ABSTRACT
In early days it was very hard to get stock data but now it is very easy to retrieve data from yahoo finance. Here in this model I used Tata Motors stock price historical data to predict the stock price.

Keywords
Regression, Stock, Prediction, etc

1. INTRODUCTION
If you have a money and you want to invest but don’t know where to invest then you need a proper knowledge to invest your money. Stock market always fluctuate and no one knows the future. These stock price depends on many factor such as opening price, close price, low cost, high cost and many others smaller factor like company’s events, political events, natural disaster and world wars. Stock price prediction is very hard and important issue to investigate by company and researchers. At early days linear regression was used to make a predictive model.

After some time polynomial regression, neural network begin to use in many models but the failure of these models leads to dawn of neural network but after some years a company again started using neural network with a new name called deep learning and making predictive models and try to increase their performance

2. DATA SET
A collection of data is taken from the banking sector. The Data set is in ARFF (Attribute-Relation File Format) format that is acceptable by Weka. ARFF file is composed of tags that include the name, types of attributes, values and data itself. For this paper we are using 8 attributes like opening price, closing price, low, high etc.

The table below represents the data set that we have used:

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company name</td>
<td>Company in which you want to invest</td>
<td>Character</td>
</tr>
<tr>
<td>Date</td>
<td>Date wise data</td>
<td>Date</td>
</tr>
<tr>
<td>Open price</td>
<td>Market opening price</td>
<td>Integer</td>
</tr>
<tr>
<td>Close Price</td>
<td>Market Closing Price</td>
<td>Integer</td>
</tr>
<tr>
<td>High</td>
<td>Highest Price of the date</td>
<td>Integer</td>
</tr>
<tr>
<td>Low</td>
<td>Lowest Price of the date</td>
<td>Integer</td>
</tr>
<tr>
<td>Volume</td>
<td>Trading Volume</td>
<td>Integer</td>
</tr>
<tr>
<td>Performance</td>
<td>System performance and result</td>
<td>Float</td>
</tr>
</tbody>
</table>

Now in machine learning model, we first apply the training data set, in this data set the model is trained with known examples. The entries of new applicants will act as a test data which are to be filled at the time of submitting the application.

3. METHODOLOGY
In this project our team is worked on Recurrent neural network and LSTM(long short term memory) to predict stock price and try to increase the accuracy.

RNN (Recurrent Neural Network):- It is a part of Artificial neural network in which the connections between nodes form a directed graph from a time sequence. It is a supervised learning that is used in Alexa, speech recognition, image recognition and many more new technologies.

Basically RNNs is a network of neuron like nodes organized into consecutive layers. Each node in a given
layer is connected with a directed (one-way) connection to every other node in the next successive layer. Each node (neuron) has a time-varying real-valued activation. Each and every connection can be modified for valued prediction. Nodes are either input layer (receiving data from outside the network), output layer, or hidden layer.

In reinforcement learning settings, no teacher provides target signals. Rather a fitness function is used to evaluate the recurrent neural network performance that effect its input stream by output units which is connected to actuators and affect the environment. It is more like we play a game and performance is based on number of points won.

Each sequence produces an error as the sum of the deviations of all target signals from the corresponding activations computed by the network. For a training set of numerous sequences, the total error is the sum of the errors of all individual sequences.

Human don’t understand directly anything they get trained just like mathematics you practice and then you get able to solve any question just like that our machine or model need to learn and practice from various data and based on that data it predict the possible outcome.

Recurrent Neural Network work in a loop to build from basic to building.

\[
\begin{align*}
\text{h}_t &= \text{A} \\
\text{X}_t &\rightarrow \\
\end{align*}
\]

Fig 1: An unrolled recurrent neural network

3.1 LSTM:- LONG SHORT-TERM MEMORY is an architecture of RNN in deep learning. LSTM is not a feed forward neural network, it is a feedback connection that store the whole sequence of data. It is nowadays used in handwriting recognition anomaly detection and also used in alexa.

It is composed of 3 gates --

1. A cell
2. Input gate
3. Output gate
4. Forget gate

Cell is used to store data and the 3 gates is used to regulate the data/information in and out of the model.

below architecture is a chain-like architecture which reveals that recurrent neural networks are intimately related to sequences and lists. These are the natural architecture of recurrent neural network we used in our project in our dataset.

4. Model Design And Implementation

This model is based on RNN with regression and intend to make prediction using libraries like numpy, pandas, keras and to plot graph we use matplotlib.

The numpy is used to convert the data into an array, pandas is used to take input in the form of .csv or in an excel format. The data is too complex for calculation so we need to convert it into simple. In this project we used min-max function to normalize our data and reshaped it in between 0 and 1. This feature scaling is done to minimize the loss in worst case scenario.

Then we create 60 timestamp to map the data i.e, we take 59 data and map it to the 60th data. Our system require a 3-Dimensional data so we reshape our 2-Dimensional data into 3-Dimensional by adding a new axis as 1. Keras library is used to make nerve structure of our system in this model we used five neuron to train the and use LSTM to store the data.

We used 250 epoch to tune our model to get maximum accuracy and minimum error. Epoch is a parameter which defines the number of times that the learning algorithm will work through the whole dataset. In this model we used 1259 data entry to train our system and 21 days entry to test our model.

4.1 Test 1.0

In our first test run we get 67.21% accuracy and got the following graph:-
Our system predict the next month data as shown in the figure:

4.2 Test 2.0
In the same model when we train our model on tata motors stock price value we get 54.96% accuracy.

This is less because of covid-19 pandemic. Our result for latest data shows less accuracy.

And our next month prediction statistics graph is shown as:

5. Conclusion
In this project, prediction based on LSTM based network model is proposed to predict the price of stock simultaneously. This project is based on deep recurrent neural network. The algorithm framework, experiment design and model structure are shown. We used three data set to verify the applicability of our Associated Net-model. Experiment shows that average accuracy of associated Net-model is much better than previous research paper. Not only that it can predict multiple values simultaneously and average accuracy of each predicted value is over 65%. This model achieves good effect, there are still some aspects can be improved. Like, the model can be optimized by reducing the total loss.

This loss calculation method does not take into count the relationship between each sub-loss, as well as some details when the total loss is the minimum, such as the extreme situation of each sub-loss and move in the direction of loss reduction.

Moreover, stock market is heavily based on human emotions. They feel and based on various factor and other physical degree can be considered for calculating the possibility. Various social media platform can be used for collecting the news and articles, where emotions are gauged with it. This project can be linked with the human emotion based factor to better train weights and further improve accuracy.

6. References
[3] Data on website :-
[4] Data on website :-
   https://en.wikipedia.org/?title=LSTM&redirect=no