**STOCK PREDICTION**



**UNIVERSITY OF ENGINEERING**

**&**

**MANAGEMENT, JAIPUR**

STOCK PREDICTION

Submitted in the partial fulfillment of the degree of

**BACHELOR OF TECHNOLOGY**

In

**COMPUTER SCIENCE & ENGINEERING**

Under

**UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR**

BY

**Nitin Yadav**

**University Roll no: 12019002001030**

**University Registration no:** 204201900200030

**Md Mahfuz Alam**

**University Roll no: 12019002001019**

UNDER THE GUIDANCE OF

**MS JYOTI KHANDELWAL**

COMPUTER SCIENCE & ENGINEERING



UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

**Approval Certificate**

This is to certify that the project report entitled “**Stock Prediction**” submitted by **Nitin Yadav** (Roll:**12019002001030**) and **Md Mahfuz Alam (12019002001019)**in partial fulfillment of the requirements of the degree of **Bachelor of Technology** in **Computer Science & Engineering** from University **of Engineering and Management, Jaipur** was carried out in a systematic and procedural manner to the best of our knowledge. It is a bona fide work of the candidate and was carried out under our supervision and guidance during the academic session of 2021-2022.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Prof. Jyoti Khandelwal**

Project Guide, Assistant Professor (CSE)

UEM, JAIPUR

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Prof. Mrinal Kanti Sarkar Prof. A Mukherjee**

HOD (CSE) Dean

UEM, JAIPUR UEM, JAIPUR

**ACKNOWLEDGEMENT**

The endless thanks goes to Lord Almighty for all the blessings he has showered onto me, which has enabled me to write this last note in my research work. During the period of my research, as in the rest of my life, I have been blessed by Almighty with some extraordinary people who have spun a web of support around me. Words can never be enough in expressing how grateful I am to those incredible people in my life who made this thesis possible. I would like an attempt to thank them for making my time during my research in the Institute a period I will treasure. I am deeply indebted to my research supervisor, Professor Jyoti Khandelwal guide me such an interesting thesis topic. Each meeting with her added in valuable aspects to the implementation and broadened my perspective. She has guided me with his invaluable suggestions, lightened up the way in my darkest times and encouraged me a lot in the academic life.

Nitin Yadav

**ABSTRACT**

It has never been easy to invest in a set of assets, the abnormally of financial market does not allow simple models to predict

future asset values with higher accuracy. Machine learning, which consist of making computers perform tasks that normally

requiring human intelligence is currently the dominant trend in scientific research. This article aims to build a model using

Recurrent Neural Networks (RNN) and especially Long-Short Term Memory model (LSTM) to predict future stock market

values. The main objective of this paper is to see in which precision a Machine learning algorithm can predict and how much the

epochs can improve our model.

Table of Contents

1. Introduction
2. Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)
3. Methodology and Data
4. Results and Discussion
5. Conclusion
6. Bibliography

**INTRODUCTION**

Several studies have been the subject of using machine learning in the quantitative financial, predicting prices of managing and constricting entire portfolio of assets, as well as, investment process, and many other operations can be covered by machine learning algorithms. In general machine learning is a term used for all algorithm’s methods using computers to reveal patterns based only on data and not using any programming instructions. For quantitative finance and specially assets selections several models supply a large number of methods that can be used with

machine learning to forecast future assets value. This type of models offers a mechanism that combine weak sources of information and make it a strange tool that can be used efficiently. Recently, the combination of statistics and

learning models have polished several machine learning algorithms, such as acritical neural networks, gradient boosted regression trees, support vector machines and, random forecast. These algorithms can reveal complex

patterns characterized by non-linearity as well as some relations that are difficult to detect with linear algorithms. These algorithms also prove more effectiveness and multi collinearity than the linear regressions ones. A large number of studies is currently active on the subject of machine learning methods used in finance, some studies used tree-based models to predict portfolio returns, others used deep learning in the production of future values of financial assets. Also, some authors over viewed the forecasting of returns using of ADaBoost algorithm. Others proceeds to forecast stock returns using unique decision-making model for day trading investments on the stock market the model developed by the authors use the support vector machine (SVM) method, and the mean-variance (MV) method for portfolio selection . Another paper conversed deep learning models for smart indexing. Also, some study has covered a large number of trends and Applications of Machine Learning in Quantitative Finance, the literature review covered by this paper consist of return forecasting portfolio construction, ethics, fraud detection, decision making, language processing and sentiment analysis. These models don't depend one long term memory (passed sequences of data), in this regard a class of machine learning algorithms based on Recurrent Neural Network prove to be very useful in financial market price prediction and forecasting. A paper has compares

the accuracy of auto regressive integrated moving average ARIMA and LSTM, as illustrative techniques when forecasting time series data. These techniques were executed on a set of financial data and the results showed that LSTM was far more superior to ARIMA. Our paper aim to use ML algorithm based on LSTM RNN to forecast the adjusted closing prices for a portfolio of assets, the main objective here is to obtain the most accurate trained algorithm, to predict future values for our portfolio.

**Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM)**

Long Short-Term Memory (LSTM) is one of many types of Recurrent Neural Network RNN, it’s also capable of

catching data from past stages and use it for future predictions . In general, an Artificial Neural Network (ANN)

consists of three layers:

1) Input layer,

2) Hidden layers,

3) Output layer.

In a NN that only contains one hidden layer the number of nodes in the input layer always depend on the dimension of the data, the nodes of the input layer connect to the hidden layer via links called ‘synapses’. The relation between every two nodes from (input to the hidden layer), has a coefficient called weight, which is the

decision maker for signals. The process of learning is naturally a continues adjustment of weights, after completing the process of learning, the Artificial NN will have optimal weights for each synapses. The hidden layer nodes apply a sigmoid or tangent hyperbolic (tanh) function on the sum of weights coming from

the input layer which is called the activation function, this transformation will generate values, with a minimized error rate between the train and test data using the SoftMax function. The values obtained after this transformation constitute the output layer of our NN, these value may not be the best output, in this case a back propagation process will be applied to target the optimal value of error, the back

propagation process connect the output layer to the hidden layer, sending a signal conforming the best weight with the optimal error for the number of epochs decided. This process will be repeated trying to improve our predictions

and minimize the prediction error. After completi+ng this process, the model will be trained. The classes of NN that predict future value base on passed sequence of observations is called Recurrent Neural Network (RNN) this type of NN make use of earlier stages to learn of data and forecast futures trends. The earlier stages of data should be remembered to predict and guess future values, in this case the

hidden layer act like a stock for the past information from the sequential data. The term recurrent is used to describe the process of using elements of earlier sequences to forecast future data. RNN can’t store long time memory, so the use of the Long Short-Term Memory (LSTM) based on “memory line” proved to be very useful in forecasting cases with long time data. In a LSTM the memorization of earlier stages can be performed trough gates with along memory line incorporated. The following diagram-1 describe the composition of LSTM nodes.



The ability of memorizing sequence of data makes the LSTM a special kind of RNNs. Every LSTM node most be consisting of a set of cells responsible of storing passed data streams, the upper line in each cell links the models as transport line handing over data from the past to the present ones, the independency of cells helps the model dispose filter of add values of a cell to another. In the end the sigmoidal neural network layer composing the gates drive the cell to an optimal value by disposing or letting data pass through. Each sigmoid layer has a binary value (0 or 1) with 0 “let nothing pass through”; and 1 “let everything pass through.” The goal here is to control the state of each cell

1. **Methodology and Data**

The data in this paper consist of the daily opening prices of any stocks real time data extracted from yahoo finance, for GOOGL our data series cover the period going from

2018-01-01 to Till Date and for NKE the data cover the period from 1/4/2010 to 12/19/2019.

To build our model we are going to use the LSTM RNN, our model uses 70% of data for training and the other 30% of data for testing. For training we use mean squared error to optimize our model.Also, we used different Epochs for training data (12 epochs, 25 epochs, 50 epochs and 100 epochs) our model will be structured as follow:

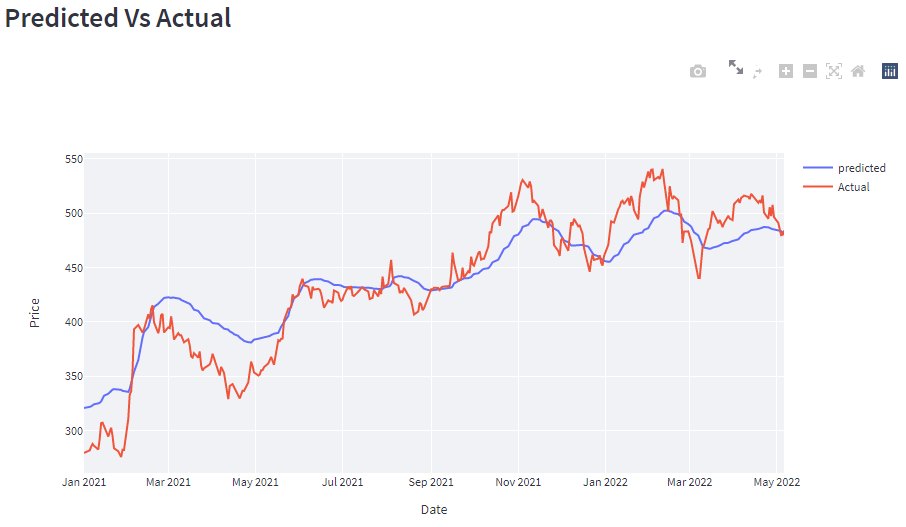
|  |  |  |
| --- | --- | --- |
| **Layer (Type)** | **Out Shape** | **Parameters** |
| lstm\_1 (LSTM) | (None, 100, 50) | 10400 |
| dropout (Dropout) | (None, 100, 50) | 0 |
| lstm\_2 (LSTM) | (None, 100, 60) | 26640 |
| dropout\_1 (Dropout) | (None, 100, 60) | 0 |
| lstm\_3 (LSTM) | (None, 100, 80) | 45120 |
| dropout\_2 (Dropout) | (None, 100, 80) | 0 |
| lstm\_4 (LSTM) | (None, 120) | 96480 |
| dropout\_3 (Dropout) | (None, 120) | 0 |
| dense (Dense) | (None, 1) | 121 |

# Results And Discussion

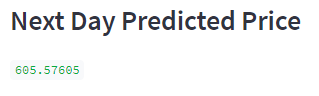
1. **Closing Price VS Time Chart**



1. **Original Value VS Predicted Value**



1. **Next Day Prediction**



**Conclusion**

This paper proposes RNN based on LSTM built to forecast future values of any stock, the result of our model has shown some promising result. The testing result conform that our model is capable of tracing the evolution of opening prices for both assets. For our future work we will try to find the best sets for bout data length and number of training epochs that beater suit our assets and maximize our predictions accuracy.

# BIBLIOGRAPHY