iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.43403

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Stock Market Prediction & Analysis

Varun Gaur[1], Sharad Bharadwaj[2], Utsav Gaur[3], Sushant Gupta[4]
[1, 2, 3, 4]*Department of CSE, Krishna Engineering College, Ghaziabad*

Abstract: ***Stock trading is one of the most essential activities in the financial sector. The act of attempting to anticipate the future value of a stock or other financial instrument is known as stock market prediction. A financial exchange-traded instrument. This document illustrates how Machine Learning is used to predict a stock. The time series analysis or technical and fundamental analysis is used most stockbrokers use when deciding on a stock predictions. To forecast the outcome, the computer language is employed. Python is a stock market that uses machine learning. This paper is about We suggest a Machine Learning (ML) strategy that will be cost-effective. taught from publicly available stock data and intelligence and then applies what they've learned to make an accurate prediction. This work use machine learning in this setting. Support Vector Machine (SVM) is a technology for predicting Stock prices for large and small cap companies, as well as in the three different markets, using daily and weekly pricing Frequencies that are up to date.***

Keywords: Support Vector Machine, Stock Market, Machine Learning, Predictions

I. INTRODUCTION

In 2016, about 15% of the cash market trading value was supplied by over 2.2 million Hong Kong stock investors. HK$1.6 trillion is the total cash market trading turnover. Retail investors, in particular, made $240 billion in buy or sell investment choices in 2016 [1].

There are numerous investment decisions made in Hong Kong that involve big sums of money. Retail investors devote a significant amount of effort to researching investment alternatives. Professional financial counselling services are available to wealthy investors, but the charges are exorbitant for most regular investors. As a result, ordinary investors must educate themselves about the market and make their own decisions.

In modern society, this makes investing extremely stressful. Humans, unfortunately, are illogical creatures. Decisions are impacted by cognitive biases or emotional emotions without quantitative, data-driven models, resulting in avoidable losses. Even if investors are cautious, the majority lack the abilities to evaluate the massive amounts of data required to make sound decisions. To avoid traps, institutional investors use sophisticated models supported by technology, whereas retail investors do not have access to such tools and frequently fall behind the market.

A. Goals

Our application's ultimate purpose is to provide retail investors with a third-party investment tool that employs machine learning to assist them in navigating the volatile stock market. The goal of the project is to make cutting-edge machine learning technology more accessible to retail investors. There is no such thing as an exact prediction. As a result, the stock prices' upper and lower bounds will be provided to show the trading range that investors should be looking at. With the use of technology, this application acts as a supplemental quantitative tool for investors to observe the market in a new light. This project is separated into two parts: a research component and an application component, with the goal of providing retail investors with stock price predictions based on several machine learning models in a user-friendly manner.

B. Problem Statement

Everyone want to be rich in his life with low efforts and great advantages. Similarly, we want to look in our future with inner most desire as we do not want to take risks or we want to decrease risk factor. Stock market is a place where selling and purchasing can provide future aims of life (Kang Zhang et al, 2019). Now the question is that how we can get advantages from stock market? Or what are the steps that can give us stocks market predictions before taking yourself in risk zoon (Yue-gang Song et al, 2018).

C. Stock Market Prediction (SMP)

If stock market trend predicted then we can avoid wastage of money. SMP is a process of predicting future on the base of past data. Prediction decreases the risk level to investors and increases the confidence level for investment. If they predicted goals before reach then they can avoid loss of money. All these consideration work as SMP.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

The major goal of the proposed work is to determine the link between different existing time series algorithms, particularly ARIMA and Holt Winter, and stock prices. For the investments, a decent risk-free range of stock prices is investigated, and thus higher accuracy of the model can be seen. To produce a range of pricing for stock market consumers, a combination of two different time series analysis methods is used to produce distinct results for shares in the stock market. The key advantage of these models is that they are not sophisticated in nature and that they estimate values merely based on past stock prices for non-seasonal or seasonal securities. Some limitations in this experiment include the work that never considers other situations such as news about any new market strategy or a media release related to any company that may be affected by stock prices.

Multi-category news items are utilised as features in this paper to create a model for stock price trend prediction. The multi-category events are based on a feature word dictionary that has already been developed. We've also looked at the association between stock price fluctuations and particular multi-category news using both neural networks and SVM models. The predefined multi-category news events outperformed the baseline bag-of-words feature in predicting stock price trend in the experiments. According to this - study, short term prediction is superior to long term prediction. Machine Learning Technique for Predicting Share Prices: Publication Year: 2018 Jeevan B, Naresh E, Vijaya Kumar B P, and Prashanth Kambli are the authors. IEEE Journal Title: 2018 IEEE Summary: This study is mainly focused on the approach of forecasting the stock value on NSE data utilising various criteria such as current market price, price-earning ratio, base value, and other anonymous events using Long Short Term Memory (LSTM) and Recurrent Neural Networks (RNN). An RNN graph is used to compare the genuine data and the forecasted data to determine the model's efficiency. Machine learning to predict stock price as seen in the model, which is able to predict the stock price extremely near to the actual price by capturing comprehensive features and employing various methodologies. The model train will recognise and group all NSE data from the internet and offer input according to the user's preferences. This RNN-based architecture proven to be quite effective at projecting stock prices by altering the configuration accordingly, and it also uses a backpropagation approach to avoid data mixing while gathering and grouping data.

Stock Price Prediction by Machine Learning is used to forecast the future value of a stock using machine learning techniques such as LSTM. This machine-learning system is designed to provide the best stock future price prediction. In this proposed system, LSTM is capable of detecting changes in the behaviour of the stock price during the specified period. Propose [3] a stock price prediction normalisation based on machine learning. Yahoo Finance provided the data for the analysis. It has around 9 lakh records with the requisite Stock price and other pertinent information. Every day of the year, the data reflected the stock price at various time intervals. Date, symbol, open price, close price, low price, high price, and volume are all included. Only one company's data was considered in this study. All of the data was stored in a CSV file, which was read and turned into a data frame using Python's Pandas package. The data was normalised and separated into training and testing sets using Python's sklearn module. The experiment set was limited to 20% of the total dataset. The Regression-based Model and the LSTM are the subjects of this study. The Regression-based Model is used to predict continuous values from a set of autonomous values. Regression makes predictions by employing a given linear function to predict continuous values of the most important among them. The LSTM design can detect changes in trends that are visible in the results. The best model for the suggested methodology is determined to be LSTM. This demonstrates that the suggested system is capable of detecting data interrelationships. The changes that occur in the stock market may not always follow the same cycle or be in a regular pattern. The time of life will vary, and the trend's existence is dependent on the companies and sectors. This form of trend and cycle research will yield more profit for investors. We must employ networks like LSTM to analyse varied data because they rely on current information.

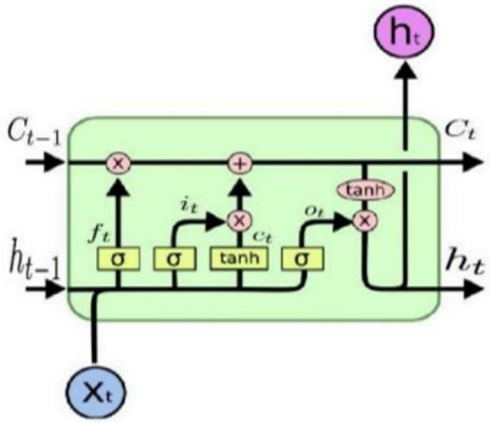
III. METHODOLOGY

The system shown here is made up of five modules: -

1) Create a Dataset
2) Initial preparation
3) Data division
4) Create and train Lstm, CNN, and LSTM+CNN hybrid models
5) Produce the expected outcome.

Price of open, high, low, close, adjusted closing price are obtained from a large dataset and provided as input to the models for training to pre-process the data using normalisation and one hot encoding techniques. Following that, the data is separated into two sets: training and testing, with an 80:20 ratio. The data is then utilised to train a model using three different approaches: LSTM, CNN, and a hybrid LSTM+CNNS method. Finally, the Root Mean Square Error is used to analyse all of these modules.

A. The LSTM Model in Action



A recurrent neural network is a type of long short-term memory. The output from the previous step is used as input in the next phase in RNN. It addressed the issue of RNN long-term dependency, in which the RNN cannot anticipate the word retain in long-term memory but can provide more accurate forecasts based on recent data. RNN does not provide an economical performance as the gap length grows. By default, LSTM will store the information for a long time. It's utilised for time-series data processing, prediction, and classification.

LSTM has a chain organisation that consists of four neural networks and several memory blocks known as cells.

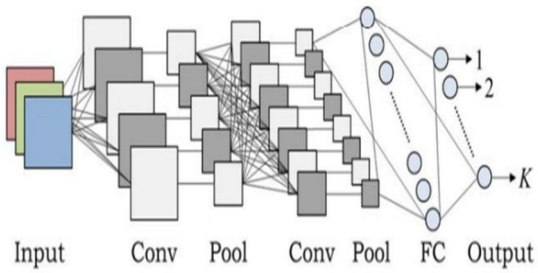
A memory cell is a novel structure in the LSTM. The memory cell decides what information to preserve and when reading, writing, and forgetting are permitted.

There are three main gates in a memory cell: o An input gate allows a new value to be entered into the memory cell. o Forget gate: the memory cell retains a value. o The memory cell's output gate value is used to compute the output.

B. CNN Model Operation

CNN model layer 1

1) MAX Pooling
2) Dropout
3) Flatten



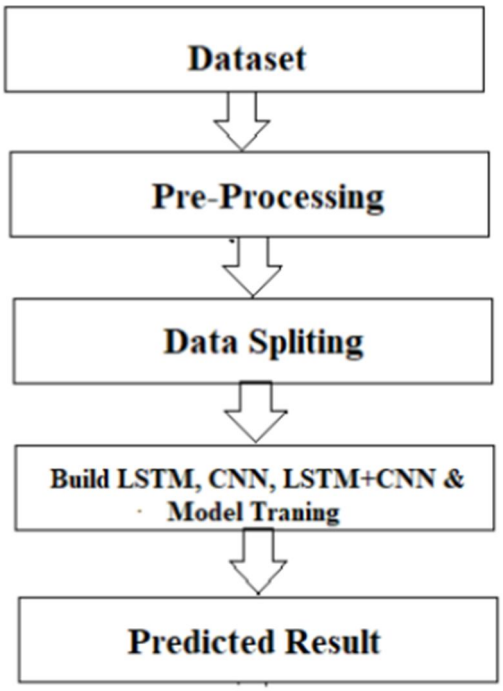
4) Dense
5) Activation
6) Convolution

a) Convolution: Extract the features from the input image using Convolution. It displays the results as a matrix.
b) MAX Pooling: The largest element from a corrected feature map is used in MAX polling.
c) Dropout: occurs when neurons are neglected at random during training.
d) Flatten: In a fully linked layer, flatten the feed output. It provides data in the form of a list.
e) Dense: A linear operation in which each input is weighted and related to each output. A nonlinear activation function followed.
f) Activation: It predicted the probability 0 and 1 using the sigmoid function.

C. LSTM + CNN Hybrid Approach

In the hybrid technique, Convolutional Neural Networks (CNNs) aid in the selection of appropriate solutions, while Long Short-Term Memory (LSTM) networks have demonstrated their ability to learn sequential data. Each method is said to yield better results. CNNs must apply filters to each input layer in order to obtain basic options, and CNNs have demonstrated improvements in computer vision, natural language processing, and other applications. CNN could be a useful tool for identifying features and improving prediction accuracy. The ability of LSTMs to learn data series by taking into account past outputs Multiple convolutional filters glide over the matrix to create a new feature map, and the filters come in a variety of sizes to generate various features. The Maxpooling layer is responsible for calculating the highest value as a filter's related feature. The Max-pooling layer's output vectors are fed into LSTM networks, which measure feature sequences' long-run dependencies. One of the advantages of LSTMs is their capacity to capture consecutive data while taking previous data into account. As inputs, this layer takes the dropout layer's output vectors. This layer has a fixed number of units or cells, and each cell's input is the output from the dropout layer. The outputs from LSTMs are merged and concatenated in one matrix before being transferred to a fully linked layer, which has the same number of units as the network. The fully connected layer converts the array into a single output in the range of 0 to 1, which is then categorised using the sigmoid function.

IV. FLOW CHART



V. CONCLUSION &FUTURE WORK

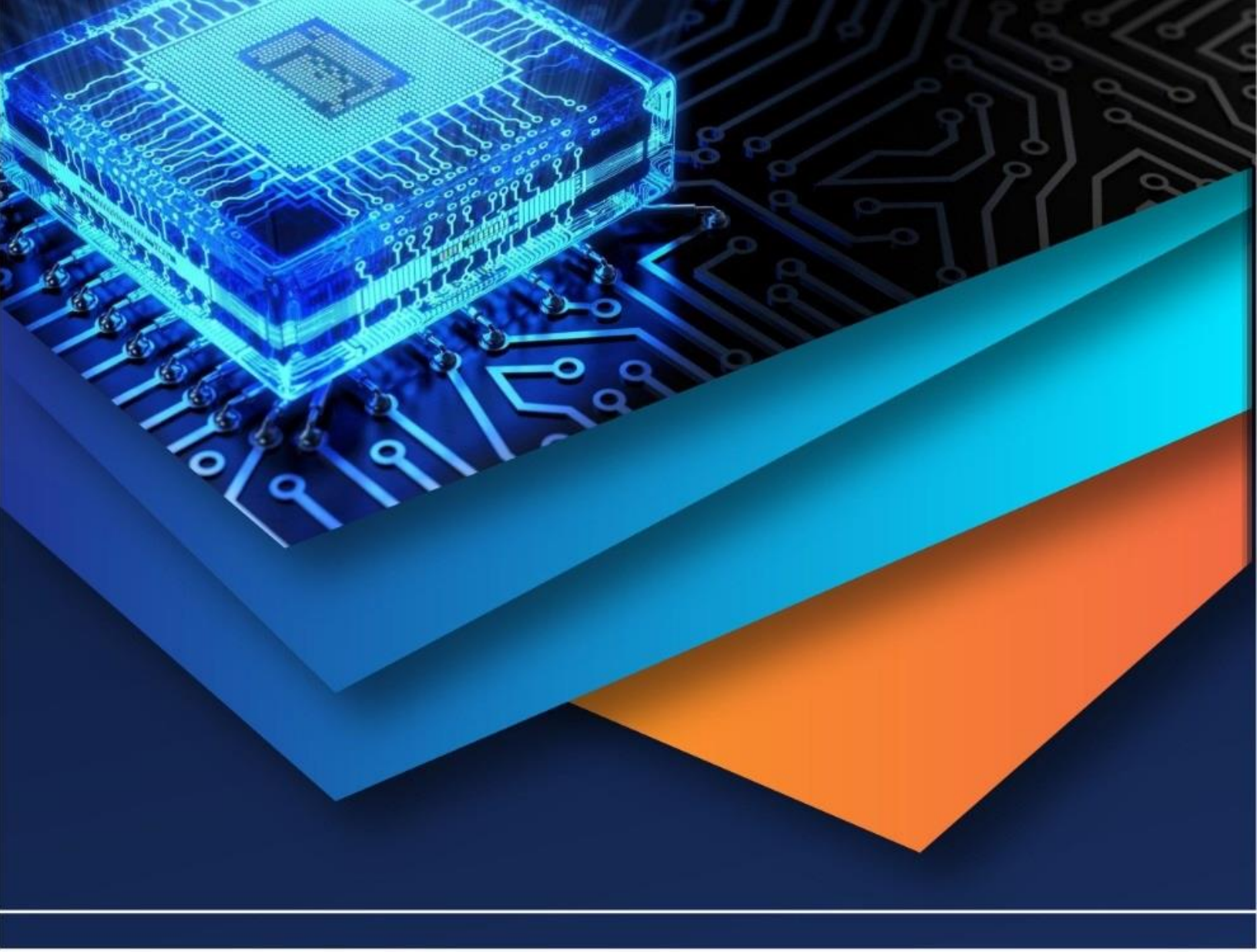
We will compare machine learning models such as the LSTM model, the CNN model, and the hybrid approach of the LSTM + CNN model in this paper. To anticipate the stock future value, we usually train the model using data from NSE listed firms. This demonstrates that the proposed method can distinguish between data interrelations. Furthermore, the results show that the hybrid strategy of LSTM+CNN model is capable of detecting changes in trends. The Hybrid strategy of LSTM+CNN is known as the best model for the projected method. It makes predictions based on the information available at the time. Even while the other two models, LSTM and CNN, are used in many other time-dependent data analysis applications, they do not outperform the Hybrid technique of LSTM+CNN architecture in this case. This is frequently due to the stock market's rapid fluctuations. Stock market fluctuations do not necessarily follow a predictable pattern or follow a continuous cycle. The existence of trends and the duration of their existence will vary depending on the firms and sectors. The understanding of these types of cycles and patterns can help investors make more money. To improve the accuracy of forecasted stock prices, we will incorporate more stock market data and compare more models in future work. In the future, the model can be taught with more varied and detailed data to improve accuracy. Other algorithms can be used with the presented to produce a new hybrid model.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] "Stock price prediction using LSTM, RNN and CNN-sliding window model - IEEE Conference Publication." https://ieeexplore.ieee.org/document/8126078 (accessed Dec. 27, 2019).
[2] J. Jagwani, M. Gupta, H. Sachdeva, and A. Singhal, "Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend," in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, Jun. 2018, pp. 462–467, doi: 10.1109/ICCONS.2018.8663035
[3] I. Parmar et al., "Stock Dec. 2018, pp. 574–576, doi: 10.1109/ICSCCC.2018.8703332.
[4] Y. Lei, K. Zhou, and Y. Liu, "Multi-Category Events Driven Stock Price Trends Prediction," in 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Nanjing, China, Nov. 2018, pp. 497–501, doi: 10.1109/CCIS.2018.8691392.
[5] B. Jeevan, E. Naresh, B. P. V. kumar, and P. Kambli, "Share Price Prediction using Machine Learning Technique," in 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C), Bangalore, India, Oct. 2018, pp. 1–4, doi: 10.1109/CIMCA.2018.8739647.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)