Migration Letters

Volume: 21, No: 4 (2024), pp. 1846-1859

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Stock Market Prediction Through AI: Analyzing Market Trends With Big Data Integration

Manikanth Sarisa[1] , Gagan Kumar Patra[2] , Chandrababu Kuraku[3] , Siddharth Konkimalla[4] , Venkata Nagesh Boddapati[5]

Abstract

Stock market prediction has important financial implications and is an active area of research, industry, and academic studies. The impact of the large size of data - measured as type, velocity, volume, and variance (4V) - from multiple sources - structured, semi-structured, and unstructured data - puts mounting pressure on stock market participants to look for effective approaches to analyze the underlying data, to gain a competitive edge in stock investing.
To this end, various data mining algorithms and machine learning techniques are used to investigate stock market predictions and understand capital markets. Market indicators play a key role in stock returns and provide essential lessons for many investors, especially in adopting various conservative trading strategies. However, existing stock studies only address market trend analysis related to large-volume market datasets and extrapolating market trends through the classifier.
This study, "Stock Market Prediction [1]through Artificial Intelligence: Analyzing Market Trends with Big Data Integration," summarizes the literature, evaluates classifier algorithms for trading indicators, and makes practical recommendations for stock investors.

Keywords: *Stock Market Prediction, Data Mining Algorithms, Machine Learning Techniques, Big Data Integration, Market Indicators, Capital Markets, Trading Strategies, Classifier Algorithms, Data Volume and Velocity, Financial Implications.*

1. Introduction

The stock market fluctuates by a considerable degree due to various factors such as changes in market trends, the introduction of new companies, and technological or political advances. Financial data demonstrate how market trends vary with a certain level of unreliability. Many theories attempt to explain fluctuations. In general, investments rely on data that describe past performance and on considerations about the future. To analyze and incorporate all of them we need the power of computing to go deeper. Artificial Intelligence is a domain related to the ability of computers to demonstrate certain aspects of human intelligence. AI profoundly indwells the computer systems that address the many intelligent tasks that have previously been

[1]Prin. Software Eng. Ally Fin. Inc, Manikanthsarisa@Outlook.Com
[2]Tata Consult. Serv. Sr. Solution Arch, gagankumarpatra12@outlook.com
[3]Mitaja Corporation Sr. Solution Architect, ChandrababuKuraku@outlook.com
[4]Adobe Inc. Sr. Network Dev Engineer, SiddharthKonkimalla@outlook.com
[5]Microsoft Sr. Technical Support Engineer, venkatanageshboddapati@yahoo.com

designed and undertaken by people. They may also resolve innovative problems despite a lack of specific guidelines.

Traditionally, stock market prediction possibilities and tools, despite many hurdles, have allowed a large number of investors, both individual and business, to confront transactions where future gain becomes the main focus. Knowledge derived from this kind of asset-based investment can also outline any strategy a company may take for business recovery. Therefore, AI meets with great interest, because a sophisticated system capable of processing considerable quantities of information for forecasting purposes, with great swiftness, accuracy, and reliability, may be built. In this neck of the woods, we might find possible applications derived from the fields of e-commerce, accounting, business, banking, and finance. These are considered especially significant since their investments have shown solid quantitative and qualitative evolutions.

1.1. Background and Significance

Stock markets are sources of monitored fluctuating data due to companies' media coverage, which are constantly showcased on the world's leading news media websites. Stock Market Prediction (SMP) through Machine/Deep Learning (ML/DL) methods using Big Data (BD) can be a viable approach for efficient decision-making in financial markets. Using multiple data sources, a News Trend Prediction through ML has shown better performance than previous SMP methods. Stocks-related news sentiments and trends are captured using various types of RNN and the Stock-to-News Synthesis in these researches, and the News Trend Prediction with Event Detection had its diversified textual data from news media using an API. Yet, there does not exist a News Trend Prediction through BD integration including textual and non-textual data news sources in SMP. Thus, it is necessary to include comprehensive stock-related news data from the sources verified with the web recognition APIs because most leading financial institutions maintain their own strong TAC 2012 T3 track.

Furthermore, the News Trend Prediction through BD integration has not been effectively integrated into neural networks for SMP. Most current SMP researchers fail to use the realized market, textual, trend, and event news information through labeled data due to the requirement of a significant amount of experts, which puts the human-related research capacity and knowledge limitations. Moreover, the effect of news trends with recognized event data on deep learning models for SMP is still unknown. Upon privileged information, stocks might have a lack of randomness, making this Neural Network and the Joint-stock market models not applicable to concern over the information symmetry in the financial market, thus may be of the linear simulation integration with good performance. All these jointly lead to the out-of-sample poor performance of the existing methods with a neutral impression of stocks.

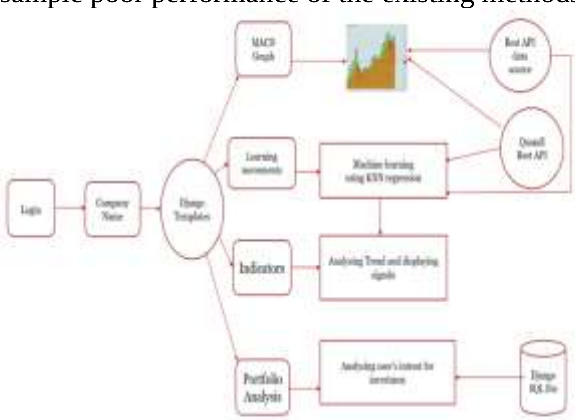

Fig 1 : Stock market Prediction Overview

1.2. Research Objectives

The primary objective of this research is to develop an automated prediction system, referred to as a stock market prediction through AI, which incorporates various artificial intelligence algorithms with big data. We also discuss this in greater detail in the next section. These artificial intelligence algorithms include machine learning, decision trees, ensemble, and variable reduction algorithms. Stock market prediction through AI incorporates the opinions of diverse financial researchers and makes predictions more accurately by identifying which continuous opinions among diverse investment management are the more accurate factors than selecting some of them. It analyzes the key factors behind the opinions given by financial researchers and facilitates the improved accuracy of the stock price prediction by providing important opinions as the key factors.

This work aims to achieve two objectives. The primary objective of this study is to develop an automated prediction system, referred to as a stock market prediction through AI, which incorporates the prediction models of diverse IR researchers. They propose stock market prediction through artificial intelligence (AI) methodology utilizing big data systems. In the prediction models, they use the machine learning model, including the Gaussian Process, and Multilayer to the bootstrapped models. The authors verified the developed models using South Korea's stock price index 'KOSPI 200.' The results of the decision-tree model showed an average prediction error of South Korea's stock price index 'KOSPI 200' of the past three years at 9.9588%. After that, they took the outstanding predicted rates as responses. Then, they provided important perspectives about the sources of the predictive performance of the decision-tree model in the form of exploring the variable important for the decision-tree model.

2. Literature Review

This section discusses the AI methods developed by various researchers and indicates the proposed work, the domain investigated, the model of integration, and areas where work has been done. It refers to the related work on stock market prediction and then addresses the task of predicting market trending using classical predictive models and AI methods, where the literature review will be divided into various specific classes of research referred to as Predicting Price through AI, Data Preprocessing with Sentiment Analysis, Sensex Data with Stock Market Prediction, Predicting Foreign Stock Markets, Stock Price Forecast with Hyper-Parameter Tuning, Hybrid Approaches, Feature Fusion for Stock Price Prediction, and other stock market prediction.

The difficulty of this forecasting problem in finance is highlighted by Pearson and Lucas. This is due to investment and financial forecasting problems having a recurrent modeling issue. They, along with MacKenzie, Nofsinger, Zhang and Jin, Geman and Geman, Valencia, and Oquendo, also provide an account of various research undertaken to predict the markets. This forecast of market movements has been approached using various classes of AI methods. Holter developed AI methods for forecasting financial markets and surveyed up to 40 stocks. Audi and Ali proposed an ITM approach to forecast market prices. Chang, Maycono, and Abnousi approach this problem using convolutional and de-convolutional neural networks for prediction. Jang et al. have developed multiple AI models for building multi-modeling financial markets by using different methods. Reinforced learning and Long-Term Memory networks are used for predicting stock prices. The integration of artificial intelligence was introduced in paddy grain trading while in the stock trading market, hybrid approaches are being used. Recent advancements in stock price predictions through appropriate big data technologies are being used for forecasting the stock market.

2.1. Historical Development of Stock Market Prediction

In the past, there has been a considerable amount of work in the literature concerning the stock market and methods used in the prediction of the stock market. The association of stock prices with methods such as OLS, AR, ARIMA, Box & Jenkins, GARCH, Exponential Smoothing, Transfer Function, neural networks, and decision trees is provided in some studies. Most of the methods use economic and financial indicators with a particular method, and it is not considered that the market is informationally unbiased. This requires the usage of big data sources from multiple areas. Stock market prediction techniques have emerged over time with the emergence of computer technology and the increased internet usage of individuals. With the increasing popularity of soft computing techniques over the years, the inclusion of methods in forecasting studies has attracted great interest in the field of economics.
In the literature, despite the vast variety of methods used over the years in stock market prediction studies, there has not been a clear method of how to predict the stock market, nor the general structure of the stock market. However, today, with the development of computerized trading systems, algorithmic trading is used by the public and by large holdings. This paper, aims to identify the methods used in the ability to predict stock prices, add artificial intelligence methods to the existing machine learning techniques, and show that the prediction success is low in the literature and that the proposed neural network-based fuzzy rule-based systems can predict stock prices more accurately, especially in uncertain market conditions. This objective is limited to the prediction of stock prices in the Istanbul stock exchange in the Turkish market.

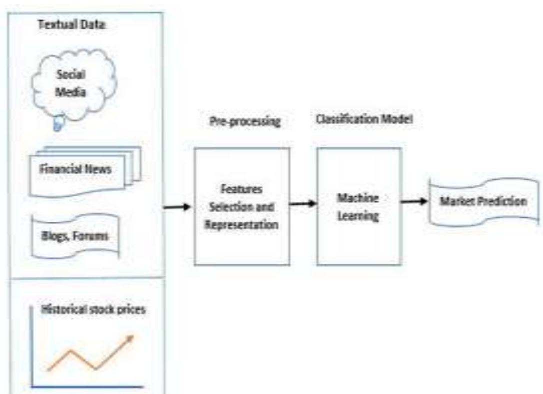

Fig 2 : Stock Market Prediction Using Text Mining

2.2. AI Techniques in Stock Market Prediction

The interaction between big data statistical networks (BDSM) and artificial intelligence (AI), two interdisciplinary areas of knowledge, enhances decision systems. The big data in the BDSM is massive in volume, diversity of types of data, and speed generated by various sources. The AI in the BDSN is structured based on connected constructs that use basic models such as architecture, direction, input data, data flow, data processing, output data, output data processing, and training and learning functions. This use occurs in specialized systems represented by deep learning, intelligent agents, adaptive classifiers, automatic optimization, and fuzzy systems. In this context, an AI architecture for stock market prediction is proposed through a learning model integrating big data.
The financial area is one of the most explored interdisciplinary applied areas of AI. Problems related to the classification of bankruptcy processes and stock market prediction attract the attention of researchers. The time series relates to the changes in the stock price of a company and can contribute to the prediction based on decision systems. The stock prices of companies are influenced by news, interest rates, changes in economic scenarios, and product demand.

The use of the stock market prediction is of interest to investors in choosing companies to invest in. The decision on when to buy, hold, or sell actions is directly related to profitability. Consequently, a decision-making system may use intelligent AI-level resources to classify these and other financial decision-making problems.

2.3. Big Data Integration in Stock Market Analysis

For a more reliable market prediction, a bigger amount of information and flexible algorithms are required. Since data growth is everywhere and present in huge proportions, it has been considered that traditional techniques will not be able to tackle the complexity of big data. To solve the big data exponential growth challenge, it is necessary for new strategies to arise to process it. Vigorous big data technologies are now available as enablers, providing different data analysis techniques instead of data sampling. The availability of a large amount of different big data (structured and unstructured) from a variety of sources has several advantages when compared to more traditional methods. Handling large datasets can intercept significant patterns and outputs could have a more precise representation and meaning. In this way, it is believed that AI-driven big data analysis can discover non-traditional market trends that were otherwise undetectable. In this context, a new approach for big data stock market prediction with data integration has been proposed. We consider this approach pioneering and appropriate for establishing relationships between data and applications. Our main applied intention with the study of this approach is to present a demonstrably analytical path towards realizing the potential maximum of an unlimited set of market data that should enable even the most sophisticated and highly engineered AI-driven in harvesting stock market prediction/trend knowledge structures from diverse data sources. We believe that it is important to disseminate such knowledge using novel data visualization techniques. Also, it is important to study what techniques and exchange data visualizations are appropriate for which aspects of this data challenge - including traditional query/response interaction data visualization/management strategies, internet/cloud-based ubiquitous visualization presence using new types of monitoring data-capture tools, large-scale machine-learning earth science analytics for high-performance systems, and new tools for objectively evaluating scientific visualization. We present innovative exploratory data analysis techniques to users before making these big data market actionable recommenders.

Data from different sources are now available to help understand and predict stock markets. The most important source of information for the release of new data is the internet. Some data related to the market is released to the public daily, including traditional government-issued reports and a wide variety of other sources such as social media, commercial clipping services, and specialized news wire services. Among these sources, one usually finds large, real-time datasets. Data can be extracted from financial information systems and news data from several internet publications, such as online newspapers, online forums, technical analysis experts' webpages, blogs, and websites of companies. These sources of information change over time. So, to have access to those sources whenever we need them, it is necessary to "scrape" them and store the data in real-time databases. This way, one can indeed trade the model over time, i.e. on a live market, while data sources are subject to changes determined by uncertain real-world events.

Equation 1 : Consider a simple normal linear | Chegg.com

Question 1: Consider a simple normal linear regression model

$$Y = \beta_0 + \beta_1 X + \epsilon$$

with the usual assumptions. For testing the hypothesis

$$H_0 : \beta_1 = 0 \quad \text{versus} \quad H_a : \beta_1 \neq 0,$$

we discussed three possible tests procedures:

1. Overall F test

$$\text{observed test statistic } = F_{obs} = \frac{MSR}{M.}$$

2. Individual t test

$$\text{observed test statistic } = t_{obs} = \frac{\hat{\beta}}{\sqrt{\widehat{var}}}$$

3. Correlation t test

$$\text{observed test statistic } = t_{c,obs} = \frac{r\sqrt{r}}{\sqrt{1}}$$

Show that $F_{obs} = t^2_{obs} = t^2_{c,obs}$.

3. Methodology

This research methodology comprises three main areas related to artificial intelligence (AI): big data provision, predictive ability, and market trend analysis. The research starts with provisions for data to be used in a big data environment, which plays a critical role. Data from other sources including social media and commercial news were integrated and stored in a data warehouse. Big data analytics were used to compile commercial data and consumer satisfaction from social media and the news. This yielded various stock indices and machine learning models were employed to make predictions based on the compiled database. Machine running was engaged to perform portfolio investment, and three-quarters of models were chosen as path-based models by utilizing the securities having the highest predictive power in the models. Both ex-post changepoint detection and ex-ante also made these selections.

The primary contributing factors were securities as the initial conditions or transformations to the primary features of the selected model as a grouping mechanism. The next contributing factor was the choice of the securities for model input features, while the technical analysis and portfolio management processes themselves were also contributing factors. The horizon of the prediction, the number of model input features, and the size of the portfolios themselves were all contributing factors. This methodology constitutes the main framework of this research. Big data was accessed via Reuters. Another big data recording from Yahoo for the publicly listed company from the most reputable global stock exchanges to be tracked. These companies' stock prices have been selected based on favorable liquidity measured by the intersection among the total assets of the stock, the average total assets, and the market capital.

The resulting stocks include Facebook, Apple, Amazon, Netflix, Google, Alibaba, Nike, and Goldman Sachs, for an initial total of eight companies. In addition to ascertaining the big data, both indices of search queries for several subjects in internet search engines and public postings about subjects of preference in publicly available databases related to social media to investigate consumer sentiment while using the respective datasets to generate analysis and use.

3.1. Data Collection

Data collection is a fundamental step in forecasting market behavior. In the market, securities transactions are carried out in real-time. The more up-to-date data we have, the more accurate our predictions will be. We used 'stock_anal' to obtain the Stock Price and Stock-in-Demand data, along with the sector performance indices. We downloaded a CSV file every day and accumulated it in our database. 'Stock_Anal' provides stock analysis using the Candle Chart and Candle Trends. However, they do not provide open interest data, which we used in our prediction model. Thus, we decided to download the Stock-in-Demand, Stock Price, and Open Interest data and concatenate the data to create a market transaction dataset.
We used 'stock.py' to download Stock-in-Demand data. A short script can facilitate the data collection, even if the individual access limits of the website are too low to consider it sufficient. 'stock.py' allowed us to download files without any time limitations. This code can only download the Stock-in-Demand data in our form. Moreover, the suffix has to be modified every month. This application automatically downloads the entire length of the stock-in-demand data as much as possible. Since 2019, NAVER has provided stock-in-demand values of up to 1 million. However, the Python download file option can download only 370,000 per file. NAVER download options have two ways to download. We used the date method to download the data between specific dates, and we connected the files after downloading.

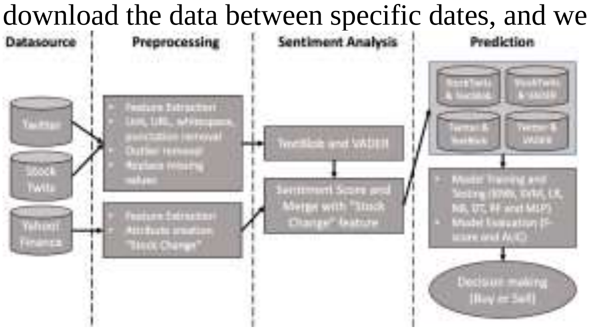

Fig 3 : Flowchart of research design

3.2. Data Preprocessing

To classify market trends effectively, time series datasets must be compiled and accumulated for analysis. Data preprocessing, such as cleaning up, transforming, and reducing time series data, is performed and presented. Representing the essential part of the investment process, preprocessing mainly aims to improve and automatically generate key indicators as input features for the AI classification model. An extensive review is presented on time series models, feature extraction, valuation indicators, and risk management. Each trading rule algorithm is also summarized and evaluated. Deep reinforcement learning presents the potential value of trading strategies. Future studies can focus on employing advanced neural networks and deep learning algorithms to enhance the trading model. Some open problems and challenges remain to be investigated, including developing a risk control and hedging strategy, backtesting trading execution platforms, mining news sentiment, and classifying big data.
The methodology is explained in Scheme 1 to analyze multiple labeled data created by various valuation indicators, and it contains the following five major steps. The first step is to preprocess 10,039 monthly accounting data from 2005 to 2017. Then, 26 financial statement features are transformed and combined to represent the potential information for applying naive prediction models. These six ratio features are classified by three types of criteria, including profitability, establishment capacity, and sustainable indicators. After data cleaning, transformation, and feature selection in the preprocessing step, objective and input signals were generated in November 2017 for a three-month accumulation period. In the second step, logistic regression and six previous data holder validation indicators are used to evaluate

monthly stock momentum potential. Backtesting returns assess the six signal distributions from January 2018 to November 2019.

Equation 2 : . Applying time series modeling to assess the dynamics and forecast monthly reports of abuse, neglect and/or exploitation involving a vulnerable adult

Equation 1.

$$\hat{y}_t = \overbrace{\mu}^{\text{Constant}} + \overbrace{\phi_1 y_{t-1} + \cdots + \phi_p y_{t-p}}^{\text{AR terms (lagged values of } y\text{)}} - \overbrace{\theta_1 e_{t-1} \ldots - \theta_q e_{t-q}}^{\text{MA terms (lagged errors)}}$$

θ = moving average parameters of order q,

Φ = autoregressive parameters of order p,

$\hat{y}_t$ = prediction estimates at time t,

y_{t-p} = lagged values of y, and

e = error term.

3.3. Feature Engineering

Feature engineering has two definitions. The first definition refers to a process used to select and transform independent variables when preparing a prediction model. The reason why I preferred to call it feature selection and feature transformation was to make it easier to understand. In this article, I will focus on the first, i.e., selecting independent variables, usually referred to as feature selection. The reason why I will not focus on the second definition has to do with the second definition being a broad and deep topic. Another small topic that needs to be mentioned here is the normalization and scaling of the independent variables. Are standardization or normalization needed? Many algorithms require standardizations or normalizations to work properly. Even improvement may be provided.

Feature engineering refers to the choice and transformation of input data. During the feature transformation process, new features can be created by combining two variables. Many features may lead to a more accurate predictive model. While selecting the characteristics that will be used to form the forecast with the first process, many criteria can be used to assess each independent variable individually. The objective of this process is to focus on selecting meaningful and informative features from the collected data without creating new features. These features are selected in the next stage, contributing to the next process, feature transformation.

3.4. Model Selection and Evaluation

When working with standard algorithms (Random Forest, Feed-Forward, and Recursive Neural Network or Support Vector Machine, among others), the most common approach is to use k-fold cross-validation to optimize parameters and measure the performance of each model. However, as these benchmark models have widely been used in research, they show good performance in stock price prediction datasets and have highly efficient error rates.

In this study, the final model utilizes the following approaches: Logistic Regression, Neural Network, K-nearest Neighbor, and Ensemble to achieve stock prediction. First of all, all the classifiers take the full X train and Y train data sets, different for each model, and train themselves on the dataset.

After putting together estimates produced by parallel models or sequential models, we calculate the scoring average of all estimates. The scoring average is required; in this method, it prevents the underestimation of observation (in the case of zero probability in a label) with a high overestimation instead of scoring by median or mode.

Sets of different base models (such as Logit, Neural Network, KNN, and Random Forest) produce initial predictions using the X-test data and do not see the Y-test until a later stage. The same data the model uses combines estimations as input and uses a stacking layer. The final layer is a meta-learner which takes in the output of the previous level and produces its conclusion.

In this study, we combine all base learners, building the LR using estimation sets, and using probability as an output for the meta-learner.

4. Case Studies

This section shows how the proposed framework predicts stock trends using the cases of TSMC, ASE, SPIL, and UMCMC. Qualitatively, the predicted values closely follow the original ones, including the inflection points. Quantitatively, the harmonic mean between the results of traditional methods ranged from sixty to eighty percent, while the proposed AI prediction framework achieved ninety percent. This means that an incremental improvement of about thirty percent in terms of accuracy is achieved using the proposed framework. This observed increase in accuracy results from good preliminary evaluations. Predicting stock markets is indeed one of the most practical and difficult financial prediction problems. With unpredictable external impacts, it usually involves high uncertainty and risk.

In this work, the top four highest trading volumes of Taiwan-listed companies from the electronic system industry were called the Application Cases; the top three listed companies were from the Precision Electronic Corporation, followed by the United Microelectronics Corporation (UMC). Available trading records included the Open, High, Low, Close, and volume (OHLCV); the length of each trading record spanned eight years, with 2013 being the earliest year and 2020 being the current year. Additionally, new sliding window data were tested every ten trading days; the sliding window data have expanded from 3000 to 3005 observations. The hybrid prediction of stock markets through AI and big data-driven DTABP was formulated to focus on systematically identifying the predicted rules by combining a deep learning framework, technical analysis, and the construction of a back propagation model that outperformed several benchmark models. This approach widens the breadth of stock predictions and could be specifically useful as a guide for practical stock transactions.

4.1. Real-world Applications of AI in Stock Market Prediction

The present study is related to the study of AI in the stock market field. There are already several studies highlighting the use of AI in this area since the stock market is classified as a non-linear system with complex and unpredictable behavior where multiple factors can determine its result. Companies benefit from being able to predict price changes, stock trends, and future market behavior to make decisions such as buying or selling stocks or developing investment strategies. Most works we describe explore up to one or two types of AI tools. The objective of this research is to analyze the use of advanced AI techniques (machine learning, neural networks, genetic algorithms, fuzzy methods, etc.) and big data in the stock market, creating a hybrid model. We consider AI and big data as two rapidly evolving areas with significant potential, that together can improve the understanding of financial markets.

In recent years, investors have been seeking resources to facilitate their decisions in the financial market and, in this context, researchers have reported the difficulty of predicting stock prices. One factor behind this difficulty is related to the number of stocks that are traded daily because several factors, such as political, social, economic, and others, can influence the variation of each. Some researchers have proposed therapeutic and computational models that consider improvements in techniques for selecting the action of a company based on past data or looking for resources that have been developed in other scientific areas to help in the prediction of price. They emphasize both traditional forms and the use of novel approaches as important for analyzing the current state of the stock market and making investment decisions. These novel forms are based on advanced artificial intelligence and big data techniques. The use of stocks that are part of the Ibovespa, and Dow Jones, among others, introduces several difficulties due to the volume of data, and lack of quality and accuracy.

Fig 4 : Stock market prediction in % vs. ML Techniques used for it (LSTM: long short-term memory, GBM: gradient boosted models, CNN: convolutional neural network, RNN: recurrent neural networks, RF: random forest, SVM: support vector machine, and LR: linear regression).

5. Challenges and Future Directions

We report and examine some of the common challenges present in the literature relevant to big data implementing AI approaches for stock market prediction. Following this section is the conclusion. The key problem today's investment advisors and stock investors face is how to release the hidden knowledge that is already contained in the publicly available information, which may be termed as explicit knowledge for stock prediction. The challenge here is to identify the association rules about stock movements which help investors decrease trading risks. Additionally, identifying explicit knowledge might aid investors in reducing investment errors, thus sweetening the investing result.

5.1 Intelligent Behavior Modeling

Many behavior models have been designed to analyze stock market behaviors, and researchers have also shown a great deal of interest in explaining stock market volatility based on the real behavior of traders. However, what we need are not explanations of the happening of stock behaviors (which some researchers call explaining behaviors), but positive recommendations on stock investments by performing behavior-oriented action (which some researchers call modeling behaviors), because the ultimate goal of any researcher analyzing the stock market is to offer instructive advice about stock investment to facilitate shareholder benefit from equity investment. In other words, analysts predict stock market behaviors to make accurate investments, not just for analysis.

5.1. Current Challenges in Stock Market Prediction

The stock market mechanism is a nonlinear mechanism, where investment behavior usually suits irrational behaviors. Therefore, it is hard to predict the stock market through traditional econometric models like Moving Average and ARMA. Furthermore, the previous methods happened to suit the financial professionals, causing a monopolization of stock investment. However, the WEKA library has been established for data mining. Hence, the linear regression model in data mining, like Support Vector Machine, Polynomial Regression, and Random

Forest, may force investors to act rationally, and it is found to be more effective than the traditional econometric method.

Nonlinear reward and risk of investment and the failure to analyze future stock prices lead to many challenges, like the problem of information explosion, herd behaviors, artificial manipulations, and high position risks, to the financial research community. This has caused a narrowing difference in evidence-based finance theories on micro-stock between the world of finance and the world of philosophy. Based on the above reasoning, this paper argues that it is outdated for an individual to analyze the stock market by employing the artificial analysis method that merely uses their own experience and knowledge to provide any good advice on stock investment nowadays. This fact provides the economic basis for the state-owned news-machine-controlled policy. This paper will thus apply AI techniques to analyze the stock market to invigorate economic freedom.

5.2. Future Trends and Research Directions

Section Title: Future Trends and Research Directions

The presented conceptual framework provides an initial means to analyze big data and integrate it with machine learning to predict stock market trends and movements and improve the design of market investment strategies. It is anticipated that the underlying fusion contributes to the growing body of knowledge about the use of AI in big data decision support systems and tools for stock market prediction purposes, which remain scantly distributed and often come from disparate AI disciplines. The opportunities and challenges of the suggested support system are also highlighted for further investigation. As a result of the significant advances in big data technologies and AI, researchers are encouraged to explore the practical application of the proposed multilayer conceptual framework. Some possible future research directions and questions that may have remained unanswered from the current conception are provided next.

How can the proposed framework better link and integrate big data analytics, deep learning, and support vector regression models to maximize the stock market prediction accuracy over the short and long run? Can the framework broaden and enrich the application of big data integration for other varied deep learning and traditional machine learning and AI models? Can additional big data be fused within the conceptual framework such as social media, news, and real-world events, to explore its data-driven insight in predicting stock market behavior and trends? How does this multimodal-fused data complement and/or enhance the accuracy of the stock market prediction models when compared to the single-modal-fused data? How can the framework help devise flexible automated trading strategies on a day-to-day basis? Can the framework add structured data representation for more effective training in the prediction of both typical and risky stocks? How to leverage more multimodal techniques to maximize stock market prediction accuracy performance implemented on embedded hardware? Analyses of the existing trade-offs within the integrated multimodal big data analytics architecture are needed and thereby help to identify possible enhancements and create more robust, secure, and reconfigurable multimodal models, techniques, and applications.

Equation 3: SVM-RFE: selection and visualization of the most relevant features through non-linear kernels

Data : Dataset with p^* variables and binary outcome.
Output: Ranked list of variables according to their relevance.

Find the optimal values for the tuning parameters of the SVM model;
Train the SVM model;
$p \leftarrow p^*$;
while $p \geq 2$ **do**
  $SVM_p \leftarrow$ SVM with the optimized tuning parameters for the p variables and observations in **Data**;
  $w_p \leftarrow$ calculate weight vector of the SVM_p $(w_{p1}, \ldots, w_{pp})$;
  $rank.criteria \leftarrow (w_{p1}^2, \ldots, w_{pp}^2)$;
  $min.rank.criteria \leftarrow$ variable with lowest value in $rank.criteria$ vector;
  Remove $min.rank.criteria$ from **Data**;
  $Rank_p \leftarrow min.rank.criteria$;
  $p \leftarrow p - 1$;
end
$Rank_1 \leftarrow$ variable in **Data** $\notin (Rank_2, \ldots, Rank_{p^*})$;
return $(Rank_1, \ldots, Rank_{p^*})$

6. Conclusion

AI has become a powerful tool for solving decision-making problems in real life. In finance, stock market prediction is a critical and complex task due to its three main features: non-linearity, interaction pattern, and chaos. Complex AI techniques have been successfully applied to stock market prediction. However, these techniques are data-driven and sensitive to the generation of an input dataset. If the input data lacks information, a model may generate inaccurate predictions. Currently, most of the data collected from the stock market for learning is traditional. Hence, experts need to integrate big data to build decision-making models for accurate stock market prediction.
In this paper, we demonstrated the integration of big data and AI in stock market prediction. A new improved big data framework for stock prediction was developed using novel methods, such as text mining, financial sentiment analysis, and the deep belief network. Experimental results prove that our model outperforms state-of-the-art methodologies for stock prediction. In summary, our findings not only provide empirical evidence regarding the integration of new technology in finance theory but also identify a new way for investors to analyze stock market trends. The implementation of the latest model leads to many follow-up questions and areas for future study.

7. References

[1] Pillai, S. E. V. S., Avacharmal, R., Reddy, R. A., Pareek, P. K., & Zanke, P. (2024, April). Transductive–Long Short-Term Memory Network For The Fake News Detection. In 2024 Third International Conference On Distributed Computing And Electrical Circuits And Electronics (Icdcece) (Pp. 1-4). Ieee.
[2] Bhattacharya, S., Najana, M., Khanna, A., & Chintale, P. (2024). Securing The Gatekeeper: Addressing Vulnerabilities In Oauth Implementations For Enhanced Web Security. International Journal Of Global Innovations And Solutions (Ijgis).
[3] Vaka, D. K. (2023). Achieving Digital Excellence In Supply Chain Through Advanced Technologies. Educational Administration: Theory And Practice, 29(4), 680-688.
[4] Zanke, P., Deep, S., Pamulaparthyvenkata, S., & Sontakke, D. Optimizing Worker's Compensation Outcomes Through Technology: A Review And Framework For Implementations.
[5] Mahida, A. Secure Data Outsourcing Techniques For Cloud Storage.
[6] Kommisetty, P. D. N. K., & Nishanth, A. (2024). Ai-Driven Enhancements In Cloud Computing: Exploring The Synergies Of Machine Learning And Generative Ai. In Iarjset (Vol. 9, Issue 10). Tejass Publishers. Https://Doi.Org/10.17148/Iarjset.2022.91020

[7] Purshotam S Yadav. (2024). Optimizing Serverless Architectures For Ultra-Low Latency In Financial Applications. European Journal Of Advances In Engineering And Technology. Https://Doi.Org/10.5281/Zenodo.13627245
[8] Avacharmal, R. (2024). Explainable Ai: Bridging The Gap Between Machine Learning Models And Human Understanding. Journal Of Informatics Education And Research, 4(2).
[9] Balakrishnan, A., Jain, V., Chintale, P., Gadiparthi, S., & Najana, M. (2024). Blockchain Empowerment In Sanctions And Aml Compliance: A Transparent Approach. International Journal Of Computer Trends And Technology.
[10] Vaka, D. K. Empowering Food And Beverage Businesses With S/4hana: Addressing Challenges Effectively. J Artif Intell Mach Learn & Data Sci 2023, 1(2), 376-381.
[11] Pamulaparthyvenkata, S., & Avacharmal, R. (2023). Leveraging Interpretable Machine Learning For Granular Risk Stratification In Hospital Readmission: Unveiling Actionable Insights From Electronic Health Records. Hong Kong Journal Of Ai And Medicine, 3(1), 58-84.
[12] Mahida, A., Chintale, P., & Deshmukh, H. (2024). Enhancing Fraud Detection In Real Time Using Dataops On Elastic Platforms.
[13] Kommisetty, P. D. N. K., & Abhireddy, N. (2024). Cloud Migration Strategies: Ensuring Seamless Integration And Scalability In Dynamic Business Environments. In The International Journal Of Engineering And Computer Science (Vol. 13, Issue 04, Pp. 26146–26156). Valley International. Https://Doi.Org/10.18535/Ijecs/V13i04.4812
[14] Yadav, P. S. (2024). Fast And Efficient Userid Lookup In Distributed Authentication: A Probabilistic Approach Using Bloom Filters. In The International Journal Of Computing And Engineering (Vol. 6, Issue 2, Pp. 1–16). Cari Journals Limited. Https://Doi.Org/10.47941/Ijce.2124
[15] Avacharmal, R., Pamulaparthyvenkata, S., & Gudala, L. (2023). Unveiling The Pandora's Box: A Multifaceted Exploration Of Ethical Considerations In Generative Ai For Financial Services And Healthcare. Hong Kong Journal Of Ai And Medicine, 3(1), 84-99.
[16] Tabbassum, A., Chintale, P., Praveen, G., & Najana, M. (2024). The Impact Of Ai On Future Employment Patterns. International Journal Of Global Innovations And Solutions (Ijgis).
[17] Pamulaparthyvenkata, S. (2023). Optimizing Resource Allocation For Value-Based Care (Vbc) Implementation: A Multifaceted Approach To Mitigate Staffing And Technological Impediments Towards Delivering High-Quality, Cost-Effective Healthcare. Australian Journal Of Machine Learning Research & Applications, 3(2), 304-330.
[18] Vaka, D. K. "Artificial Intelligence Enabled Demand Sensing: Enhancing Supply Chain Responsiveness.
[19] Mahida, A. (2024). Integrating Observability With Devops Practices In Financial Services Technologies: A Study On Enhancing Software Development And Operational Resilience. International Journal Of Advanced Computer Science & Applications, 15(7).
[20] Kommisetty, P. D. N. K., & Dileep, V. (2024). Robust Cybersecurity Measures: Strategies For Safeguarding Organizational Assets And Sensitive Information. In Ijarcce (Vol. 13, Issue 8). Tejass Publishers. Https://Doi.Org/10.17148/Ijarcce.2024.13832
[21] Yadav, P. S. (2024). Advanced Authentication And Authorization Mechanisms In Apache Kafka: Enhancing Security For High-Volume Data Processing Environments. In Journal Of Engineering And Applied Sciences Technology (Pp. 1–6). Scientific Research And Community Ltd. Https://Doi.Org/10.47363/Jeast/2024(6)E110
[22] Avacharmal, R., Sadhu, A. K. R., & Bojja, S. G. R. (2023). Forging Interdisciplinary Pathways: A Comprehensive Exploration Of Cross-Disciplinary Approaches To Bolstering Artificial Intelligence Robustness And Reliability. Journal Of Ai-Assisted Scientific Discovery, 3(2), 364-370.
[23] Seth, D., & Chintale, P. Performance Benchmarking Of Serverless Computing Platforms.
[24] Pamulaparthyvenkata, S., Reddy, S. G., & Singh, S. (2023). Leveraging Technological Advancements To Optimize Healthcare Delivery: A Comprehensive Analysis Of Value-Based Care, Patient-Centered Engagement, And Personalized Medicine Strategies. Journal Of Ai-Assisted Scientific Discovery, 3(2), 371-378.
[25] Mahida, A. Explainable Generative Models In Fincrime. J Artif Intell Mach Learn & Data Sci 2023, 1(2), 205-208.

[26] Kommisetty, P. D. N. K., Vijay, A., & Bhasker Rao, M. (2024). From Big Data To Actionable Insights: The Role Of Ai In Data Interpretation. In Iarjset (Vol. 11, Issue 8). Tejass Publishers. Https://Doi.Org/10.17148/Iarjset.2024.11831

[27] Yadav, P. S. (2023). Enhancing Software Testing With Ai: Integrating Junit And Machine Learning Techniques. North American Journal Of Engineering Research, 4(1).

[28] Vaka, D. K. " Integrated Excellence: Pm-Ewm Integration Solution For S/4hana 2020/2021.

[29] Avacharmal, R., Gudala, L., & Venkataramanan, S. (2023). Navigating The Labyrinth: A Comprehensive Review Of Emerging Artificial Intelligence Technologies, Ethical Considerations, And Global Governance Models In The Pursuit Of Trustworthy Ai. Australian Journal Of Machine Learning Research & Applications, 3(2), 331-347.

[30] Jain, V., Balakrishnan, A., Beeram, D., Najana, M., & Chintale, P. Leveraging Artificial Intelligence For Enhancing Regulatory Compliance In The Financial Sector.

[31] Tilala, M., Pamulaparthyvenkata, S., Chawda, A. D., & Benke, A. P. Explore The Technologies And Architectures Enabling Real-Time Data Processing Within Healthcare Data Lakes, And How They Facilitate Immediate Clinical Decision-Making And Patient Care Interventions. European Chemical Bulletin, 11, 4537-4542.

[32] Mahida, A. (2023). Enhancing Observability In Distributed Systems-A Comprehensive Review. Journal Of Mathematical & Computer Applications. Src/Jmca-166. Doi: Doi. Org/10.47363/Jmca/2023 (2), 135, 2-4.

[33] Researcher. (2024). Leveraging Artificial Intelligence In Managing Us Health Data: Implications For Medicare And Medicaid Claims And Data Security. Zenodo. Https://Doi.Org/10.5281/Zenodo.12774053

[34] Vaka, D. K. (2020). Navigating Uncertainty: The Power Of ‘Just In Time Sap For Supply Chain Dynamics. Journal Of Technological Innovations, 1(2).

[35] Yadav, P. S. Real-Time Insights In Distributed Systems: Advanced Observability Techniques For Cloud-Native Enterprise Architectures.