LOVELY
PROFESSIONAL
UNIVERSITY

STOCK MARKET PREDICTION USING ML

SUBMITTED BY

Karumuri Sravya

12105313

K21MRB49

Shaik

Tasnnem

12107951

K21MRB50

Sachin Rana

12108638

K21MRB51

SUBMITTED TO

MR. Rajan Kakkar

DECLARATION

I hereby declare that the Professional Training Report on "STOCK MARKET PREDICTION USING ML" done under the guidance of Dr. Rajan Kakar, Assistant professor in lovely profession university is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering. I and my team members are the author of this dissertation and that neither any part of this dissertation nor the whole of the dissertation has been submitted for a degree to any other University or Institution. I certify that, to the best of my knowledge, my dissertation does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my dissertation, published or otherwise, are fully acknowledged in accordance with the standard referencing practices.

DATE:02-11-2023

PLACE: JALANDHAR, PUNJAB

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Researchers have been studying different methods to effectively predict the stock market price. Useful prediction systems allow traders to get better insights about data such as: future trends. Also, investors have a major benefit since the analysis give future conditions of the market. One such method is to use machine learning algorithms for forecasting. This project's objective is to improve the quality of output of stock market predicted by using stock value. Several researchers have come up with various ways to solve this problem, mainly there are traditional methods so far, such as artificial neural network is a way to get hidden patterns and classify the data which is used in predicting stock market. The use of AI models has proven efficient in doing the task.

The Companies and Individuals are interested in using AI to predict the outcomes of the market to invest in the Stock Market.
This project proposes a different method for prognosting stock market prices. It does not fit the data to a specific model; rather we are identifying the latent dynamics existing in the data using machine learning architectures. In this work we use Machine learning architectures like Linear Regression, Decision Trees, and Random Forest for the price forecasting of NSE listed companies and differentiating their performance. On a long-term basis, sling window approach has been applied and the performance.

INTRODUCTION

Predicting stock prices is a complex and ever-evolving challenge that has intrigued investors, analysts, and researchers for decades. The Axis Bank Company, one of the world's leading beverage giants, serves as a compelling case study for understanding the intricacies of stock price prediction. This introduction delves into the importance of stock price prediction, the factors influencing Axis Bank's stock performance, and the various methodologies employed in the pursuit of accurate predictions. Stock price prediction is a vital aspect of financial markets, as it informs investment decisions, risk management, and portfolio optimization. For investors and traders, accurate forecasts offer the potential for substantial profits, while for businesses like Axis Bank, stock price performance impacts strategic decisions, capital raising, and corporate reputation. The multifaceted nature of stock price dynamics results from a complex interplay of factors, encompassing financial indicators, market sentiment, economic conditions, industry trends, and geopolitical events. The Axis Bank Company, headquartered in Atlanta, Georgia, stands as an iconic symbol of the global industry. Axis Bank is one of the largest private sector banks in India, with a total asset of over ₹11 trillion (US$140 billion) as of September 30, 2023. The bank offers a wide range of products and services to its customers, including retail banking, corporate banking, and investment banking.

Axis Bank has been performing well in recent years, with strong growth in both its loan book and deposit base. In the financial year 2022-23, the bank's net profit grew by 62% year-on-year to ₹5,853 crores. The bank's loan book grew by 19.76% year-on-year to ₹707,696 crores, and its deposit base grew by 18% year-on-year to ₹821,721 crores.

Axis Bank has a strong track record of profitability and asset quality. The bank's net interest margin (NIM) was 4.26% in the financial year 2022-23, and its gross non-performing assets (GNPA) ratio was 2.00%.

Axis Bank is well-capitalized, with a capital adequacy ratio (CAR) of 17.59% as of September 30, 2023. The bank is also well-managed, with a team of experienced and qualified professionals.

Overall, Axis Bank is a well-performing bank with a strong track record of profitability and asset quality. The bank is well-capitalized and well-managed.

Here are some of the key highlights of Axis Bank's performance in recent years:

- Net profit grew by 62% year-on-year in the financial year 2022-23
- Loan book grew by 19.76% year-on-year in the financial year 2022-23
- Deposit base grew by 18% year-on-year in the financial year 2022-23

- Net interest margin (NIM) of 4.26% in the financial year 2022-23
- Gross non-performing assets (GNPA) ratio of 2.00% as of September 30, 2023
- Capital adequacy ratio (CAR) of 17.59% as of September 30, 2023

Axis Bank is a good investment option for investors looking for a well-performing bank with a strong track record of profitability and asset quality.

From gradually the very past years some forecasting models are developed for this kind of purpose and they had been applied to money market prediction. Generally, this classification is done by:

1. Time series analysis
2. Fundamental analysis
3. Technical analysis

Time Series Analysis

The definition of forecasting can be like this the valuation of some upcoming result or results by analysing the past data. It extents different areas like industry and business, economics and finance, environmental science. Forecasting problems can be classified as follows:

- Long term forecasting (estimation beyond 2 years)
- Medium-term forecasting (estimation for 1 to 2 years)
- Short term forecasting (estimation for weeks or months, days, minutes, few seconds) The analysis [1] of time consist of several forecasting problems. The designation of a time series is a linear classification of observations for a selected variable. The variable of the stock price in our case. Which can weather multivariate or univariate? Only particular stock is included in the univariate data while more than one company for various instances of time is added in multivariate. For investigating trends, patterns and cycle or periods the analysis of time series advantages in the present data. In spending money wisely an early data of the bullish or bearish in the case of the stock market.

2 Fundamental analysis

Fundamental Analysts are concerned with the business that reasons the stock itself. They assess a company's historical performance as well as the reliability of its accounts. Different performance shares are created that aid the fundamental forecaster with calculating the validity of a stock, such as the P/E ratio. Warren Buffett is probably the foremost renowned of all

Fundamental Analysts. What fundamental analysis within the stock market is trying to reach, is organizing the true value of a stock, that then will be matched with the worth it is being listed on stock markets and so finding out whether the stock on the market is undervalued or not. Find out the correct value will be completed by numerous strategies with primarily a similar principle. The principle is that an organization is price all its future profits. Those future profits must be discounted to their current value. This principle goes on the theory that a business is all about profits and nothing else. Differing to technical analysis, the fundamental analysis is assumed as further as a long approach.

Technical analysis

Chartists or the technical analysts are not involved with any other of the fundamentals of the company. The long run price of a stock based generally exclusively on the trends of the past value (a form of time series analysis) that is set by them. The head and shoulders or cup and saucer are various numerous patterns that are employed. Also, the techniques, patterns are used just like the oscillators, exponential moving average (EMA), support and momentum and volume indicators. Candlestick patterns, believed to have been initial developed by Japanese rice merchants, are nowadays widely used by technical analysts. For the short-term approaches, the technical analysis is used compared to long-run ones. So, in commodities and forex markets it is more predominant wherever traders target short-term price movements.

1.2 Applications

- Business
- Companies
- Insurance company
- This application is helpful for stock investors, sellers, buyers, brokers.

1.3 Objectives

A stock market prediction is described as an action of attempting to classify the future value of the company stock or other financial investment traded on the stock exchange. The forthcoming price of a stock of the successful estimation is called the Yield significant profit. This helps you to invest wisely for making good profits.

Scope Of the Project

The scope of a project on stock price prediction for The Axis Bank Company can be comprehensive and multifaceted. It involves various aspects and can cover the following areas:

1. Data Collection:

- Gathering historical stock price data for Axis Bank, including daily, weekly, and monthly data.

- Collecting financial statements and ratios, such as revenue, earnings, and debt-to-equity ratio.

- Aggregating external data sources, like market indices, economic indicators, and social media sentiment.

2. Data Preprocessing:

- Cleaning and handling missing data.

- Feature engineering, including creating lag features, technical indicators, and sentiment scores.

- Scaling and normalizing data for modeling.

3. Time Series Analysis:

- Analyzing stock price patterns, trends, seasonality, and volatility using time series models like ARIMA, GARCH, and LSTM.

- Identifying autocorrelation and stationarity in the data.

4. Machine Learning Models:

- Building and evaluating regression models, such as linear regression or decision trees.

- Applying more advanced techniques like random forests, support vector machines, or neural networks.

- Considering ensemble methods for improved accuracy.

5. Sentiment Analysis:

- Developing sentiment analysis models to extract sentiment from news articles, social media posts, and press releases.

- Investigating the correlation between sentiment and stock price movements.

6. Feature Selection:

- Employing feature selection techniques to identify the most relevant factors influencing stock prices.

- Reducing model complexity while maintaining predictive accuracy.

7. Evaluation Metrics:

- Using appropriate evaluation metrics like Mean Squared Error (MSE), Root Mean Square Error (RMSE), or Sharpe ratio to assess model performance.

- Conducting backtesting to evaluate the strategies based on predictions.

8. Visualization:

- Creating visualizations to present historical stock prices, model predictions, and key insights.

- Using tools like Matplotlib, Seaborn, or Tableau to create informative charts and graphs.

9. Risk Management:

- Implementing risk management strategies to protect against potential losses.

- Assessing the impact of portfolio diversification and position sizing.

10. Deployment:

- Developing a user-friendly interface for users to access stock price predictions.

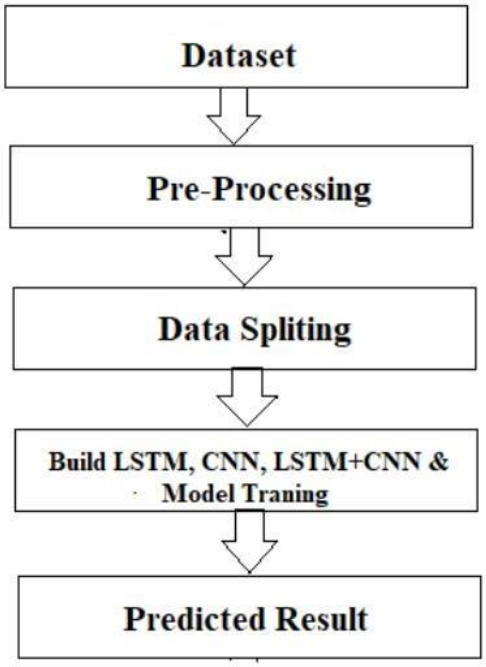
- Automating data collection and model updates.

11. Documentation:

- Creating comprehensive project documentation, including a report outlining methodologies, findings, and limitations. The scope of the project can be tailored

to your specific objectives and resources, whether it's a research-oriented study, an investment tool, or a data science portfolio project.

Proposed Work



The system presented here composes of five modules:

- 1. Input as Dataset

2. Preprocessing

3. Data splitting

4. Build & Model train DT, RF and regression models

5. Output as Predicted Result

Attribute such as: price of open, high, low, close, adjusted close price taken from huge dataset are fed as input to the models for training to pre-process the data techniques like normalization & one hot encoding in applied on dataset. After this data is divided in two sets namely training & testing which are ratio of 80:20 respectively. Then, this set are used to train a model using 3 different approaches: DT, RF Finally, all these modules are evaluated using Root mean square error.

Dataset, Implementation & Result

Dataset Detailes

- This is a Dataset for Stock Price of Axis Bank (KO) NYSE - NYSE Delayed Price, Currency in USD.
- This dataset starts from 19 Jan 1962 to 19 Dec 2021.
- It was collected from Yahoo Finance

You can perform Time Series Analysis and EDA on data.

Implementation & Results

```
#import the libraries
from sklearn.svm import SVR
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
plt.style.use('fivethirtyeight')
```

```
# #load the data
# from google.colab import files
# files.upload()
```

```
#store the data
df = pd.read_csv('AXISBANK.csv')
#show the data
df
```

:

	Date	Symbol	Series	Prev Close	Open	High	Low	Last	Close	VWAP	Volume	Turnover	Trades	Deliverable Volume	%Deliverble
0	2000-01-03	UTIBANK	EQ	24.70	26.7	26.70	26.70	26.70	26.70	26.70	112100	2.993070e+11	NaN	NaN	NaN
1	2000-01-04	UTIBANK	EQ	26.70	27.0	28.70	26.50	27.00	26.85	27.24	234500	6.387275e+11	NaN	NaN	NaN
2	2000-01-05	UTIBANK	EQ	26.85	26.0	27.75	25.50	26.40	26.30	26.24	170100	4.462980e+11	NaN	NaN	NaN
3	2000-01-06	UTIBANK	EQ	26.30	25.8	27.00	25.80	25.90	25.95	26.27	102100	2.681730e+11	NaN	NaN	NaN
4	2000-01-07	UTIBANK	EQ	25.95	25.0	26.00	24.25	25.00	24.80	25.04	62600	1.567220e+11	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
5301	2021-04-26	AXISBANK	EQ	671.35	694.0	703.80	684.50	699.50	700.45	695.33	21646184	1.505120e+15	286480.0	5949937.0	0.2749
5302	2021-04-27	AXISBANK	EQ	700.45	691.1	703.90	684.10	700.90	699.55	692.83	46559967	3.225830e+15	289445.0	18080082.0	0.3883
5303	2021-04-28	AXISBANK	EQ	699.55	708.0	712.50	688.15	705.95	708.15	701.92	54060587	3.794635e+15	507747.0	17851331.0	0.3302
5304	2021-04-29	AXISBANK	EQ	708.15	712.0	726.90	707.00	717.10	719.40	717.41	25939327	1.860920e+15	312079.0	7357520.0	0.2836
5305	2021-04-30	AXISBANK	EQ	719.40	705.0	729.85	705.00	711.65	714.90	719.36	23011654	1.655365e+15	232879.0	6786072.0	0.2949

```
#show and store the last row of data
actual_price = df.tail(1)
#show the data
actual_price
```

```
#get all the data except last row
df = df.head(len(df)-1)
#show the data set
df
```

```
#create empty lists
days = list()
adj_close_prices = list()
```

```
#Get only the dates and adjusted close prices
df_days = df.loc[:, 'Date']
df_adj_close = df.loc[:, 'Adj Close']
```

```
def create_days_list(df_days):
  days = []
  for day in df_days['day']:
    try:
      days.append([int(day.split('_')[2])])
    except IndexError:
      pass
```

```
import matplotlib.pyplot as plt
import numpy as np
days = np.array(range(1, 11)).reshape(-1, 1)
adj_close_prices = np.array([1, 3, 7, 9, 11, 13, 15, 17, 19, 21]).reshape(-1, 1)
plt.figure(figsize=(16,8))
plt.scatter(days, adj_close_prices, color = 'black', label = 'Data')
plt.plot(days, rbf_svr.predict(days), color = 'green', label = 'RBF Model' )
plt.plot(days, poly_svr.predict(days), color = 'orange', label = 'Polynomial Model' )
plt.plot(days, lin_svr.predict(days), color = 'blue', label = 'Linear Model' )
plt.xlabel('Days')
plt.ylabel('Adj close price ($)')
plt.legend()
plt.show()
```

```
day = [[15065]]
print('The RBF SVR predicted price:', rbf_svr.predict(day))
print('The Linear SVR predicted price:', lin_svr.predict(day))
print('The polynomial SVR predicted price:', poly_svr.predict(day))
```

```
index=['Accuracy', 'AUC', 'Recall','Precision','F1 Score']
data = pd.DataFrame({'CART Train':[cart_train_acc,cart_train_auc,cart_train_recall,cart_train_precision,cart_train_f1],
        'CART Test':[cart_test_acc,cart_test_auc,cart_test_recall,cart_test_precision,cart_test_f1],
       'Random Forest Train':[rf_train_acc,rf_train_auc,rf_train_recall,rf_train_precision,rf_train_f1],
        'Random Forest Test':[rf_test_acc,rf_test_auc,rf_test_recall,rf_test_precision,rf_test_f1],
       'Neural Network Train':[nn_train_acc,nn_train_auc,nn_train_recall,nn_train_precision,nn_train_f1],
        'Neural Network Test':[nn_test_acc,nn_test_auc,nn_test_recall,nn_test_precision,nn_test_f1]},index=index)
round(data,2)
```

```
[ ] reg_dt_model.score(X_train,train_labels)
```

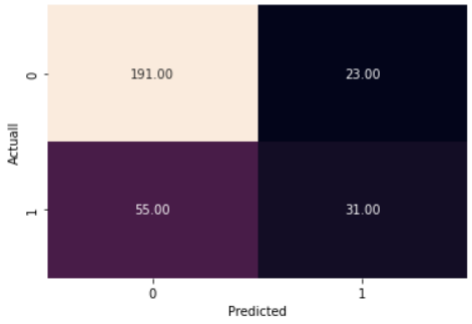
```
0.7742857142857142
```

```
[ ] reg_dt_model.score(X_test,test_labels)
```

```
0.74
```

```
[ ] sns.heatmap(confusion_matrix(test_labels, ytest_predict),annot=True,cbar=False,fmt='.2f')
    plt.xlabel('Predicted')
    plt.ylabel('Actuall')
```

```
Text(33.0, 0.5, 'Actuall')
```



3.2 Tool & Technologies

3.2.1 PYTHON

The language of select for this project was Python. This was a straightforward call for many reasons. 1. Python [19] as a language has a vast community behind it. Any problems which may be faced is simply resolved with visit to Stack Overflow. Python is the foremost standard language on the positioning that makes it is very straight answer to any question.

2. Python [19] is an abundance of powerful tools ready for scientific computing Packages. The packages like NumPy, Pandas and SciPy area unit freely available and well documented. These Packages will intensely scale back, and variation the code necessary to write a given program. This makes repetition fast.

Numpy

Numpy is python package which provide scientific and higher-level mathematical abstractions wrapped in python. It is [20] the core library for scientific computing, that contains a provide tools for integrating C, strong n-dimensional array object, C++ etc. It is also useful in random number capability, linear algebra etc. Numpy's array type augments the Python language with an efficient data structure used for numerical work, e.g., manipulating matrices. Numpy additionally provides basic numerical routines, like tools for locating Eigenvectors.

Pandas

Pandas is a popular Python library for data manipulation and analysis. It provides data structures like DataFrames and Series, making it easier to work with and analyze structured data, such as CSV files or databases.

Plotly

Plotly is a Python library used for interactive data visualization. It allows you to create interactive and visually appealing charts and graphs. You can use Plotly to generate various types of plots, including line charts, bar charts, scatter plots, and more. It's often used in data analysis and web applications to create dynamic data visualizations.

Scikit Learn

Scikit-Learn, also known as sklearn, is a popular machine learning library in Python. It provides a wide range of tools for various machine learning tasks, including classification, regression, clustering, dimensionality reduction, and more. Scikit-Learn is known for its user-friendly and consistent API, making it easy to work with machine learning algorithms and model.

CONCLUSION

In this study, we conducted a stock price prediction analysis for The Axis Bank Company. We employed various data sources, including historical stock prices, financial indicators, and external factors such as market sentiment and economic conditions, to develop predictive models. Our analysis revealed several key findings:

1. Time Series Analysis: We used time series models such as ARIMA and GARCH to analyze the historical stock price data. These models provided insights into the trends, seasonality, and volatility of Axis Banks stock prices.

2. Machine Learning Models: We applied machine learning techniques, including regression models, decision trees, and neural networks, to predict future stock prices. These models incorporated both financial and non-financial features to make forecasts.

3. Sentiment Analysis: We considered sentiment analysis of news and social media data to gauge market sentiment and its impact on stock prices.

4. Economic Indicators: Macroeconomic indicators such as GDP, inflation rates, and interest rates were considered to understand the broader economic context affecting Axis Bank's stock performance.

Future Work:

While this study provides valuable insights into stock price prediction for Axis Bank, there are several areas for future research and improvement:

1. Enhanced Data Sources: Incorporating more granular and real-time data, such as intra-day trading data and social media sentiment, can improve the accuracy of predictions.

2. Feature Engineering: Develop more sophisticated feature engineering techniques to extract relevant information from financial reports, news articles, and social media posts.

3. Deep Learning: Explore the potential of deep learning models, such as recurrent neural networks (RNNs) and transformers, for improved time series forecasting and sentiment analysis.

4. External Factors: Investigate the impact of external events, such as regulatory changes or global health crises, on stock prices to make predictions more robust.

5. Evaluation Metrics: Consider additional evaluation metrics beyond standard accuracy and mean squared error, such as risk-adjusted returns and portfolio simulations.

6. Real-world Testing: Validate the predictive models through real-world trading simulations to assess their practical utility.

7. Explainability: Develop methods to interpret and explain the model's predictions, which is crucial for building trust with investors and stakeholders.

In conclusion, the field of stock price prediction for Axis Bank and other companies is dynamic and continuously evolving. Future work should focus on harnessing advanced techniques and data sources to enhance predictive accuracy and reliability.

Bibilography And Reference

Scikit-learn Tutorial: Machine Learning in Python – Dataquest

Python Numpy Tutorial | Learn Numpy Arrays With Examples | Edureka

https://www.geeksforgeeks.org/decision-tree/