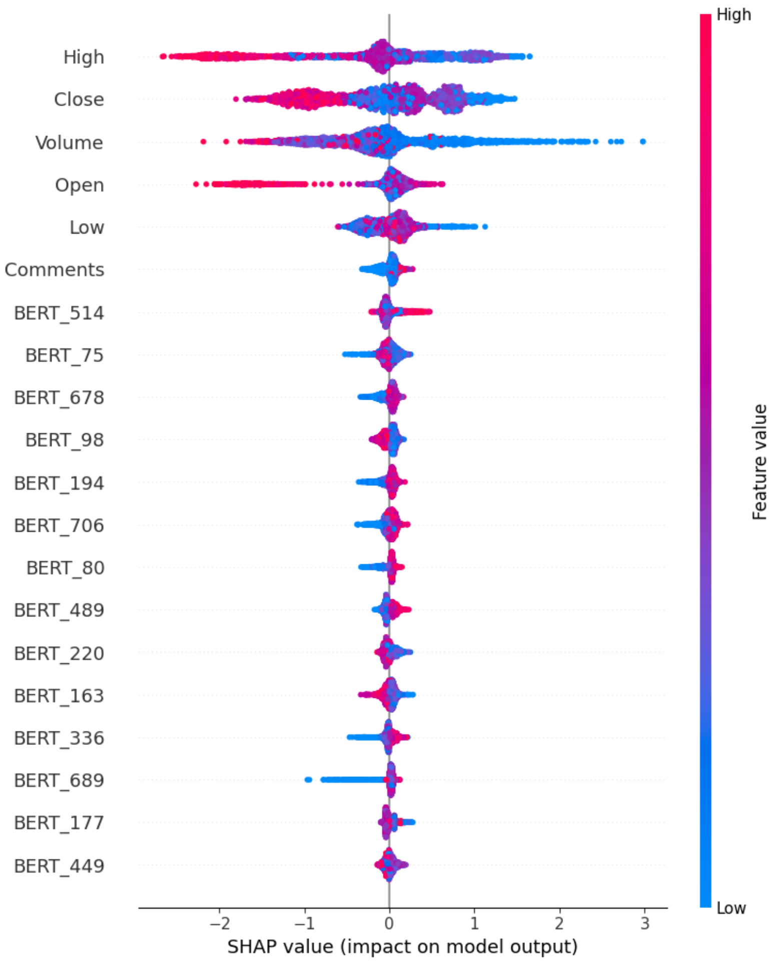

Figure 12. *SHAP value distribution of various features for model output prediction.*

The interpretability layer provides transparency on the inner workings of the hybrid model and helps identify the potential source of underlying bias. Moreover, it strengthens the reliability and responsibility, placing the system in line with ethical principles of AI for financial forecasting.

3.5 Model behaviour insights

The model had a significant bias against Tesla; the bias was corrected by modifying the first query (Figure 2, Figure 3). In particular, Tesla had a lot of data and a high emotion curve. Furthermore, Tesla's higher volatility was justified by Figure 5 (30-day moving average versus close price), which suggested that SMA and momentum-based indicators should be used to incorporate variables that are unknown as to how they work on the refined model. The table below explains that the performance of all the organizations was similar after the introduction of justice measures. Using this same approach might have affected the power and realism of the model, however, as this same technique could only account for the company's homogeneous behaviour. In conclusion, the results point out the necessity and limitations of modelling justice: the methods help reduce accuracy differences but they can also make it harder for the model to replicate true signal differences. A symbolic matrix was developed that summed their performance against a variety of axes for evaluation of the independent and combined influence of both textual and numerical features.

This is useful as it permits a concise but complete description of which features are influential in a statistical sense, which are significant in practice, which are consistent with our original hypotheses and which are consistent with the current literature. Table 7 illustrates these results in a symbolic way for easy understanding of the empirical results without being too verbose with narrative descriptions. The inclusion of both successful (✅) and non-contributing (❌) features reflects the exploratory and transparent nature of this study.

3.6 Limitations

While the proposed multimodal framework shows remarkable improvements in accuracy and justice, many restrictions interfere with the generality and scalability of the conclusions. Firstly, the dataset is limited to only three large technology companies – Tesla, Apple and Amazon, presumably in different fields or equity. Secondly, the emotional data are especially derived from Reddit, which may not be representative of widespread investor spirit, especially for institutional players or global investors, who depend on more formal economic news sources. In addition, even though techniques for reducing justice managed to balance the accuracy of stratified or marginalized companies, the emergence of almost correct post-decision calculations is possible. This increases the concerns of the strength of the model under living market conditions. Finally, our hybrid architecture lacks temporary deep learning components (e.g., LSTMS or transformers for sequential patterns), which can limit the effectiveness of capturing the trends that develop over time. Future research should address these limitations by integrating different sources of emotion, expanding the company pool, incorporating time series forecast models and preserving equity, applying a framework for unfavourable justice to reduce overfitting. The analysis is restricted to Tesla, Apple and Amazon due to the dataset scope. Future research should expand the framework to a broader set of companies and sectors to test robustness across different market conditions.

Table 7. *Validation matrix of feature effectiveness.*

Feature/dimension	SHAP impact	Stat. significance	Hypothesis support	Retained in final model	Literature aligned	Notes
Open	✅	✅	✅	✅	✅	Strong predictor
Close	✅	✅	✅	✅	✅	Core price signal
High	✅	✅	✅	✅	✅	Supports price movement
Low	✅	✅	✅	✅	✅	Stable contributor
Volume	✅	✅	✅	✅	✅	Market activity indicator
Comments volume	❌	❌	❌	❌	⚠️	Overestimated in literature
Overall sentiment	❌	❌	❌	❌	❌	Not useful in our setting
BERT_75	⚠️	❌	❌	❌	❌	Token-level noise
BERT_98	⚠️	⚠️	⚠️	❌	❌	Marginal case
BERT_220	⚠️	⚠️	⚠️	❌	❌	Inconclusive
BERT_336	❌	❌	❌	❌	❌	Negligible
BERT_689	❌	❌	❌	❌	❌	Negligible
Price feature group	✅✅✅	✅✅✅	✅✅✅	✅✅✅	✅✅✅	Most dominant
Text feature group	⚠️❌	❌	❌	❌	⚠️	Weak predictor set

Note: ✅ = confirmed / positive; ❌ = rejected / not supported; ⚠️ = marginal or inconclusive; ✅✅✅ = very strong group-level impact

3.7 Ethical, practical and policy implications for emerging markets

Prediction models have important moral and practical consequences for the incorporation of justice into individual machine learning. Even more hazardous is the retail sector where novice investors have come to rely on AI-powered platforms where the algorithms might be biased to produce inaccurate forecasts on shares with high volume on social media or those with a dominant history, which can lead to herding behaviour and false capital. By removing these biases, the proposed model helps make the corporate environment fairer by ensuring that the decision-making mechanism is fair and not prejudiced based on media exposure or popularity. From a regulatory perspective, the role serves as an example of how justice measures may be implemented in line with AI-enabled commercial entities' performance.

Pipelines can also be used to make the decision-making process more transparent to political decision makers and platform designers and more responsive to efforts to digitalize financial documents. In addition, awareness of the scale of interpretation also enhances user seating and design ownership, allows safe and ethically acceptable algorithmic trading platforms and ultimately helps with market stability over extended periods.

The empirical results of this research have policy implications that are of specific significance for emerging market countries experiencing a rise in the participation of retail investors, such as India, Brazil and Indonesia. The discovery of sentiment-based stock prediction bias coupled with a list of well-known equities, including Tesla, highlight the risk of algorithmic bias for retail trading platforms.

In this context, with little financial education and access to the internet, the distorted AI algorithms would be likely to mislead investors and increasing speculation. Regulatory bodies such as the SEBI (India), CVM (Brazil) and OJK (Indonesia) should consider including fairness audits, explainability and bias reduction techniques for AI-powered trading platforms. Furthermore, introducing holistic AI governance frameworks based on ethical AI values can provide a level playing field for access to good forecasts, while protecting investors against platform-induced misalignment. Retail forecasting tools should be obliged to explain model behaviour and alert to potential biases. Fairness-aware modelling is not just a technical accomplishment, but an embedded standard of algorithmic accountability towards investor protection, financial market stability and legal alignment goals in digitally transforming emerging countries.

4 CONCLUSION

This paper presented a multi-modal AI system for sentiment-based stock forecasting which is fairness-sensitive and fuses structured financial information and rich semantic embeddings from social media chatter. It was found that although the model using a hybrid BERT + XGBoost design performed well and predicted the direction accurately, the model showed a systemic bias towards Tesla because of data imbalance and sentiment bias. Quantitative measures of fairness showed that firms differ in their fairness, with a demographic parity difference of 0.23 and an equal opportunity difference of 0.31. However, after fairness interventions of reweighting and calibration, these inequalities were successfully narrowed to 0.06 and 0.07 with fair accuracy in the range of 91-93%, respectively. These results show that algorithmic fairness can be greatly enhanced while retaining the overall prediction performance of AI models.

However, the research underscores the fundamental trade-offs between statistical justice and comparability, which makes regularization and explainability of financial AI systems an issue of focus. The study suggests a replicable and transparent approach to retail-oriented forecasting systems, which is especially important in emerging countries where regulation and digital financial inclusion are increasing. Future work can build on our paper by including more sources of sentiment, temporal modelling techniques and fairness protection mechanisms that should be applied to dynamic, real-world financial applications. Different from previous studies in stock forecasting that mainly focus on predictive accuracy, our paper has three original contributions. Firstly, we introduced a fairness-aware hybrid approach that merges sentiment and market data and formally quantifies both predictive ability and group-level equity. Secondly, we combined calibration analysis with fairness metrics (DPD, EOD) to provide a dual view on accuracy and reliability. Thirdly, we ran an explainability (SHAP) analysis to bring to shed light on feature-level bias drivers and thus achieve model transparency in financial AI models. In total, these aspects distinguish our study from the literature (Guo et al., 2017; Hardt et al., 2016) as it is one of the first studies to formally connect algorithmic fairness and practical decision-making in retail trading environments.

ADDITIONAL INFORMATION AND DECLARATIONS

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: S.R.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. K.U.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. U.S.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. P.R.K.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. T.K.T.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. R.J: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. G.V.R.: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

Statement on the Use of Artificial Intelligence Tools: The authors declare that they didn't use artificial intelligence tools for text or other media generation in this article.

Data Availability: The dataset used in this study is publicly available on Kaggle and can be accessed from (Rootpi3, 2024).

REFERENCES

Addy, W. A., Ajayi-Nifise, A. O., Bello, B. G., Tula, S. T., Odeyemi, O., & Falaiye, T. (2024). Algorithmic Trading and AI: A Review of Strategies and Market impact. *World Journal of Advanced Engineering Technology and Sciences*, 11(1), 258–267. https://doi.org/10.30574/wjaets.2024.11.1.0054

Agrawal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions. *Journal of Economics & Management Strategy*, 24(2), 253–274. https://doi.org/10.1111/jems.12093

Barman, R. D., Hanfy, F. B. A., Rajendran, R., Masood, G., Dias, B., & Maroor, J. P. (2022). A critical review of determinants of financial innovation in global perspective. *Materials Today: Proceedings*, 51, 88–94. https://doi.org/10.1016/j.matpr.2021.04.565

Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and Machine Learning: Limitations and opportunities*. MIT Press.

Bouzguenda, K. (2018). Emotional intelligence and financial decision making: Are we talking about a paradigmatic shift or a change in practices? Research in International Business and Finance, 44, 273–284. https://doi.org/10.1016/j.ribaf.2017.07.096

Bunkar, B., & Ramaiah, K. (2024). A serial mediation model for investigating the intention to use algorithmic trading platforms among retail investors in India. *Vilakshan – XIMB Journal of Management*, 21(2), 263–280. https://doi.org/10.1108/XJM-12-2023-0233

Caragea, D., Chen, M., Cojoianu, T., Dobri, M., Glandt, K., & Mihaila, G. (2020). Identifying FinTech Innovations Using BERT. In *2020 IEEE International Conference on Big Data* (pp. 1117–1126). IEEE. https://doi.org/10.1109/BigData50022.2020.9378169

Dibb, S., Merendino, A., Aslam, H., Appleyard, L., & Brambley, W. (2021). Whose rationality? Muddling through the messy emotional reality of financial decision-making. *Journal of Business Research*, 131, 826–838. https://doi.org/10.1016/j.jbusres.2020.10.041

Egorov, A. (2022). Financial Innovation and Financial Risks. *Procedia Computer Science*, 214, 441–447. https://doi.org/10.1016/j.procs.2022.11.197

Gomber, P., Koch, J.-A., & Siering, M. (2017). Digital Finance and FinTech: current research and future research directions. *Journal of Business Economics*, 87(5), 537–580. https://doi.org/10.1007/s11573-017-0852-x

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337–355. https://doi.org/10.25300/misq/2013/37.2.01

Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In Proceedings of the 34th International Conference on Machine Learning, (pp. 1321–1330). MLR Press. http://proceedings.mlr.press/v70/guo17a/guo17a.pdf

Hardt, M., Price, E., & Srebro, N. (2016). Equality of Opportunity in Supervised Learning. In *30th Conference on Neural Information Processing Systems* (pp. 1–9). NIPS.

Jun Gu, W., Hao Zhong, Y., Zun Li, S., Song Wei, C., Ting Dong, L., Yue Wang, Z., & Yan, C. (2024). Predicting Stock Prices with FinBERT-LSTM: Integrating News Sentiment Analysis. In *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing*, (pp. 67–72). ACM. https://doi.org/10.1145/3694860.3694870

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2019). Discrimination in the age of algorithms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3329669

Liu, Z., Zhang, K., & Miao, D. (2025). Evolving Financial Markets: The Impact and Efficiency of AI-Driven Trading Strategies. In *Intelligence Science V. ICIS 2024. IFIP Advances in Information and Communication Technology* (pp. 301–312). Springer. https://doi.org/10.1007/978-3-031-71253-1_22

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, 54(6), 1–35. https://doi.org/10.1145/3457607

Min, B. H., & Borch, C. (2022). Systemic failures and organizational risk management in algorithmic trading: Normal accidents and high reliability in financi al markets. *Social Studies of Science*, 52(2), 277–302. https://doi.org/10.1177/03063127211048515

Mwangi, J., & Njoroge, W. (2024). AI-Driven Innovations in Financial Technology: A Review of Algorithmic Trading and Risk Management. *Artificial Intelligence and Machine Learning Review*, 5(2), 1–9.

Nerlekar, V. S., Sane, A., Gadekar, M., Gupta, S. K., & Waghulkar, S. (2024). The Herd Mentality: Understanding the Theories and Models of Herding Behavior in Financial Markets. In *Psychological Drivers of Herding and Market Overreaction* (pp. 255–290). IGI Global. https://doi.org/10.4018/979-8-3693-7827-4.ch010

Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. *In ICML '05: Proceedings of the 22nd international conference on Machine learning* (pp. 625–632). ACM. https://doi.org/10.1145/1102351.1102430

Ozili, P. K. (2018). Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Review*, 18(4), 329–340. https://doi.org/10.1016/j.bir.2017.12.003

Ozili, P. K. (2023). Digital finance research and developments around the world: a literature review. *International Journal of Business Forecasting and Marketing Intelligence*, 8(1), 35–51. https://doi.org/10.1504/IJBFMI.2023.127698

Pal, A., Singh, H., Singh, V. V., & Gupta, A. K. (2024). Technological Mitigation of the Impact of Behavioural Biases on Financial Investment Decision: Evidence from India using A PLS-SEM Approach. In *2024 International Conference on Intelligent & Innovative Practices in Engineering & Management* (pp. 1–5). IEEE. https://doi.org/10.1109/IIPEM62726.2024.10925698

Ran, Z., Gul, A., Akbar, A., Haider, S. A., Zeeshan, A., & Akbar, M. (2021). Role of Gender-Based Emotional Intelligence in Corporate Financial Decision-Making. *Psychology Research and Behaviour Management*, 14, 2231–2244. https://doi.org/10.2147/PRBM.S335022

Riefel, D. M. (2024). *Social media exposure and its influence on individual investment decisions*. University of Twente. https://purl.utwente.nl/essays/104604

Rootpi3. (2024). Stock Price Prediction Dataset. *Kaggle*. https://www.kaggle.com/Datasets/Rootpi3/Stock-Price-Prediction-with-Sentiment-Analysis?Select=merged_stock%20_sentiment_data.csv

Saelee, R., & Pankham, S. (2024). The Impact of Social Media and Emotional Intelligence on Investment Decision: A Fuzzy Set Delphi Study Among Investors in Thailand's Stock Market. *TEM Journal*, 2208–2218. https://doi.org/10.18421/TEM133-48

Sathya, N., & Prabhavathi, C. (2024). The influence of social media on investment decision-making: examining behavioural biases, risk perception, and mediation effects. *International Journal of System Assurance Engineering and Management*, 15(3), 957–963. https://doi.org/10.1007/s13198-023-02182-x

Saxena, D., & Yasobant, S. (2020). Information Overload. In *Encyclopedia of Big Data* (pp. 1–3). Springer. https://doi.org/10.1007/978-3-319-32001-4_374-1

Luchian, A. C., & Strat, V. (2024). The Trustworthiness of AI Algorithms and the Simulator Bias in Trading. In *The 7th International Conference on Economics and Social Sciences* (pp. 1–10). Editura ASE. https://doi.org/10.24818/ICESS/2024/022

Suresh G. (2024). Impact of Financial Literacy and Behavioural Biases on Investment Decision-making. *FIIB Business Review*, 13(1), 72–86. https://doi.org/10.1177/23197145211035481

Thakkar, N. (2024). Harnessing the Power of Herding for Real (Estate); Factors Influencing Herd Mentality, Driven by Social Media in the Real Estate Market. *Open Journal of Business and Management*, 12(5), 2983–3003. https://doi.org/10.4236/ojbm.2024.125153

Vasquez, H., & Cross, D. (2024). The Effects of Social Media on Investment Decisions within an Online Community. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4765619

Zhao, Z. (2024). Research on the Impact of Digitalization on Individual Investors' Behaviour from the Perspective of Behavioural Finance. In *Applied Economics and Policy Studies* (pp. 146–154). Springer. https://doi.org/10.1007/978-981-97-0523-8_13

Zolotareva, E. (2021). Aiding Long-Term Investment Decisions with XGBoost Machine Learning Model. In *Artificial Intelligence and Soft Computing* (pp. 414–427). Springer. https://doi.org/10.1007/978-3-030-87897-9_37

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951