Stock Price Analysis and Prediction of Apple Company

Abstract—**The Stock market process is full of uncertainty, expectations and is affected by many factors. Technical analysis is done using historical data of stock prices of different companies by applying machine learning algorithms.Here we collect the data-set from quandl apple stock price 2013-2017. Then the output which will get after applying algorithms will analyze and the stock values are analyzed. The learned application can then be used to make future predictions about stock values. This paper investigates various techniques for the stock market prediction using LSTM (Long Short Time Memory). It can be shown that this method is able to predict and forecast the the stock price, and it can be used on any real-time data set.**

Index Terms—**stock, price, analysis, close, dataset, prediction**

I. Introduction

A stock market is a public market where you can buy and sell shares for publicly listed companies. Stock markets serve as an indicator of the state of the economy.It is a widely used source for people to invest money in companies with high growth potential. Stock Price Prediction using machine learning helps you discover the future value of company stock and other financial assets traded on an exchange. Technical analysis inspects past data and volumes of stock prices while fundamental analysis not only considers stock statistics but also evaluates industry's performance, political events, and economic circumstances (Patel et al., 2015; Milosevic, 2016). Fundamental analysis is more realistic because it evaluates the market in a broader scope[1].Based on 26 Wall Street analysts offering 12 month price targets for Apple in the last 3 months.The average price target is dollar 162.12 with high and low forecast of dollar 185.0 and 90.00.The average price target represents a 10.96 percent change from the last price of dollar 146.11[2].

II. Literature Review

In this section we will discuss about some recent analysis and prediction on stock prediction. Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange.Currently, most of the study is based on forecasting of stock price trends using neural networks based on ARIMA. . The ARIMA model was selected because of its wide suitability of the model. The parameters selected for the ARIMA model have been done using Akaike information criterion (AIC) and Akaike information criterion corrected (AICc).An additional reason in support of the ARIMA model is its usefulness in predicting or forecasting the next day, mainly in the study along with experiments a specific stock market associated toward stock price prediction. However, This study has highlighted ARIMA model to predict the Apple stock price in the short term. After collecting enough real data to create stock market data, an ARIMA model is implemented in the data set used to improve the short term forecast. The application of the model in the case of bank action data has allowed verifying its accuracy and demonstrating its presentation capabilities[3].

In this research[4], create a novel framework for stock prediction (dynABE). dynABE explores domain-specific areas and its test by three cobalt-related companies, Compare the performance of dynABE to three baseline models commonly used for stock prediction, namely support vector machine, neural network, and random forest and it achieves the best-case miss classification error of 31.12 and an annualized absolute return of 359.55 percent with zero maximum drawdown. Four main contributions of this research to current works: the exploration of domainspecific information for high-frequency predictions; the establishment of an effective, first-level ensemble learning framework; the proposal of "advisors" for a second-level ensemble; and an online update strategy for dynamic flexibility. For this research, historical data from the entire year of 2015 is used for training Data of 2016 and the first half of 2017 is used as the validation set and choose three cobalt company. In features set makeup and advisor section illustrates how they choose the advisors and the feature set for the critical metal case study. After that, they propose a general standard for data selection if one wishes to implement dynABE.In methodology it gave importance ensemble learning framework for each advisor and the online update strategy for combining the advisors. This paper combines the results of three rankings with majority vote and adopt a simple voting strategy: for any feature, top 20 percent as the selected feature set. Its also build five base model - Linear regression, Logistic regression, support vector machine, extreme gradient boosting, rotation forest. In order to reduce the model variance, its further implement bootstrap aggregation. Lastly use miss classification errors as the evaluation metric, which is the percentage of the trading days for which the models predict the wrong stock trend. This paper shows online update tuning, weight history, accuracy history, trading strategy returns for Jinchuan, Zijin Mining, Sumitomo. It's also shows the comparison of stacking and online update errors, comparing baseline model and dynABE. For future work this paper want to develop a tuning strategy for the update frequency, decay rate, and diversity bias.

Stock price data is not normal and it has some characteristics such as skewness, kurtosis, fat tail and non linearity for this the main goal of this article[5] is to predict stock price indices using (ANN) and train it with some new metaheuristic algorithms such as (SSO) and (BA). These two are successful and brilliant results in various fields and researches such as prediction of the stock price and prediction of interest rate; on the other hand, they have some properties including their approximate and usually non-deterministic nature and also, they are not problem-specific and flexible too Then, (GA) as a heuristic algorithm for feature selection and choosing the best and most related indicators. By using (GA) speed rate of calculation is increased and also the network will be prevented getting into local minima or maxima trap. This paper also gives importance in Neural Network which is based on learning which means each time it tries to reduce their error based on trial and error. This paper also shows some pseudocode of (GA-ANN), (BA) and they also used (BA), (SSA) for operating in many ways. Technical analysis, fundamental analysis and statistical methods, Time series method are used for stock price prediction. Heuristic algorithms are another set of methods being used for prediction. The other methods are metaheuristic algorithms comparing with exact methods, there is no guarantee that metaheuristics can find global optimum of an optimization problem also used Hybrid metaheuristic ANN for stock price predict. To evaluate the performance of the hybrid algorithm, the obtained results are compared with results of ARIMA as a time series model to predict the stock price. This paper also added some limitation of ARIMA, BPNN, CART, GP ,GRNN, Hierarchical, HMM, KNN, LR, LSTM, MLP, PSO, RBF, RF, RNN, SOM, SVM, SVR, ANN limitation. After this, this paper used some loss functions such as mean absolute error (MAE) as error evaluation criteria. On the other hand, used some time series models forecasting like ARMA and ARIMA for prediction of stock price. Finally, compared the results with each other means ANN-Metaheuristic algorithms and time series models. The statistical population of research have five most important and international indices which were SP500, DAX, FTSE100, Nasdaq and DJI.

Stock market prediction is a challenging task as it requires deep insights for extraction of news events, analysis of historic data, and impact of news events on stock price trends.For this paper[6], it highlights the significance of deep neural network based prediction techniques to capture the hidden relationship between textual and numerical data.The Istanbul Stock Exchange stock prices volatility is forecasted using ANN and SVM prediction models.SVM and Least Squares SVM (LS-SVM). It is observed that each of the algorithms has its own limitations.Here also use the kNN algorithm which is a complete training dataset is used to predict class for every test example. This makes KNN inefficient in terms of time and memory. It summarizes the techniques that only consider numerical financial data for stock prediction. Then it discusses the feature extraction from textual data. It also summarizes the prediction algorithms that exploit the combination of numerical and textual data for prediction.

This paper[7] studies the possibilities of making prediction of stock market prices using historical data and machine learning algorithms.An accuracy analysis was also conducted to determine how useful can these types of supervised machine learning algorithms could be in the financial field. Stock market data of the Apple Inc. using random trees andmultilayer perceptron algorithms to perform the predictions of closing prices.The final goal is to manage the extraction of quantitative data with relevant information from the stock market. This paper have used historical price data of the Open, Close, High, Low and Volume of the last 250 trading sessions. Regarding to the machine learning algorithms, they have used the following WEKA packages. The attribute Close is the one to be forecasted and compared to its real data, so that the accuracy of the algorithms can be tested and measured for errors and over fitting. Both executions of WEKA's algorithms fit the actual historical Price data (Correlation factor of 0.9998 for the first one and 0.9976 for the second with a maximum adjustment possible of 1.0) very closely. This transparency of information would facilitate the application machine learning algorithms and artificial intelligence to Ecuadorian financial securities to further research their application in other markets, which may help reduce costs of market inefficiencies.

Stock Prediction change rapidly and unpredictably that makes prediction quite rigid. This research work[8] is organized in eight sections. Fundamental and technical analyses are two basic approaches used for stock trend prediction. Shallow learning composition layers are few like (SVM) and (ANN) where deep learning technique contains hidden layers like (CNN), event extraction. This research study emphasizes the domain of stock prediction using machine learning so the other relevant terms are "machine learning", "artificial intelligence", "artificial neural network", "deep learning", "stock price", "news headlines", "event extraction", "text mining", "sentiment analysis", "sentiment lexicon", and "time series analysis". General methodology in first term is data collection, second one is performing data processing. Then they observed noise can be caused by human error or machine error while outliers can be caused by experimental error. Determine a difficulty that text mining for extracting fact is unstructured from the data but it can perform by using sentiment analysis. Machine learning algorithms they prefer (NB), (KNN), (SVM), (ANN), (DNN), (CNN), (RNN), (LSTM). Used Accuracy, Precision, Recall, F1- score, (MSE), (RMSE), (MAE) for evaluate also added time series consists of four components: (T), (S), (C) (I). A financial indicators can be used directly in prediction models as a dependent or independent variable .Moreover new features are also derived from the existing one such as gain . This paper presents an extensive study of stock trend prediction using news and stock prices also has four major contribution forecasting using time series data and textual data, pre processing and feature extraction in textual and numerical data, techniques for stock trend prediction using numerical and textual features.

The study[9] of Hamilton a plethora of studies have analyzed the interrelation between macroeconomic activity and oil price changes, most of which demonstrated a negative correlation. Moreover, a number of researchers have examined the role of crude oil prices in monetary policy and impacts of oil prices on exchange rates. However, there are comparably fewer studies on the relationship between oil prices and stock markets. According to the analyses, oil price shocks influence various industries' stock prices differently and the relationships between oil price shocks and financial markets are, for many countries, complex and ambiguous. A commonly held view is that an upward trend in oil price is beneficial for upstream oil companies, whose cash flows are directly related to the difference between oil price and crude oil lifting cost, yet has an adverse effect on downstream companies, including refining crude oil and marketing products, and many other industries.

A large literature has identified a number of predictors that are useful to predict future stock returns. Those include dividend yield and dividend-price ratio, price-earnings ratio, short interest rate ,term and default spreads and consumption-wealth ratio . Besides predictors, forecasting techniques also play an important role in determining forecast accuracy. According to Mallikarjuna and Rao[10], traditional regression techniques generally outperform others including artificial intelligence and frequency domain mod-els in providing accurate forecasts. In terms of stock volatility, academic researchers used to make the forecasts by tra-ditional GARCH models using indicators based on the past behavior of stock price and volatility. More recent studies become aware of issues such as parametric assumptions, leverage and asymmetric effects, and power transformations and long memory. In this paper, we introduce GARCH models for volatility forecasting because we aim to test for the instability of the volatility process, which is primarily built upon those mod-eling techniques.

III. Methodology

As it is a prediction of apple stock price we will use LSTM model as we have to make time series prediction model.Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Link: https://scikit-learn.org/stable/modules/neural_networks_supervised.html

IV. Data Set

Apple stock price from September 3,2013 to December 28, 2017. Link: https://www.quandl.com/data/EOD/AAPL The data will be divided into two parts train and test data respectively. The train dataset is for model estimation and the test is for testing our predicted model. The variables are weekly date as a factor variable and Adj.close price of the stock as a numeric variable.

```
RangeIndex: 1090 entries, 0 to 1089
Data columns (total 11 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Date        1090 non-null   object
 1   Open        1090 non-null   float64
 2   High        1090 non-null   float64
 3   Low         1090 non-null   float64
 4   Close       1090 non-null   float64
 5   Volume      1090 non-null   int64
 6   Adj_Open    1090 non-null   float64
 7   Adj_High    1090 non-null   float64
 8   Adj_Low     1090 non-null   float64
 9   Adj_Close   1090 non-null   float64
 10  Adj_Volume  1090 non-null   int64
dtypes: float64(8), int64(2), object(1)
```

Fig. 1. Dataset Features

Fig.1 shows the features or columns of our dataset. In total, there are 11 features in our dataset, where 8 of them are float, two integers, and one object type feature. There are total of 1090 data in our dataset.

	Open	High	Low	Close	Volume	Adj_Open	Adj_High	Adj_Low	Adj_Close	Adj_Volume
count	1090.000000	1090.000000	1090.000000	1090.000000	1.090000e+03	1090.000000	1090.000000	1090.000000	1090.000000	1.090000e+03
mean	194.564537	196.114735	193.016676	194.613130	3.605114e+07	26.283677	26.493692	26.068855	26.289305	1.931334e+08
std	160.562848	161.812898	159.407381	160.647536	2.194473e+07	6.593449	6.619745	6.557756	6.595409	1.112891e+08
min	90.000000	90.700000	89.470000	90.280000	5.704900e+06	14.144639	14.515653	14.121273	14.212842	4.590369e+07
25%	108.042500	108.942500	106.972500	108.007500	2.120714e+07	22.099713	22.362631	21.879019	22.130410	1.133450e+08
50%	121.585000	122.504950	120.640000	121.755000	3.248325e+07	25.699500	25.972960	25.430346	25.707787	1.681963e+08
75%	156.905000	157.809625	155.744975	156.515000	4.673778e+07	29.235805	29.448022	29.041285	29.269451	2.394642e+08
max	649.900000	651.260000	644.470000	647.350000	1.895606e+08	41.991703	42.492889	41.931753	42.305844	1.065523e+09

Fig. 2. Different values of our dataset

Fig.2 shows the mean average, standard deviation, 1st,2nd, 3rd quartile, and minimum and maximum value of our dataset, which is very important for many analysis tasks.

V. Result and Analysis

In this section, we will show our data analysis result and discuss it. First of all, we have shown the change in the price of stick over time. Fig.3 shows the overtime change of the price of all the 10 price features.

Secondly, we have found out the moving average of various stokes. Fig.4 shows the moving average for 10, 20, and 50 days.

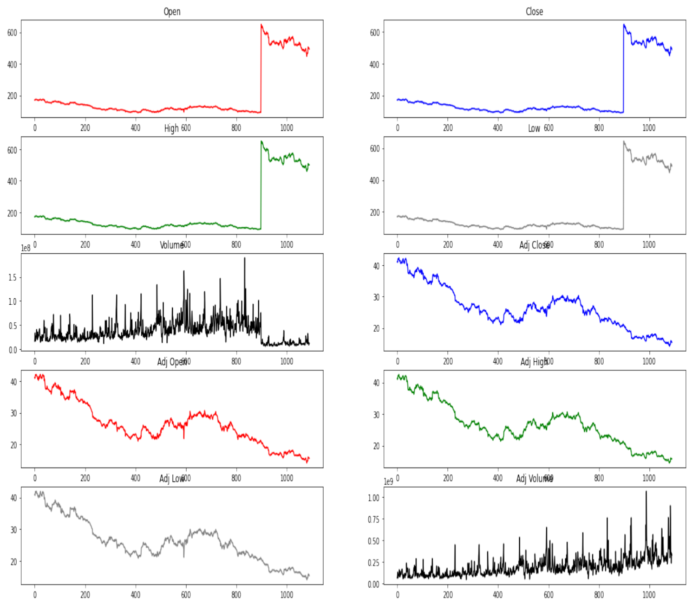

Fig. 3. Overtime Change in stock price

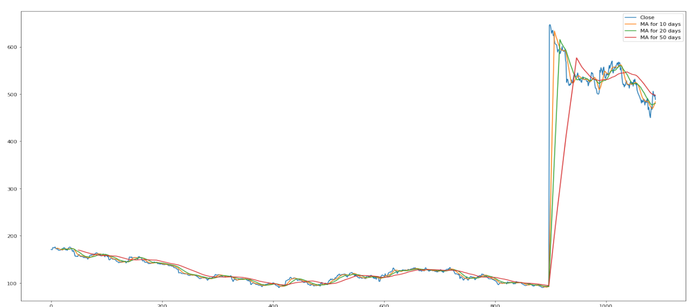

Fig. 4. Moving Average of the stokes

Thirdly, we showed the daily return of the stock on average. Let's go a little further now that we've completed some preliminary research. We're going to look at the stock's risk presently. To do so, we'll need to look at the stock's daily fluctuations rather than its absolute value. Let's get started by retrieving the daily returns for the Apple stock. Fig 5 shows the daily return of the apple stock.

Fig.6 shows the same daily return of Fig.5 by using histogram and kde plotting on the same figure.

Fourthly, we have tried to analyze how much value we put at risk by investing in a particular stock. We may measure risk in various ways, but one of the most fundamental is to compare the expected return to the standard deviation of the daily returns, which we can do using the data we've acquired on daily percentage returns. Fig.7 shows the comparison between risk and expected return.

Finally, we have tried to predict the closing price stock price of Apple stock. We can also predict other stock prices,

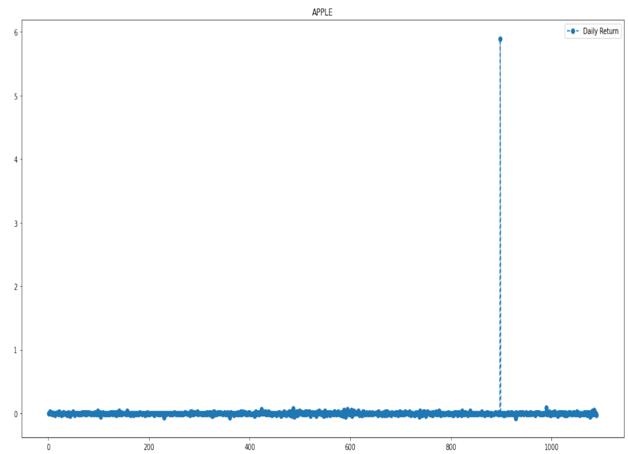

Fig. 5. Daily Return for apple stock

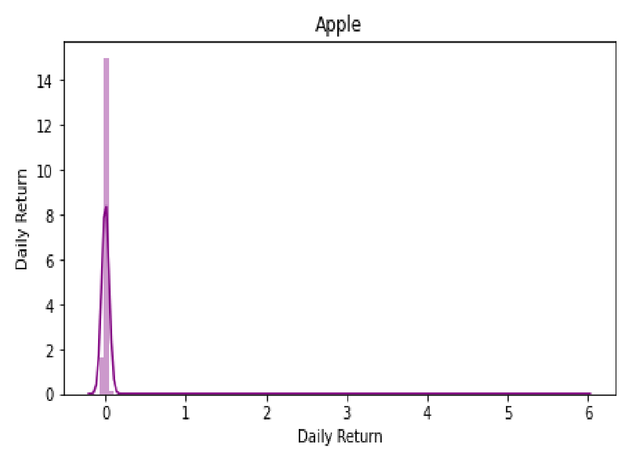

Fig. 6. Histogram and kde plot of daily return for apple stock

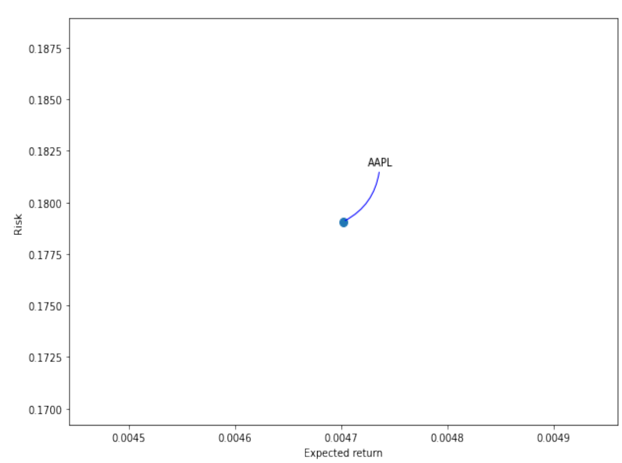

Fig. 7. Risk vs Expected Return

such as Open, but we have to start predicting with Close. Before training, we have split our dataset into train and validation, which is 95% data for training and 5% for validation. For training, we have used the LSTM model. Adam optimizer and MSE loss function used as a parameter. After creating the model, we have trained with 50 epochs. We have tried to find the RMSE value of our trained model, and it was 7.85. Fig.8 shows the training value, validation value, and prediction value. In the last portion of the graph, it can be clearly seen that the validation data and prediction data are pretty similar, which means our prediction was successful. For more clear visualization, Fig.9 shows the comparison between the actual value and the prediction value. We can see that the prediction value is pretty close to the actual value, which means our model trained very well.

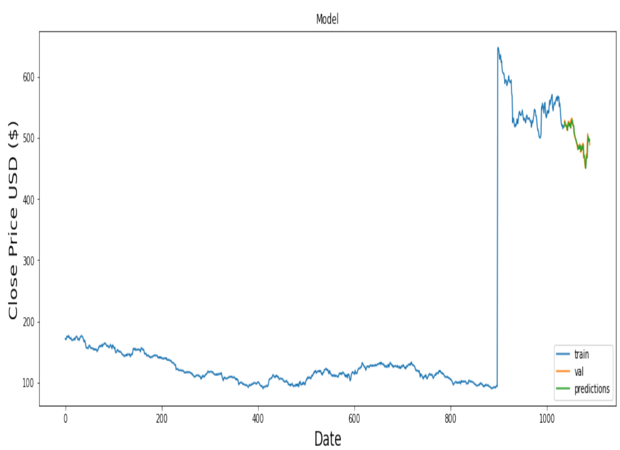

Fig. 8. Train vs Validation vs Prediction

	Close	predictions
1036	524.991	517.552734
1037	528.160	522.729309
1038	520.634	525.871094
1039	520.010	519.945312
1040	519.048	518.692261
1041	520.560	517.618103
1042	512.492	518.742004
1043	520.920	512.157166
1044	525.449	518.347351
1045	526.750	522.792419
1046	520.030	524.482849
1047	522.702	519.095215
1048	524.896	520.668640
1049	516.678	522.613892
1050	529.876	516.059509

Fig. 9. Close actual value VS Close Prediction value

CONCLUSION

REFERENCES

[1] Milosevic (2016) Milosevic N. Equity forecast: predicting long term stock price movement using machine learning. https://arxiv.org/abs/1603.00751.

[2] Apple Stock Forecast Price Targets https://www.tipranks.com/stocks/aapl/forecast

[3] Apple Stock Price Prediction using ARIMA Model http://www.ijarset.com/upload/2019/december/2-awajionyi04.pdf

[4] Dynamic Advisor-Based Ensemble (dynABE): Case study in stock trend prediction of critical metal companies.Zhengyang Dong PLoS One. 2019; 14(2): e0212487. Published online 2019 Feb 22. doi:10.1371/journal.pone.0212487 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6386270/

[5] Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models. Milad Shahvaroughi Farahani, Seyed Hossein Razavi Hajiagha https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8070984/

[6] Shazia Usmani, Jawwad A. Shamsi PeerJ Comput Sci. 2021; 7: e490. Published online 2021 May 4. doi: 10.7717/peerj-cs.490 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8114814/ref76

[7] Stock Market Data Prediction Using Machine Learning Techniques: Proceedings of ICITS 2019. February 2019 DOI:10.1007/978-3-030-11890-752Author : EdgarP.T orres, MyriamHernandezAlvarez, SangGuunY oo. https://www.researchgate.net/publication/330677469StockMarketDataP redictionU singMachineLearningT echniquesP roceedingsof

[8] News sensitive stock market prediction: literature review and suggestions. Shazia Usmani, Jawwad A. Shamsi PeerJ Comput Sci. 2021; 7: e490. Published online 2021 May 4. doi:10.7717/peerj-cs.490 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8114814/sec1title

[9] Stock market returns and oil price shocks: A CoVaR analysis based on dynamic vine copula models. Julia Kielmann, Hans Manner,corresponding and Aleksey Min https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8169420/

[10] COVID-19 and instability of stock market performance: evidence from the U.S. Hui Hong, Zhicun Bian, and Chien-Chiang Lee https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7903039/