**Tesla Stock Price Forecasting with Random Forest Analysis**

Sub- Python Programming SPRING

2024 -23128

By- Sravya. Vemireddy

Pavan Ganesh Reddy. Yeruva

Afrida Mehanaz. Shaik

Kushwanth Reddy. Nomula

May 6 2024

**Tesla Stock Price Forecasting with Random Forest Analysis**

**Abstract** :

The project aims to provide retail investors to navigate through the stock market. This is achieved through the use of Random forest machine learning algorithm. A deeper understanding of Random Forest by taking OHLC (open-high-low-close) data for Tesla stock for 5year period. This will keep unrelated discussion at a minimum and try to focus on everything that is related to our understanding of Random Forest for this post. It would be helpful if you are familiar with concepts such as mean squared error and R-Squared statistics but it’s not necessary to understand this Concept.

Chapter 1

**Introduction**

The prediction of stock price is a task which tries to predict the future stock price value of a company stock and other factors which on an average will be traded financially. If our system will tell the future price value of stock successfully then, we can gain a very significant profit. The very first reason behind why we used this Random forest algorithm in our project is that it checks every condition based on decision tree which used to choose some of the important useful and the best attributes those affects most of the problem. Hence, we can predict the best possible output.

**Introduction of Machine Learning**

Machine learning (ML) is a way (and science) of making machines to learn old and new things which are not programmed by outside. It involves mostly mathematics similarly it also involves computer science. In it, the computer gets trained on a given dataset, and then to use the training data set in predicting the value of a new dataset. For example, if we give animal data to train computer by giving it 5000 pictures of cats and 5000 other pictures those are other than a cat, and predict each time by computer that given picture is cat or other animal. Now if we give the computer a new picture, then from the training dataset above, computer will be able to predict whether this new picture is cat or other. This Procedure of training and testing that is predicting uses some of the specialized algorithms. We give the training dataset to the algorithm, and then the algorithm uses these training datasets to give future values means predictions of a new test data.

**Introduction to Random Forest** :

Random forest algorithm is basically based on set of decision trees. Random Forest is constructed by using a method known as Bagging that work similar to the figure given below. for every decision trees, we always choose a random sub set of the training dataset & make it to be fitted into the decision tree . repeat this as much time as there are numbers of tree given for any forest. now, every model is up to some extent which is Predictive on a sub set of dataset. since this sub set has taken randomly , chances of mistake or error for every model must not be Correlated and these are some of the necessary things. It will take the average of all those different types of models, then i am actually merging the information very effectively which is contained in all of those models that is represented by the full dataset.

**Chapter 2**

**Objectives**:

The main aim is To design and develop a model to predict the price of stock of any company like tesla stock of a particular day and see if it actually similar to the output and analyse manually by what value of percentage its differs. The ultimate goal of this major Project is to serve the retail investors those who use Machine Learning for helping them to get informed about the very fast changing trend of the Stock markets . the Project have aim for introducing and democratizing those recent Machine Learning technology for the Retail investor for their use. 14 | P a g e We have experimented with some features in our dataset to predict the adjusted closing price for Tesla stock of a particular day. These predictions weren’t necessarily much good but these are only used for illustration purpose only , Since any prediction is not 100 percent accurate . therefore, we took the lower bound and the upper bound of Stock price and it would be display for illustrating the treding limit that investor would be taking care for.

**Chapter 3**

**LITERATURE SURVEY**:

I have always wondered why some companies are growing at such rapid growth as Facebook and Google while other companies like yahoo and Tata Nano are not growing at that rate despite good market strategy. After reading a few things one thing that stuck in my mind is that the success of any company depends largely on the price of the stock. That is why I have decided to establish a system that can predict good releases and if I can do that then it can bring great success to many companies. According to the paper of research titled "machine learning in stock price trend forecasting" published by y. dai no- y. zhang at stanford university, he uses the parameters such as pe ratios,px Volume, px EBITDA , ten day instability, fifty day average of moving , etc in predicting the next day Stock Price and long term Stock Price. Machine Learning algorithm which has been used in this study are logistic regression, gaussian discrimination analysis, quadratic discriminant analysis, and the SVM. Accuracy measure can be defined as the no. of total days in which the project has properly divided test dataset on the total no. of test day . with the small term model which predicts the price of the just next days , with the lowest accuracy, Quadratic Discriminant Analysis is the leading of all types, with a accuracy of 58.2%. With a long-term model predicting future price , as long the time, the accuracy of the SVM will be better. with a 44-day time windows , the accuracy of the svm model has reached 79.3 percentage .

Furthermore, this has been found that the accuracy will increase as the nimver of features increases. when all the 16 features was used , the accuracy of the model gets 79 percent , whereas it decreased to 64 percentage when only 8 feature was used , and 55 percentage when just a single features were used. our model would too investigates that how timing will effect the accuracies of Price estimates on various types. Since, the model has to reach a specific level in order for it to be important for users to act as a reference, it is important that we use our model to determine what appropriate parameters and model structure is our goal of predicting prices. The research paper "Predicting the movement of stock indexes and stocks using trend deterministic data preparation & Machine Learning strategies" ehich has been Written by j. patel ,s.shah, p. thakkar, noK. According to the international journal Expert Systems with Applications, the use of dataset to determine the Stock Price movements. they took a study with the use of ten signals of Technical indicator such as input, and then by using forecasting models to predict whether stocks would increase or decrease on the next ten day, Technical Analysis indicator which includes sma, ema, momentum, stochastic sk, stochastic sk, macd, rsi, etc . The forecasting model that they used including ann, svm, random forest, and naive bayesianmodel . These models emits ‘up’ or ‘down’ movements signal.

Studies has shown that the informal forests has achieved the greatest efficiency with an 83.56% accuracy of its results Researchers have shown that macroevolution can automatically produce neural network models and hyperparameters with higher performance compared to high end artificial model. On a study conducted in year 2017 , a major emergence of the discovery of image separation in neural networks. It starts with the majority of people with 1 simple simulated models, and then gradually transforms the Population by the removal of the bad design and then creating an another design by changing the parameter of the best design on every iterations. After the hundred of total hour of using the algorithms by great computation powers, many human model have achieve technical results in cifar data set. For every iterations , just a Simple modification which change one parameters were used, allowing search in a large search space. This paper has shown that it is possible to find good models using a lot of computing power in retrieving humans Machine Learning expert and to lay the foundations for democratic Learning using automl. The project problem has been set in the prediction of the next day's stocks Price. The ‘next day’ has been chosen as the time frame like a small term Price movement often rely heavily in the trend dynamics and Price patterns, While the long term Price movement depends in the basics for stocks (example: corporate managements skill, Revenue Model , Market Demand, Macroeconomic features, etc). RMSE and other schools are described as providing in-depth information on the performance of model prediction and financial-based comparisons between investor models On the research part , the System ha been designed as to be as robusts as Possible to the simplify models testings . each models could be known by the pairs of the design option & the inputs option , which specify the configuration of the design & the input input. It speeds up the processes of test of a different combination of planning and installation. The two keys feature is selected as the inputs . 1st is the long consistent lists of the certain Raw historic datasets such as the stock prices & the daily % changes. 20 | P a g e Selected specified size determines the historical size of the retrospective time from today When future’s pricing is forecast. By referring of the Principle of Technical Analysis , as soon the Stock Price reflect all the relevantinginformations, the technicals analysis will focuses on stock tradings patterns rather than the looking on the economicical and corporate fundamental.

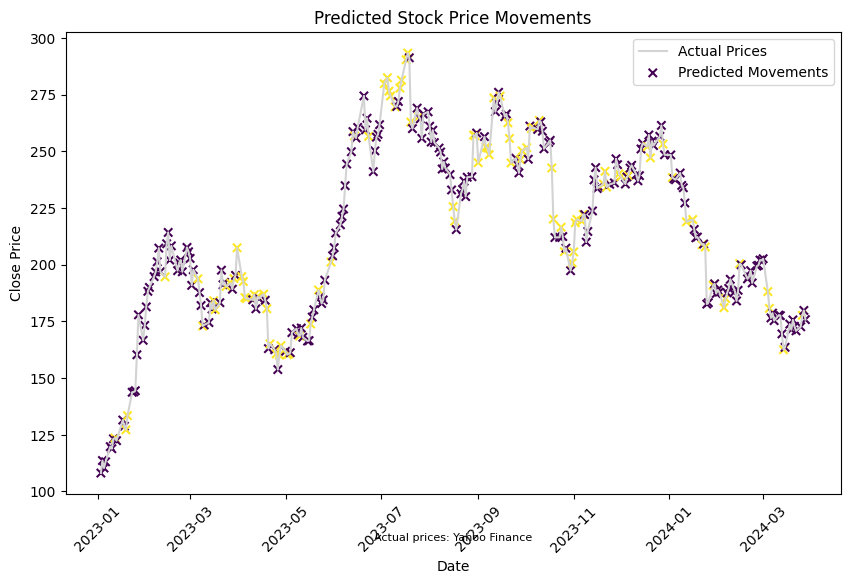
So, By determining the historical pricing period such as the inclusion of a training design , this can be pieces of the informations that is useful to find the trading Pattern & therefore prediction of the future’s Stock price trends. From the setted observation time, it has been assumed that predictable Price Movement pattern will occur in the specified historical period. The second input is the moving average of the figures. 1 of the most obvious ways to the market investor in the identification of market trends using the moving average . with a Robust System model, a different-moving average time can be use as an inclusion in stock price forecast model, for example, a set of 20, 30, 300, 600 day moving averages, which are widely used by investors. At 1 day's estimate, there will be one outflow units that is the stock on the next days. Predicted Stock Price would be the just another forecast, in predicting the Stock Price on 2nd day , the process will keep repeating until all the ten forecasts is made.

**Random Forest Classifier** :

The random forest to differentiate Machine Learning algorithm, create many Decision Tree & eventually combine those to provide a final and last Prediction . Many Decision making tree is based on many Conditional statement that help with good accuracies. With Such a large no. of Random Decision - making tree , We could avoid being overly balanced. A random forest is a way of learning together the division, retreating and other activities that work by building a pile of decision trees during training and classroom extraction which is a way of class (division) or predicting the meaning (return) of individual trees. Random decision-making forests correct the practice of over-pruning trees in their training set. The first algorithm of random forest forests was developed by Tin Kam Ho using a random subspace method, which, in its formulation, is a method of applying the "racial discrimination" method according to the categories proposed by Eugene Kleinberg The algorithm extension was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark. The extension incorporates Breiman's concept of "integration" with random selection of elements, introduced first by Ho and later independently by Amit and Geman to create a set of decision-making trees with a controlled variety.

**Chapter 4**

**Result:**



**Conclusion:**

In this study, our focus is on addressing the crucial real-world challenge of stock price prediction, particularly for companies like Tesla. We aim to leverage these predictions not only to facilitate the growth of the company itself but also to provide valuable insights for retail investors in the stock market. Through systematic efforts, we have developed a predictive system tailored specifically for forecasting Tesla's stock prices.

Our approach involved studying and evaluating two machine learning algorithms, with a particular emphasis on Linear Regression. By conducting experiments on the TSLA stock price database, we assessed the performance of our system across various metrics. The results indicate that our designed system, especially when utilizing the Linear Regression algorithm, achieves appreciable accuracy in predicting Tesla's stock prices.

Looking ahead, we envision extending our system and algorithms to predict stock prices for other companies beyond Tesla. Furthermore, there is potential for enhancing and automating stock market analysis and prediction processes, incorporating additional machine learning algorithms to improve accuracy and efficiency. Through ongoing research and development, our goal is to provide a robust and versatile solution for stock price analysis and prediction in the dynamic world of finance.

**Future Scopes**:

We target to take large numbers of feature in the future like companies future plans and checking their dependency on share capital. we do targets to grow sample size too for getting some more correct outputs to give more appropriate idea to the retail investors . As large dataset we will consider, the appropriate the sample to get Cross Validation in the criterias which will execute in our project. It would be a most necessary insrument for the future that has to come & would be recognised as a greatest invent in order to Solve the Stock Market’s Problem. The amount of person with stock market investment would increase gradually &therfore it would be a profit for the humanity. Investors and prediction makers may used this in their Validations , in Consultancy etc . people who have any doubt or only who wants to has a simple investment may took it’s Considerations in their accounts. Therefore, with current technology and in the advancement of future’s article, these problem will has a some good solution & timely using of the System will be beneficial. Hence, by the use of Machine Learning algorithms these designs are able to find an answer under the desired period of time .however in Future’s requirement on these department’s there is a large observation which could be taken in the considerations.

**References**

1. [**https://en.wikipedia.org/wiki/Random\_forest**](https://en.wikipedia.org/wiki/Random_forest)
2. [**https://www.tutorialspoint.com/machine\_learning\_with\_python/machine\_learning\_with\_python\_classification\_algorithms\_random\_forest.html**](https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_random_forest.html)
3. [**https://github.com/prnvg/Stock-Price-Prediction**](https://github.com/prnvg/Stock-Price-Prediction)
4. [**https://www.cse.ust.hk/~rossiter/fyp/103\_RO4\_Final\_201819.pdf**](https://www.cse.ust.hk/~rossiter/fyp/103_RO4_Final_201819.pdf)
5. [**https://www.slideshare.net/Tech\_MX/linear-regression-14155467**](https://www.slideshare.net/Tech_MX/linear-regression-14155467)
6. [**https://www.google.com/search?q=kaggle+datasets&oq=kaggle&aqs=chrome.2.69i60j69i57j 0i395i433j0i433j0j69i60l3.5155j1j7&sourceid=chrome&ie=UTF-8**](https://www.google.com/search?q=kaggle+datasets&oq=kaggle&aqs=chrome.2.69i60j69i57j%200i395i433j0i433j0j69i60l3.5155j1j7&sourceid=chrome&ie=UTF-8)