CS F407 – ARTIFICIAL INTELLIGENCE

PREDICTION OF STOCKS USING ARTIFICIAL INTELLIGENCE

USING

GENETIC ALGORITHM

TEAM MEMBERS ID	NAMES
1 2020AAPS0329H	Rakshith Dasari
2 2020AAPS0295H	Devesh S
3 2020A3PS0491H	Aman Jain
4 2020A1PS2412H	Pathuri Charith Goud

GROUP ID: 56

1. INTRODUCTION

For many years, investors, traders, and scholars have been interested in stock forecasting. The stock market is a complicated system that is impacted by a wide range of things, including political developments, calamities of all kinds, economic policies, and even public opinion. These elements have the potential to cause abrupt and unforeseen fluctuations in stock values. Thus, for effective trading and investing, precise stock forecasting is crucial.

In recent years, the use of artificial intelligence (AI) to stock prediction has drawn a lot of interest. A lot of data can be analysed and interpreted by AI in real-time, which can provide important information about market patterns and stock price changes. In order to forecast future stock values, artificial intelligence (AI) techniques like machine learning, deep learning, and natural language processing may be used to examine financial data, news articles, social media posts, and other pertinent information.

On the basis of past data, machine learning algorithms may be trained to spot patterns and make predictions. Based on technical analysis, which examines previous stock prices and trading volume, these algorithms can be used to forecast stock values. The goal of technical analysis is to spot trends and patterns in stock price changes that may be utilised to forecast future prices. In order to anticipate stock prices, fundamental research, which looks at firm financial statements, market movements, and economic indicators, also uses machine learning algorithms.

Artificial neural networks (ANNs) are used in deep learning, a kind of machine learning, to learn and predict the future. ANNs are able to analyse a lot of data concurrently because they are designed after the structure and operation of the human brain. To assess complex data, such as photos, audio, and spoken language, deep learning techniques can be utilised. Deep learning algorithms may be used to evaluate news articles and social media postings in the context of stock prediction in order to gauge market sentiment and forecast future stock values.

A branch of artificial intelligence called "natural language processing" (NLP) is concerned with how computers and human language interact. News stories, social media postings, and other text-based data may be analysed using NLP algorithms to find pertinent data about businesses, industries, and economic trends. It is possible to forecast future stock prices using this knowledge.

One of the most intricate and dynamic systems in the world is the stock market. Many variables, including economic statistics, current affairs, and investor mood, have an impact on the stock price index. As a result, analysts and investors now have a difficult problem when attempting to forecast stock values. Artificial neural networks (ANNs) and genetic algorithms (GAs) are two examples of machine learning algorithms that have gained prominence in recent years for stock prediction. Genetic algorithms in particular have been demonstrated to be successful in feature selection, parameter optimization, and model development, making them a potential method for stock price forecasting.

The natural selection process served as the inspiration for genetic algorithms, which are optimization algorithms. They are founded on the idea of Darwinian evolution, which holds that people with the fittest genes have a higher chance of surviving and procreating. By the use of genetic operators including crossover, mutation, and selection, a population of possible solutions is created and evolves over several generations in genetic algorithms. The fitness function, which gauges the calibre of the solution, is used by the genetic algorithm to rank each member of the population. The more fit individuals have a higher chance of surviving, reproducing, and transferring their genes to the following generation.

Researchers have recently become interested in the use of genetic algorithms in stock prediction. While searching for the most pertinent features that might affect stock prices, genetic algorithms can be utilised for feature selection. The selection of features can assist make the input data less dimensional and boost the effectiveness of the prediction model. The number of hidden neurons, learning rate, and momentum of an ANN may all be optimised using genetic algorithms. The prediction model may converge more quickly and with more accuracy as a result of this optimization process.

Using AI for stock prediction has advantages, but it also has drawbacks and restrictions. The accessibility and calibre of the data provide one of the main difficulties. Financial data may be erroneous and biassed, in addition to being complicated and challenging to gather. However, the stock market is a complicated system that is impacted by a variety of variables, some of which may be challenging to forecast. As a result, utilising AI to anticipate stock values may be challenging.

The interpretability of AI models is another problem. Understanding how the model generated its predictions might be challenging due to the difficulty in interpreting machine learning and deep learning algorithms. For traders and investors, who must comprehend the underlying reasoning behind the projections in order to make wise judgements, this may be a challenge.

A fascinating area of study that has the potential to completely change how traders and investors make investment decisions is the use of AI for stock prediction. A lot of data may be analysed to find patterns that people would miss using AI tools like machine learning, deep learning, and natural language processing. The accessibility, accuracy, and interpretability of data, as well as the limits of employing AI for stock prediction, present additional difficulties and constraints. Future stock prediction models using AI are anticipated to be increasingly advanced and precise as the field of artificial intelligence develops.

2. LITERATURE SURVEY

(Štěpánek et al. 97) explores the use of genetic algorithms to model stock market behaviour using psychological analysis and a multi-agent system. The simulation's objective is to assist traders in maximising earnings by taking advantage of brief changes in stock prices. The simulation can develop decision-making processes that precisely mimic the actions of dealers in the real world by using genetic algorithms. Simulations are crucial for education because they let students experiment with new ideas without running the risk of failure in the real world. Additionally, the system offers a platform for comparing strategies, comparing causal connections, and expanding the simulation. The final goal is to create a sophisticated stock market simulation with an intuitive user interface for testing and updating new trader behaviour programmes. Utilizing actual stock market data, the system assesses each agent's success and provides feedback to contributors so they can enhance their programmes. Overall, the simulation provides a cooperative setting where traders can exchange ideas and cooperate to get improved outcomes.

In order to increase the precision of stock price index prediction, the research paper from (Kyoung-jae Kim and Ingoo Han 93) suggests using genetic algorithms to carry out feature discretization and choose connection weights for artificial neural networks (ANNs). The suggested method reduces dimensionality through globally searched feature discretization, which addresses ANN's shortcomings in handling noisy and complex data. The findings of the study's experiments comparing the proposed approach's performance to that of conventional models demonstrate that it outperforms them. In order to improve the precision of stock price index prediction, the study emphasises the significance of concurrently optimising the learning algorithm, reducing dimensionality, and removing irrelevant patterns.

(H and K 20) seek to develop a deep learning-based stock market prediction model that can effectively learn from massive datasets. The suggested approach combines a long short-term memory (LSTM) network and a genetic algorithm (GA) to examine the temporal characteristics of stock market data and to recommend a methodical approach to picking the topology and size of the time window for the LSTM network using GA. Using the daily Korea Stock Price Index (KOSPI) data to measure performance, it is discovered that the hybrid method outperforms the benchmark model. In order to manage potential risks and attain socioeconomic sustainability, the study emphasises the significance of stock market volatility analysis. The use of big data and AI techniques in finance can provide useful insights and surmount the limitations of conventional statistical methods. The ability of LSTM networks to capture the nonlinear and unstructured nature of financial time series data is another benefit highlighted by the research for LSTM networks in time series analysis. However, it is noted that some of the drawbacks of LSTM networks include the inability to explain the final choice and the requirement to change numerous parameters. These drawbacks can be resolved by integrating genetic algorithms to give precise explanations for prediction outcomes and choose the ideal set of parameters.

(Selvamuthu et al. 10)looks at how Intelligent Trading Systems (ITS) that employ neural network learning algorithms can help traders forecast stock prices and make investment choices.

Tick-by-tick data and 15-minute data are used in the research to compare the performance of three algorithms (Levenberg-Marquardt, Scaled Conjugate Gradient, and Bayesian Regularization). The results indicate that outcomes from tick-by-tick data are better, and while Bayesian Regularization is the most accurate, it requires more time to train. Recurrent neural networks and sentiment analysis may provide even more accurate predictions, according to the study. The stock market is a complicated and dynamic system overall, but ITS and machine learning can help dealers make better choices.

(Inthachot et al.) investigates the efficacy of combining a genetic algorithm (GA) and an artificial neural network (ANN) to forecast the trend of Thailand's SET50 index through technical analysis. GA is a feature selection method that can increase the accuracy of ANN predictions by choosing better subsets of input variables. ANN is a commonly used machine learning algorithm for stock analysis. To produce more diverse subsets of input, the study suggests using technical indicators with four input variables each, each representing four past time spans of various lengths (3, 5, 10, and 15 days). Low accuracy exists in previous techniques that use a single input variable for a single fixed length of the prior time span. The results of the study demonstrate that the hybrid model is more accurate than the previous approach. The study utilises SET50 index data from 2009 to 2014 to evaluate the hybrid intelligence prediction model.

(S and C,2021)'s study suggests a stock price prediction model that blends an enhanced Long Short-Term Memory (LSTM) neural network with a Genetic Algorithm (GA) for feature selection. The GA is used to rate the significance of stock price-influencing factors, and the ideal combination of factors is then discovered through trial and error. The LSTM algorithm is then applied to stock prediction. Non-linear and non-stationary financial time series are challenging to model using conventional prediction methods. Stock price predictions have been made using machine learning techniques like decision trees, neural networks, and support vector machines. Due to their capacity to examine intricate nonlinear relationships in data, deep learning models like CNN, RNN, and LSTM have produced encouraging results. Multivariable financial timeline prediction requires effective feature selection. Empirical investigations using the China Construction Bank dataset and the CSI 300 stock dataset show that the proposed model outperforms all baseline models for time series prediction.

(Urszula Markowska-Kaczmar et al. #) outlines a genetic algorithm-evolved virtual stock market model. The model consists of agents that mimic real investors and seek to produce stock values on a virtual market that, for a brief period of time, resemble those of actual markets. The major elements of the genetic algorithm are described, including an individual, genetic operators, and fitness function. The research on the role of genetic algorithm parameters and the agent's propensity to forecast quotation values is presented in this paper. The article also covers various methods for predicting and simulating market behaviour. The use of multi-agent techniques is emphasised, and various financial instrument assumptions are made for the agents. A description of the system's potential future growth comes as the paper's conclusion. In its entirety, the article suggests a genetic algorithm-based method for simulating the stock market, which may be helpful for decision-making regarding investments.

(Huang et al. #) uses a hybrid fuzzy-genetic algorithm model for efficient stock selection is proposed, where GA is used for fuzzy model parameter optimization and feature selection. The selection of stocks that offer above-average returns on investment is a problem for the finance industry. The development of workable quantitative models for forecasting financial time series, optimising investment returns and risk management, and choosing investment tools for portfolio management is now possible thanks to recent advancements in computational intelligence and data mining. Artificial neural networks, support vector machines, evolutionary algorithms, and fuzzy inference models are a few examples of machine learning methodologies that have been created specifically for this use. These models could, however, experience overfitting, local optima, and inadequate learning capacity. The success of these models depends on feature selection because the features chosen impact the classification accuracy, computational complexity, and quantity of training examples required. In order to be effective, a model must concurrently consider the feature selection and parameter optimization of fuzzy models. The suggested model is a wrapper approach, which performs better than the filter method because feature selection is dependent on the results of the stock selection model. The encouraging findings imply that this hybrid model can progress soft computing in finance and offer a practical stock selection solution.

(Chung and Shin #) shows the value of the hybrid GA-CNN method and points to a promising line of inquiry for future work on deep learning applications to financial analysis. The use of deep learning, in particular multi-channel convolutional neural networks (CNNs), to forecast stock market fluctuations is covered in this article. Due to their capacity to extract local features from data and recognise the temporal property of the dataset, CNNs have demonstrated success in a variety of time-series issues. However, because there are so many different learning strategies, network topologies, and other regulating factors, optimising the network topology of CNNs can be difficult. The genetic algorithm (GA), which has been successful in solving optimization problems in conjunction with other machine learning techniques, is used by the authors of this study to suggest a method for systematically optimising the CNN model's parameters. The GA-CNN model that was optimised did better than other models at forecasting the KOSPI stock index.

The paper by (Hyun-jung Kim and Kyung-shik Shin #) outlines a hybrid strategy for stock market forecasting that combines genetic algorithms, time delay neural networks, and adaptive time delay neural networks (ATNNs) (GAs). Artificial neural networks (ANNs) with hidden layers are better adapted for handling non-linearities than statistical techniques, which have limitations in the prediction of non-linear issues. The BPN ANNs, however, can only acquire input-output mapping of static patterns that are time-independent. By incorporating a memory into the ANN using time-delayed links, the ATNN and TDNN models have an advantage when managing temporal patterns. However, performance might not be enhanced by counting the amount of time delays and network architecture elements in stand-alone mode. Therefore, it is suggested that GAs be used to concurrently optimise these factors for the ATNN and TDNN models. According to experimental findings, the hybrid ATNN and TDNN with GAs method outperforms the traditional ATNN, TDNN, and RNN models in stock market prediction. According to the study's findings, this method works well for handling temporal patterns while concurrently optimising network architecture elements for stock market prediction tasks.

(E. N. Desokey et al. #) examines the difficulties the stock market has in dealing with social media platforms as well as the use of social media data to forecast the stock market. In order to improve the clustering of stock market prediction, this paper suggests a new model that combines the genetic algorithm and the k-means clustering algorithm. Earlier research in this area did not produce acceptable results. The selected clustering algorithm's sum of square distances (SSD) and a Chi-Square similarity measure are both used in the suggested model to determine accuracy. According to the assessment, the suggested model had a high accuracy rate of 89.31%, which is more pleasing to investors. The suggested model will be tested in Arabic in the future, and the fuzzy model will be used for results that are more precise.

The study (R et al. #) proposes a technique for forecasting daily stock trends that combines a Random Forest classifier with a feature selection model based on the Genetic Algorithm. The model analyses the trends of 15 stocks and takes into account four foreign stock indices. The research emphasises how difficult it is to forecast stock trends because of the influence of numerous outside factors. Additionally, it covers frequently used machine learning methods for predicting stock trends, such as Support Vector Machines and Artificial Neural Networks. The findings demonstrate that the suggested model beats the dummy forecast and has a prediction accuracy of 80%. The CAC40 index performs the least well when it comes to forecasting everyday stock trends compared to the S&P 500 index. The importance of feature selection models in raising prediction efficiency and accuracy is emphasised in the article. Instead of relying on specific features, the proposed model employs global stock indices to capture changes in the global market. The model's stages include choosing global stock indices, picking a stock for forecasting, choosing features with the Genetic Algorithm, and using the Random Forest algorithm to unearth undiscovered relationships between the features and the stock's trend.

In order to forecast the stock prices of the Hong Kong Hang Seng index, the paper (Zhu and Wang #) suggests a novel trading system based on support vector regression (SVR) and multilayer perceptron (MLP) that has been enhanced with genetic algorithms (GA). The article makes the case that the proposed system outperforms conventional trading systems without prediction and fuzzy trading systems, as demonstrated by experimental findings in terms of final equity and maximum drawdown. To minimise the risk of curve fitting, the suggested system only uses price and technical indicators as input data, and the testing period is 10 years. The predictive model uses the last 8 days' PPO as input to anticipate the next 5 days' PPO in order to make trading decisions. The system is primarily based on the price percentage oscillator (PPO) trading rules.It also touches on earlier studies that were aimed at reducing mean square error (MSE) in price direction prediction and generating paper gains in trading financial markets. The proposed system differs from earlier study in that it predicts using MLP and SVR models that have undergone GA optimization. The suggested system's promising financial returns are highlighted in the paper's conclusion, along with some possible directions for future development and study.

3. CONCLUSION

In recent years, stock prediction using AI has gained popularity as a research topic, with multiple studies examining various strategies to increase forecast accuracy. Artificial neural networks (ANNs), which are capable of modelling intricate interactions between inputs and outputs, are one popular strategy. An example of a machine learning algorithm that can learn from a lot of data and produce precise predictions is an ANN. Nonetheless, the performance of ANNs may be enhanced with further modification because of their limitations in handling complicated and noisy data.

Genetic algorithms have been suggested as a way to improve artificial neural networks (ANNs) for stock prediction (GA). GA is an optimization method that draws its inspiration from genetic inheritance and natural selection. It operates by choosing the fittest members of a population of potential solutions to breed and mutate over the course of several generations.

Numerous research have looked at the application of GA to enhance ANN performance in stock prediction. A approach that employs GA to carry out feature discretization and choose connection weights for ANNs, for instance, was proposed by Kyoung-jae Kim and Ingoo Han (Kyoung-jae Kim and Ingoo Han) in their paper. The study showed that the method outperformed traditional models and stressed the significance of concurrent learning algorithm optimization, dimensionality reduction, and pattern removal.

The usage of a multi-agent system using genetic algorithms has also been suggested, as examined by Štěpánek et al. The simulation's goal is to help traders maximise profits by profiting from sudden swings in stock prices. The system creates decision-making processes that resemble the activities of actual traders using genetic algorithms. Simulations are essential to education because they let students try out novel concepts without having to worry about failing in the real world.

Additionally, by integrating a long short-term memory (LSTM) network with a GA, H and K proposed a deep learning-based stock market prediction model that successfully learns from enormous datasets. The benchmark model was exceeded by the hybrid approach, which highlighted the significance of stock market volatility analysis for minimising risks and achieving socioeconomic sustainability.

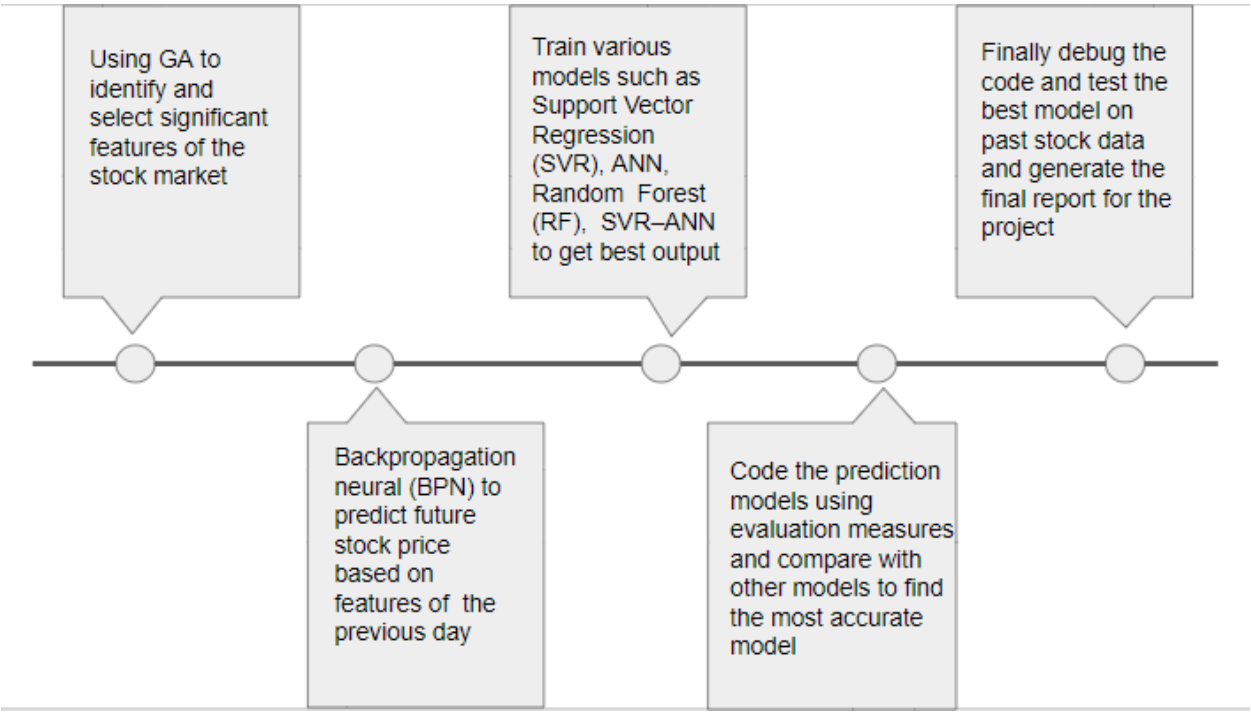
Several research have investigated the use of intelligent trading systems (ITS) that incorporate neural network learning methods in addition to the usage of GA with ANNs, as detailed in Selvamuthu et al. Using tick-by-tick data, the study examined the performances of three algorithms and discovered that Bayesian regularisation was the most accurate but took the longest to train. According to the study, sentiment analysis and recurrent neural networks may provide predictions that are even more precise.

The effectiveness of combining a GA with an ANN to predict the trend of Thailand's SET50 index using technical analysis was also examined by Inthachot et al. According to the study, the

hybrid model was more precise than earlier methods that just employed one input variable for a single set duration of the past time period.

In conclusion, stock prediction using AI approaches like genetic algorithms and artificial neural networks has shown promising outcomes. These techniques can increase the accuracy of predictions and help traders choose their investments more wisely. The application of AI in finance is anticipated to increase in the future, despite the techniques' drawbacks, such as the requirement for vast volumes of data and difficult parameter adjustment.

TIMELINE


Using GA to identify and select significant features of the stock market
Backpropagation neural (BPN) to predict future stock price based on features of the previous day
Train various models such as Support Vector Regression (SVR), ANN, Random Forest (RF), SVR–ANN to get best output
Code the prediction models using evaluation measures and compare with other models to find the most accurate model
Finally debug the code and test the best model on past stock data and generate the final report for the project

REFERENCES/BIBLIOGRAPHY

Chung, Hyejung, and Kyung-shik Shin. “Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction.” *Neural Computing and Applications*, vol. 32, 2019, https://doi.org/10.1007/s00521-019-04236-3.

E. N. Desokey, et al. “Enhancing stock prediction clustering using K-means with genetic algorithm.” no. 13th International Computer Engineering Conference (ICENCO), 2017, pp. 256-261.

H, Chung, and Shin K. “Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction.” *Sustainability*, vol. 10, no. 10, 2018, p. 20.

Huang, Chien-Feng, et al. “Feature Selection and Parameter Optimization of a Fuzzy-based Stock Selection Model Using Genetic Algorithms.” *International Journal of Fuzzy Systems*, vol. 14, 2012.

Hyun-jung Kim, and Kyung-shik Shin. “A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets.” *Applied Soft Computing*, vol. 7, no. 2, 2007, pp. 569-576, https://doi.org/10.1016/j.asoc.2006.03.004.

Inthachot, Montri, et al. “Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend`.” *Computational Intelligence and Neuroscience*, vol. 2016, 2016.

Kyoung-jae Kim, and Ingoo Han. “Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index.” *Expert Systems with Applications*, vol. 19, no. 2, 2000, pp. 125-132.

R, Abraham, et al. “Forecasting a Stock Trend Using Genetic Algorithm and Random Forest.” *Forecasting a Stock Trend Using Genetic Algorithm and Random Forest*, vol. 15(5), no. International Finance, 2022, https://www.mdpi.com/1911-8074/15/5/188.

S, Chen, and Zhou C. “Stock Prediction Based on Genetic Algorithm Feature Selection and Long Short-Term Memory Neural Network.” *IEEE Access*, vol. 9, p. 2021.

Selvamuthu, Dharmaraja, et al. “Indian stock market prediction using artificial neural networks on tick data.” *Financial Innovation*, vol. 5, no. 10, 2019, p. 10.

Štěpánek, Jiří, et al. “Application of Genetic Algorithms in Stock Market Simulation.” *Procedia - Social and Behavioral Sciences*, vol. 47, no. Cyprus International Conference on Educational Research, 2012, pp. 93-97. *Application of Genetic Algorithms in Stock Market Simulation*, https://doi.org/10.1016/j.sbspro.2012.06.619.

Urszula Markowska-Kaczmar, et al. “Genetic Algorithm as a Tool for Stock Market Modelling.” *Artificial Intelligence and Soft Computing*, vol. 5097, no. ICAISC, 2008.

Zhu, M., and L. Wang. “Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms.” no. International Joint Conference on Neural Networks (IJCNN), 2010, pp. 1-5.