OPEN ACCESS

Volume: 11

Special Issue: 1

Month: July

Year: 2023

E-ISSN: 2582-0397

P-ISSN: 2321-788X

Impact Factor: 3.025

Received: 10.05.2023

Accepted: 18.06.2023

Published: 01.07.2023

Citation:
Deepa, KR, and S. Amrutha. "Utilising Deep Learning and Machine Learning Concepts to Forecast Share Trading Changes." *Shanlax International Journal of Arts, Science and Humanities*, vol. 11, no. S1, 2023, pp. 8–13.

DOI:
https://doi.org/10.34293/sijash.v11iS1-July.6308

Utilising Deep Learning and Machine Learning Concepts to Forecast Share Trading Changes

Deepa K.R
Department of Master of Computer Applications
Rajarajeswari College of Engineering, Bengaluru.

Amrutha S
Department of Master of Computer Applications
Rajarajeswari College of Engineering, Bengaluru

Abstract

Due to a variety of deciding factors, The way the stock exchange fluctuates haslong been unclear to investors. In this work, deep learning and machine learning techniques are used. to drastically lower the risk associated with trend prediction. Four categories of stock markets, include diversified financials,petroleum, non-metallic The Tehran Stock Exchange's minerals and metals of base are picked for experimental evaluations. Nine Adaptive Boosting(Naive Bayes, K-Nearest Neighbours, Logistic Regression, eXtreme gradient booster, a Support Vector Classifier, Adaboost, XGBoost, ANN) and two. effective deep learning techniques (Long Short-term Memory (LTM) Additionally, Recurrent Neural Networks (RNN)memory (LSTM).Ten Our values for input are technical indices derived Ten years' worth of prior data information, and two methods are intended to: for employing them. First, The signals are created utilising constant information from stock trading values, which is transformed prior use, binary data. Each model for prediction depends on the data input techniques, and assessed using three metrics. The assessment findings show that for continuous data, LSTM and RNN beat additional forecasting techniques significantly. Additionally, findings indicate that such Although deep learning approaches are the best for evaluating binary data, the difference becomes less important due to the second method's significantly improved model performance.

Keywords: Machine Learning, Deep Learning, Stock Prediction, SVM, XGBoost, KNN, ANN

Introduction

Stock forecasting has traditionally been a difficult topic for statisticians and financiers. The key purpose for this forecast is to acquire stocks They are anticipated to appreciate in value, after which be sold stocks whose prices are anticipated to drop. There are often two approaches. to forecast the stock price. Among these is fundamental analysis, It is founded on a company's strategy and fundamental data such as its position in the market, its costs, and its annual growth rate. utilisation of technical analysis, which places a strong emphasis on prior stock prices and values, is the second method. This study forecasts future prices using past charts and trends.

Financial gurus have traditionally predicting historical stock market trends. However, data scientists are increasingly employed since the progress of learning methodologies to solve prediction difficulties. Additionally, computer scientists are increasingly utilising machine learning techniques to enhance the efficacy of prediction models and increase forecast accuracy.

The next move is in increasing prediction models' performance was to use deep learning. market forecasting is fraught with difficulties, and data scientists frequently run into issues while attempting to build a forecasting model. The unpredictability market for stocks and the link between investing psychology and market behaviour, provide two major concerns.

Literature Survey

H Maqsood suggested that [1] Forecasting the The stock market is a crucial element of corporate investment planning. Because Clients prefer equities over conventional assets since the high profit. since the information's nonlinear form and complicated economic regulations, great profit is frequently associated with considerable risk. since the regular volatility and quick changes on the stock exchange country's economic status, political position, and important events. The result is, investigating the results of some big events, especially global and local events, on various top stock businesses (country-by-country) remains an open study subject. In this research, look at four nations from the established, emerging, and underdeveloped economies: Pakistan, Turkey, Hong Kong, and the United States. We looked into the effects of various significant events.

We compute With information gathered from Twitter, analyse each of those instances' sentiment. The dataset, which was utilised to determine event emotions, consists of 11.42 million tweets. We used deep learning, support vector regression, and regression models for forecasting the stock market. The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are two metrics used to assess the data's accuracy. (RMSE)system's performance. The outcomes demonstrate that incorporating sentiment for these occurrences enhances performance.

W Long [2] stated that Because of the chaotic and non-stationary nature of samples, stock price modelling and prediction have been difficult goals for researchers and merchants. Because to the emergence of deep learning, purpose-built networks can now do feature learning. learning. more successfully. In this research, we provide a unique end-to-end model called multi-filters neural network (MFNN), which is made especially for price movement prediction and feature extraction from financial time series samples. Recurrent and convolutional filters are combined to form the multi-filter structure. neurons, allowing data from numerous feature areas and market perspectives to be acquired.

On the We employ Using the CSI 300 index of We simulate signal-based trading and employ our MFNN for challenging market prediction tasks on the Chinese stock market. Our network outperforms single-structure networks (convolutional, recurrent, and LSTM), statistical models, and traditional machine learning models in terms of accuracy, profitability, and
stability.. networks.

The primary goal this article's purpose is to develop a Cellular Automaton model as there are numerous types of stockbroker interact, and where the use and exchange of information between investors describe the complexity measured through the estimation of the Hurst exponent. This exponent represents an efficient or random market when it has a value equal to 0.5. Thanks to the various roposalsit might be determined in this investigation that a rational component must exist in the simulator generate an efficient behavior.

In this paper[4], we suggest and construct a fusion model to forecast financial market behaviour by merging the Hidden Markov Model (HMM), Artificial Neural Networks (ANN), and Genetic Algorithms (GA). The created tool can be utilised for thorough stock market analysisThe daily stock price changes are made. using ANN into separate sets of information that are then given to HMM. To optimise the HMM's starting settings, we look to GA. To find and locate related patterns in the historical data, the trained HMM is used. Measured are the cost variations between the matching days and the corresponding next day. To generate a forecast for the necessary following day, The price differences of related patterns are weighted to provide an average. For a number of securities in the IT sector, forecasts are compiled, and they are contrasted with a traditional forecasting technique.

A subgroup of learning algorithms known as SVMs, or support vector machines are distinguished by their capacity control of the decision function. usage of kernel functions, and solution sparsity. We study By predicting the weekly movement direction of the NIKKEI 225 index, this study used SVM to examine the reliability of financial movement direction. against assess SVM's predicting abilities, We contrast it versus Elman Backpropagation Neural Networks, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. The trials' findings show that SVM performs superior to other categorization techniques. Additionally, we provide a model for merging SVM with the other classification techniques. The combining model outperforms all other predicting models.

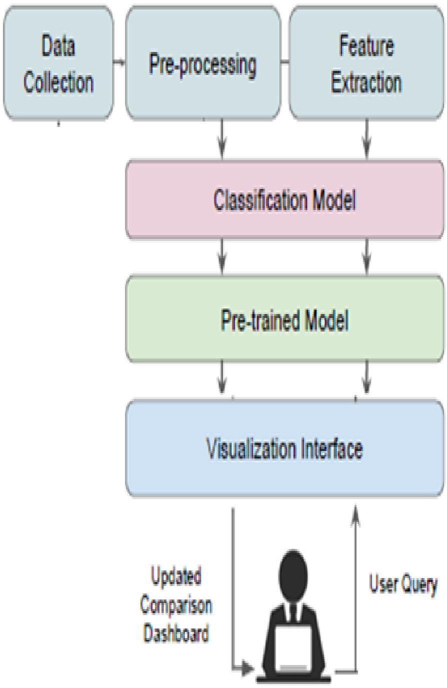

Figure 1 Proposed Architecture

Existing Model

Tsai et al. employed two different ensemble classifiers, heterogeneous and homogeneous, in light of using bagging and majority vote approaches. They also investigate the effectiveness of models using macroeconomic variables and financial statistics from the stock exchange in Taiwan. The outcomes demonstrated that Compared to solo classifiers, ensemble classifiers outperformed them in relation to investment returns and prediction accuracy.

Ballings and co.. compared AdaBoost, Random Forest, and kernel factory performance to single models including SVM, KNN, Logistic Regression, and ANN. They forecast European firm pricing for the coming year. The final findings demonstrated that none of the other models outperformed by Random Forest.

For the classification challenge, Basak et al. used the XGBoost and Random Forest algorithms to predict whether the stock would increase or drop based on prior data. Consequently, the findings, numerous companies' prediction performances have improved over those of their predecessors.

Overall, existing system modern methods of machine learning have frequently focused on macroeconomic or technical data without taking proper preprocessing techniques into account when detecting stock index or value change. No prominent results predicted.

Proposed Methodology

In this piece, we compare different models of machine learning anddeep learning techniques' a stock market forecast accuracy. Our models receive a lot of technical inputs. To assess the effect of preprocessing, the study uses two alternative ways for inputs: binary data and continuous data. The former uses stock trading data (open, close, high, and low values), whereas the latter uses a preprocessing stage to convert continuous data to binary data. Each technical indicator has its own unique qualities based on the fundamentals of the market. a distinct potential for upward or downward movement..

The effectiveness of the aforementioned models is evaluated using The most effective each parameter's tweaking strategy is then shown, in addition to categorization metrics for both methods. Four stock market groupings' historical Prior year's data are used for all experimental studies that are essential for investors..

As a consequence, deep In both methods, learning algorithms show a technical aptitude for estimating stock movement. particularly for continuous data.

The purpose of this research, which applies a variety deep learning and machine learning techniques, is to enhance the output of predicting stock group movement and trend. methodologies, is unique, in our opinion..

We identify When binary data is provided as input values to the model, we enter data with a recognised trend based on each prediction's future upward or downward movement predictors. trend feature's attribute.

Implimentation

Data Collection

This marks the beginning of the actual process of building collecting data and using a machine learning model. This step is critical since it will determine how effective the model is; the more and better data we collect, the more effectively our model will function..

There are a variety of techniques for gathering the data, including web scraping, manual interventions and etc.

Data Preparation

The data will be transformed. By removing specific columns and deleting missing data. First, we'll make a list of column names that we wish to maintain.

Then we drop or eliminate all columns except the ones we wish to keep.

Finally, we eliminate or remove the rows from the gathering of data that contain missing values.

Model Selection

We require two datasets when building machine learning model: one for training and one for testing. But now we just have one. So let's divide this in half with an 80:20 ratio. We'll additionally split the data frame into feature and label columns.

We used the sklearntrain_test_split function here. Then, split the dataset with it. Also, with test_size = 0.2, the split is 80% train dataset and 20% test dataset.

The random_state option starts a random number generator, which aids in splitting the dataset. Four datasets are returned by the method. They were labelled as train_x, train_y, test_x, and test_y. We can see the split of the dataset by looking at the style of the dataset.

Here, is the linear regression error term. The error term here which used to take into consideration the range ofboth x and y, where 0 represents the y-intercept and 1 represents the slope. To put linear regression into context, in order to train the model, x represents the input training dataset, and y represents the class labels contained in the input dataset. The machine learning algorithm's purpose is then to determine the optimal values for 1 and 0 to provide the bestfit regression line. To attain the best fit, The variation between the actual and and predicted values should be as small as possible, hence this minimization issue may be expressed as:

1 n Xni=1 (prediyi) 2

g = 1 n Xni=1 (prediyi) 2

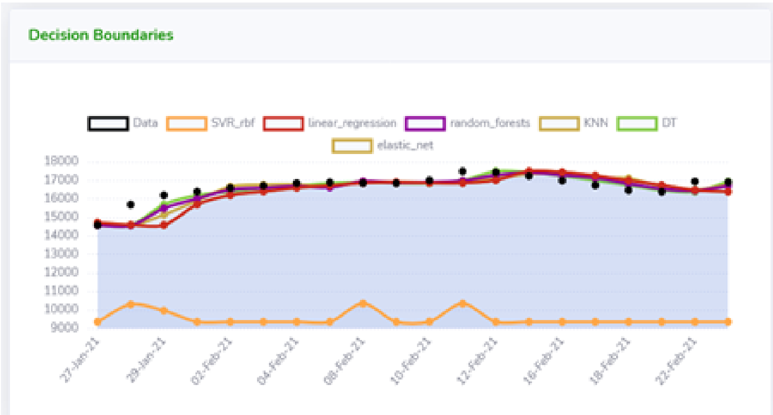

Figure 2 Decisoin Boundaries

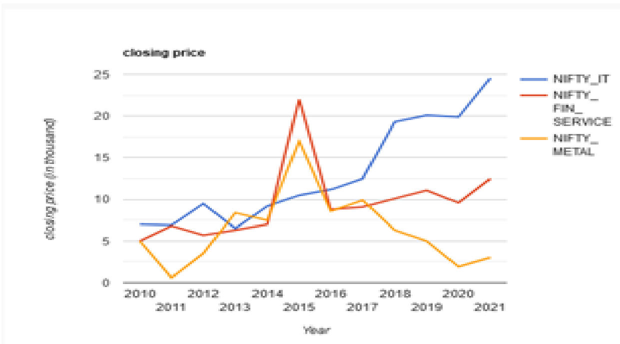

Figure 3 Result Chart

Conclusion

The goal The goal of this research was to forecast stock market activity using deep learning and machine learning. The Tehran Stock Exchange selected four stock market groupings, consisting of diverse financial sectors, non-metallic, and petroleumminerals, and basic metals. The dataset was built using ten years' worth of historical data and ten different technicalindicators. Additionally, RNN and LSTM were the two deep learning methods we employed as predictors.in addition to nine artificial intelligence (AI) models (Decision Tree, Adaboost, Random Forest, XGBoost, and SVC, NaveBayes, KNN, Logistic Regression, and ANN).For model input values, we assumed binary and continuous data in two different methods, and for evaluations, three categorization metrics. Our tests revealed that switching from continuous to binary input considerably enhanced model performance. LSTM and RNNare, in fact deep learning techniques were applied.

References

1. Technical Analysis of Financial Markets: A Comprehensive Guide to Trading Methods and Applications, J. J. Murphy. Penguin Books, 1999.
2. T. Turner, A Beginner's Guide To Day Trading Online, Simon & Schuster, New York, NY, USA, 2007.
3. H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, and K. Muhammad, ''A local and global event sentimentbased efficient stock exchange forecasting using deep learning,'' Int.
4. "Deep learning-based feature engineering for stock price movement prediction,'' W. Long, Z. Lu, and L. Cui, Knowledge-Based Systems, January 2019, vol. 164, pp. 163173.
5. J. B. Duarte, L. H. Talero Sarmiento and K. J. Sierra Juárez, "Evaluation of the Effect of Investor Psychology on an Artificial Market."
6. Ing, Ning Using Machine Learning to Trading Automation. Boston, MA, USA: Senior in Computer Science at Boston College, 2016.
7. M. R. Hassan, B. Nath, and M. Kirley, ''A HMM, ANN, and GA fusion model for the stock market forecasting,'' Expert Syst. Appl., vol. 33, no. 1, July 2007, pp. 171180.
8. "Forecasting stock market movement direction by support vector machine,'' W. Huang, Y. Nakamori, and S.-Y. Wang, Comput. Oper. Res., vol. 32, no. 10, Oct. 2005, pp. 25132522.
9. "Financial distress prediction using support vector machines: Ensemble vs. Individual,'' J. Sun and H. Li, Aug. 2012, Appl. Soft Comput., vol. 12, no. 8, pp. 22542265.
10. P. Ou and H. Wang, ''Prediction of stock market index movement using ten data mining approaches,'' Modern Appl. Sci., vol. 3, no. 12, Nov. 2009, pp. 2842.