



STOCK PRICE PREDICTION SYSTEM

A Project Work Synopsis

*Submitted in the partial fulfilment for the award of the
degree of*

BACHELOR OF ENGINEERING

IN

ARTIFICIAL INTELLIGENCE AND

MACHINE LEARNING

Submitted by:

**Abhishek Verma
20BCS6707**

SUPERVISORS NAME

Ramanjot Kaur Ma'am

**DEPARTMENT OF COMPUTER SCIENCE AND
ENGINEERING APEX INSTITUTE OF TECHNOLOGY**

**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI -
140413,
PUNJAB**

MONTH & YEAR

31 AUGUST, 2022

ABSTRACT

In this study, I aimed to create a Stock Price Prediction System. Stock Price can be something hard to guess. we have been provided with regular price and least possible price between the months of October and November of 2022, using which we aim to build a model which predicts the best possible prices of the stocks using various input features.

DECLARATION

I, **Abhishek Verma** students of **‘Bachelors of engineering in CSE (HONS.) with specialization in artificial intelligence and machine learning in association with IBM’**, session: **2020-24**, Department of Computer Science and Engineering, Apex Institute of Technology, Chandigarh University, Punjab, hereby declare that the work presented in this Project Work entitled **‘Stock Price Prediction System’** is the outcome of my own bona fide work and is correct to the best of my knowledge and this work has been undertaken taking care of Engineering Ethics. It contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Names

Abhishek Verma (20BCS6707)

ACKNOWLEDGMENT

It is indeed with great pleasure and an immense sense of gratitude that we acknowledge the help of these individuals. We are highly indebted to our Chancellor **S. SATNAM SINGH SANDHU**, VICE-CHANCELLOR Dr. **H.B. RAGHAVENDAR**, CHANDIGARH UNIVERSITY, for the facilities provided to accomplish this main project. We would like to thank our P.L. **Ramanjot Kaur** for this constructive criticism throughout our project. We feel elated in manifesting our sense of gratitude to our internal project guide **MR. NIRMALYA BASU**. He has been a constant source of inspiration for us and we are very deeply thankful to him for his support and valuable advice. We are extremely grateful to our Departmental staff members for their extreme help throughout our project. Finally, we express our thanks to all of our friends who helped us in the successful completion of this project.

Project Associates:
Abhishek Verma (20BCS6707)

LIST OF TABLES

<i>Figure</i>	<i>title</i>	<i>page</i>
---------------	--------------	-------------

2.1	Existing System Summary	9
1.4.4	Softwares used	8

Table of Contents

SR.NO		PAGE NUMBER
0	Title page Abstract List of Tables	1 2 3
1	INTRODUCTION* 1. Problem Definition 2. Project Overview/Specifications* 3. Hardware Specification 1.3.1 PC 1.4 Software Specification 1.4.1 Jupyter Notebook 1.4.2 Heroku 1.4.3 Flask 1.4.4 MS-Excel	6 7 8
2	LITERATURE SURVEY 1. Existing System Summary 2. Proposed System 3. Feasibility Study	9
3	PROBLEM FORMULATION	10
4	RESEARCH OBJECTIVES	10
5	METHODOLOGY	11
6	TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK	12
7	REFERENCES	15

1 INTRODUCTION

1.1 PROBLEM DEFINITION

This project aims to develop an application which will predicts the best possible prices of the stocks using various input features. The user will get the predicted stock price and with regular price and least possible price.



1.2 PROJECT OVERVIEW/ SPECIFICATIONS

A thorough descriptive analysis needs to be done to meet the mentioned objectives:

- The first step in LSTM is to decide which information to be omitted from the cell in that particular time step. It is decided with the help of a sigmoid function. It looks at the previous state (h_{t-1}) and the current input x_t and computes the function.
- There are two functions in the second layer. The first is the sigmoid function, and the second is the tanh function. The sigmoid function decides which values to let through (0 or 1). The tanh function gives the weightage to the values passed, deciding their level of importance from -1 to 1.

- The third step is to decide what will be the final output. First, you need to run a sigmoid layer which determines what parts of the cell state make it to the output. Then, you must put the cell state through the tanh function to push the values between -1 and 1 and multiply it by the output of the sigmoid gate.

1.3 HARDWARE SPECIFICATION

1.3.1 PC

A pc is a personal computer that can be used for multiple purposes depending on its size, capabilities, and price. They are to be operated directly by the end-user. Personal computers are single-user systems and are portable. Our web application program will be installed on the pc for our clients to use it. This makes it feasible for individual use.

1.4 SOFTWARE SPECIFICATION

1.4.1 Jupyter Notebook:

Jupyter Notebook is a web-based open-source application that is used for editing, creating running, and sharing documents that contain live codes, visualization, text, and equations. Its core supported programming languages are Julia, R, and Python. Jupyter notebook comes with an IPython kernel that allows the programmer to write programs in python. There are over 100 kernels other than IPython available for use.

1.4.2 Kaggle

Kaggle, a subsidiary of Google LLC, is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges

1.4.3 POWER BI

Power BI provides cloud-based BI (business intelligence) services, known as "Power BI Services", along with a desktop-based interface, called "Power BI Desktop". It offers data warehouse capabilities including data preparation, data discovery, and interactive dashboards.^[2] In March 2016, Microsoft released an additional service called Power BI Embedded on its Azure cloud platform.^[3] One main differentiator of the product is the

ability to load custom visualizations.

1.4.4 MS-EXCEL

Microsoft produced Microsoft Excel, a spreadsheet, for Windows, macOS, Android, and iOS. It has calculating or computing capabilities, graphing tools, pivot tables, and the Visual Basic for Applications macro programming language (VBA). The Microsoft Office programme package includes Excel.

2 LITERATURE REVIEW

2.1 Existing System Summary

Year and citation	<i>K. Tziridis T. Kalampokas G.Papakostas and K. Diamantaras "Airfare price prediction using machine learning techniques" in European Signal Processing Conference (EUSIPCO), DOI: 10.23919/EUSIPCO .2017.8081365L.</i>	<i>Li Y. Chen and Z. Li" Yawning detection for monitoring driver fatigue based on two cameras" Proc. 12th Int. IEEE Conf. Intel. Transp. Syst. pp. 1-6 Oct. 2009.</i>	<i>William Groves and Maria Gini "An agent for optimizing airline ticket purchasing" in proceedings of the 2013 international conference on autonomous agents and multi-agent systems.</i>
Article Title	A Machine Learning Approach for Flight fare Prediction: Model Regularization and Optimization	A Predictive Data Feature Exploration Based on flight fare Prediction Approach	
Purpose of the study	refined models to predict the hourly air pollution concentration on the basis of meteorological data of previous days by formulating the prediction over 24 hours as a multi-task learning (MTL) problem.	processing high-dimensional data and supporting the parallel learning, namely LightGBM, combined with the historical datasets to predict the air quality.	the proposed study puts forward a deep learning approach for quantification and prediction of ambient air quality
Tools/ Software used	Jupyter notebook, Python , Heroku url	Jupyter notebook, google collab	Jupyter , Flask , Heroku
Comparison of technique done	i. 13 years of data used ii. Model regularization iii. Multi-task learning framework used iv. Multilayer Perceptron(MLP) , Neural Network .	i. Airline data within the latest 3 years and the future 24- hour airline and flight data is used ii. Extreme Learning Machine(ELM) , Random Forest Regression . Bagging Regression Tree	i. A real-time dataset of airline was used ii. SVM , KNN , Gradient Boosting framework was implemented iii. Considers temporal sequential data of a particular airlines and flight
Findings	Airlinr and flight fare price prediction at the user end.	Flight Price and fare price of airlines at suitable rate	Prediction of the most likely flights prices.
Data set (if used)	Kaggle dataset , Airlines dataset of various flights	Web Scrapping , Data Science institute in the us, and Kaggle.	Vistara , Air Asia , GO Air , Indigo etc.
Evaluation parameters	achieve better performance than existing standard regression models and existing regularizations.	A high accurate prediction model is established using a multi-model fusion scheme	proposed KNN - SVM framework attained higher accuracy in estimating Flight based prices.

Table 2.1: Literature review

2.2 Proposed System

Here, we are providing the user friendly platform for the stock value prediction. The user first needs to select the stock file for which he/she wants to predict the opening value. We are providing many algorithms in our Interface. The user next needs to select the particular algorithm using which we would predict the opening value. Here, we are providing flexibility to user to choose multiple algorithms at a time. Next, the user needs to submit the selection. Finally, the user gets the both original opening value along with the predicted opening value for the selected algorithms. Here, if we have chosen a single algorithm the predicted values are generated only for that algorithm. But, if we have chosen multiple algorithms at a time, it gives the predicted values for all those algorithms separately. We can make comparisons among them.

3 PROBLEM FORMULATION

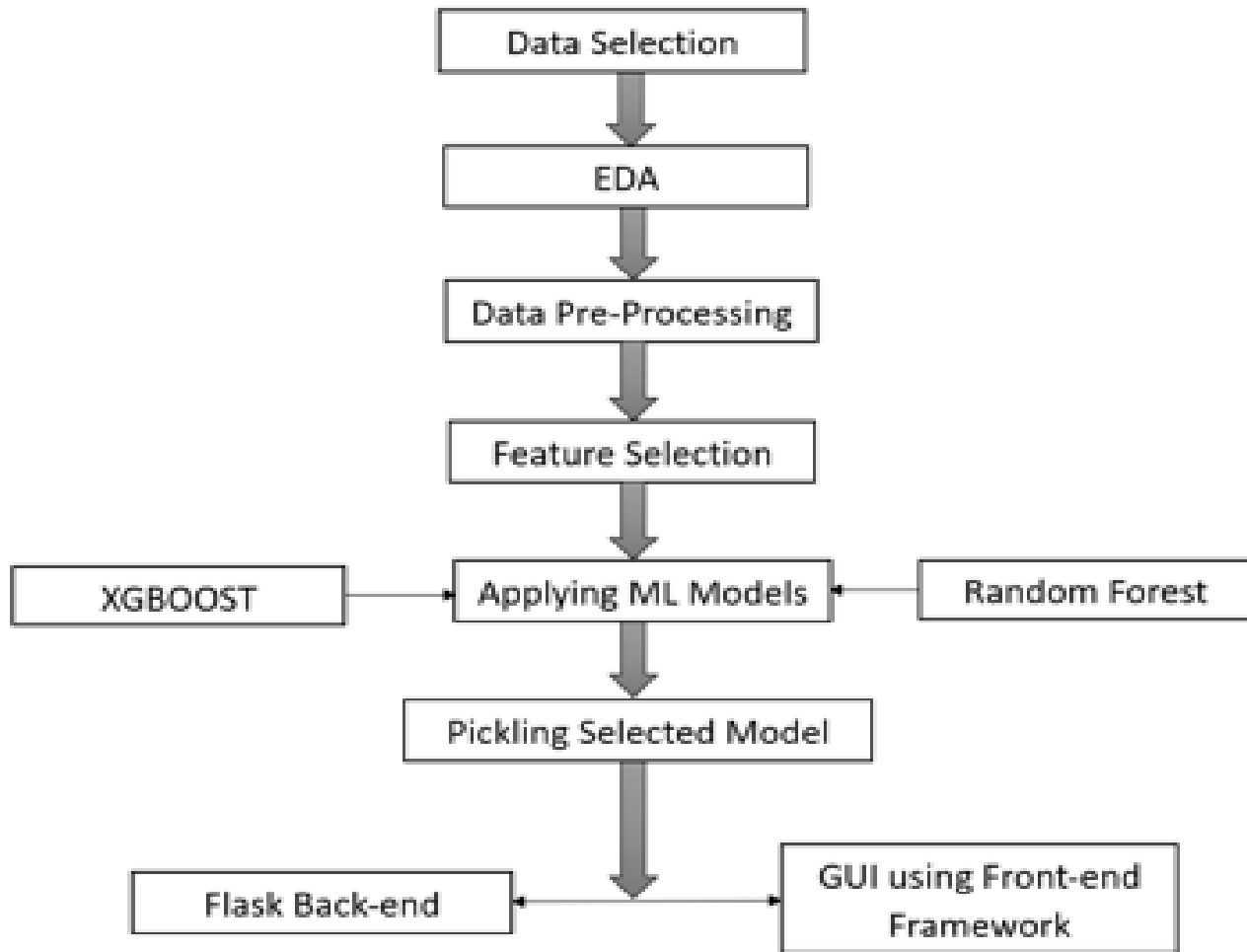


Fig. 1 Proposed System Diagram

4. RESEARCH OBJECTIVES

"What other people think" has always been an important piece of information for most of us during the decision-making process. The Internet and the Web have now (among other effects) made it possible to find out about the opinions and gestures of those in the vast pool of people that are neither our particular familiarity nor well-known professional critics — that is, people we've no way heard of. And again, further and further people are making their opinions available to non natives via the Internet. The interest that individual druggies show in online opinions about products and services, and the implicit influence similar opinions apply, is commodity that's driving force for this area of interest. And there are numerous challenges involved in this process which need to be walked each over in order to attain proper issues out of them. In this check we analysed introductory methodology that generally happens in this process and measures that are to be taken to overcome the challenges being faced

5. METHODOLOGY

The following methodology will be followed to achieve the objectives defined for the proposed research work:

The below mentioned are some parameters used in our data set:

1. *Size of Test Set*: 10683 rows & 11 columns
2. *Airline*: The name of the airline.
3. *Date of Journey*: The date of the journey.
4. *Source*: The source from which the service begins.
5. *Route*: Route of the flight, start to end.
6. *Destinations*: The destination where the service ends.
7. *Departure Time*: The time when the journey starts from the source.
8. *Arrival Time*: Time of arrival at the destination.
9. *Duration*: Total duration of the flight.
10. *Total Stops*: Total stops between the source and destination.
11. *Additional Info*: Additional information about the flight
12. *Price*: The price of the ticket

4]:

	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
0	IndiGo	24/03/2019	Banglore	New Delhi	BLR → DEL	22:20	01:10 22 Mar	2h 50m	non-stop	No info	3897
1	Air India	1/05/2019	Kolkata	Banglore	CCU → IXR → BBI → BLR	05:50	13:15	7h 25m	2 stops	No info	7662
2	Jet Airways	9/06/2019	Delhi	Cochin	DEL → LKO → BOM → COK	09:25	04:25 10 Jun	19h	2 stops	No info	13882
3	IndiGo	12/05/2019	Kolkata	Banglore	CCU → NAG → BLR	18:05	23:30	5h 25m	1 stop	No info	6218
4	IndiGo	01/03/2019	Banglore	New Delhi	BLR → NAG → DEL	16:50	21:35	4h 45m	1 stop	No info	13302

Fig. Dataset Contents

VIII. USE CASE DIAGRAM

Use Case Diagram of the project:

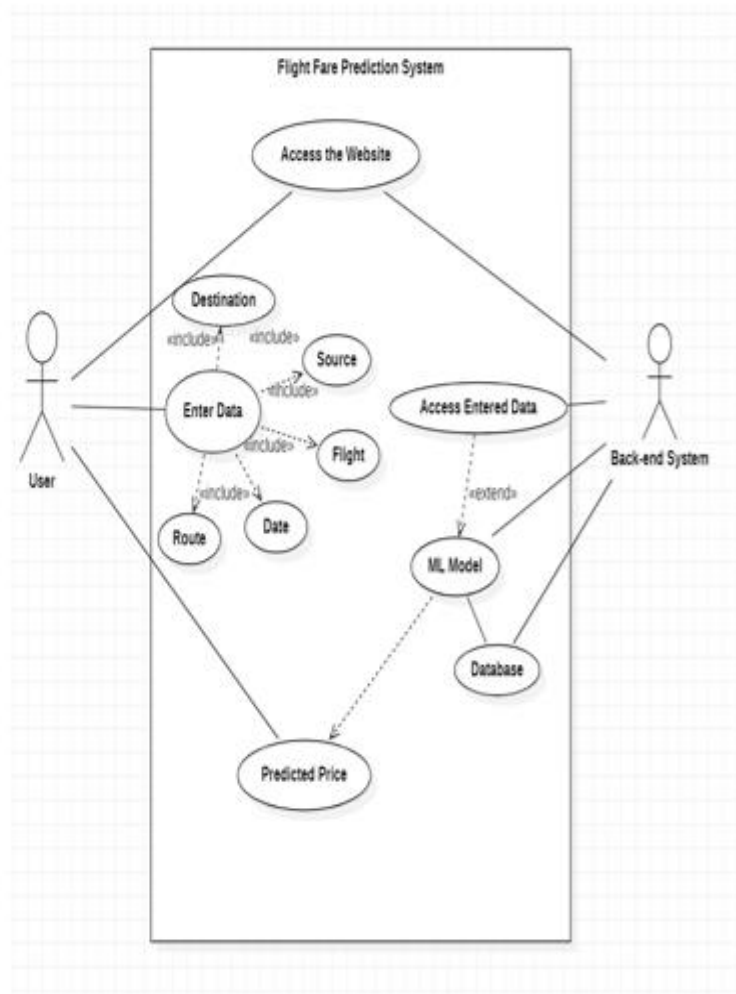


Fig. Use Case Diagram

TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: INTRODUCTION

The fiscal request is a dynamic and compound system where people can buy and vend currencies, stocks, equities and derivations over virtual platforms supported by brokers. The stock request allows investors to enjoy shares of public companies through trading either by exchange or over the counter requests. This request has given investors the chance of gaining plutocrat and having a prosperous life through investing small original quantities of plutocrat, low threat compared to the threat of opening new business or the need of high payment career. Stock requests are affected by numerous factors causing the query and high volatility in the request. Although humans can take orders and submit them to the request, automated trading

systems (ATS) that are operated by the perpetration of computer programs can perform better and with advanced instigation in submitting orders than any mortal. Still, to estimate and control the performance of ATSs, the perpetration of threat strategies and safety measures applied grounded on mortal judgements are needed. Numerous factors are incorporated and considered when developing an ATS, for case, trading strategy to be espoused, complex fine functions that reflect the state of a specific stock, machine literacy algorithms that enable the vaticination of the unborn stock value, and specific news related to the stock being analysed

CHAPTER 2: LITERATURE REVIEW

This chapter includes the literature available for Stock Price Prediction. The findings of the researchers will be highlighted which will become the basis of the current implementation.

CHAPTER 3: BACKGROUND OF PROPOSED METHOD

This chapter will provide an introduction to the concepts which are necessary to understand the proposed system.

CHAPTER 4: METHODOLOGY

This chapter will cover the technical details of the proposed approach with different flowcharts attached to it.

CHAPTER 5: EXPERIMENTAL SETUP

This chapter will provide information about the subject system and tools used for the evaluation of the proposed method.

CHAPTER 6: RESULTS AND DISCUSSION

The result of the proposed technique will be discussed in this chapter.

FLIGHT-TICKET PRICE

Departure Date

Arrival Date

Source

Destination

Stopage

Airline

CHAPTER 7: CONCLUSION

This work is concerned with prediction of stock price. These techniques have been used in this proposed system which is Support vector machine, Logistic, KNN, LSTM and Linear regression have shown the improvement in accuracy of the prediction by using these two techniques. Thereby it leads to the positive result in the prediction. Using the proper algorithm is able to predict the stock price with more accuracy. Using machine learning that leads to positive prediction of the stock price. Thereby it leads to the promising result in the prediction. Therefore, this project leads to the conclusion that one can predict the stock market price with more accuracy using machine learning. In the future the stock market prediction can be further improved by applying different algorithms to bring more accuracy. Use a real time dataset than the dataset available on a public repository that has been used in this work to predict. We want to extend this application for predicting crypto currency trading. We want to add sentiment analysis for better analysis.

CHAPTER 8: FUTURE SCOPE

- More routes can be added and the same analysis can be expanded to major airports and travel routes in India.
- The analysis can be done by increasing the data points and increasing the historical data used. That will train the model better giving better accuracies and more savings.
- More rules can be added in the Rule based learning based on our understanding of the industry, also incorporating the offer periods given by the airlines.
- Developing a more user friendly interface for various routes giving more flexibility to the users.

Competing Interests

We declare that we have no significant competing financial, professional or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

Author's Contribution

The work is a product of the intellectual environment of the whole team; and that all members have contributed in various degrees to the analytical methods used, to the research concept, and to the experiment design along with writing the manuscript.

REFERENCES

- 1]. Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose¹, Giridhar Maji, Narayan C. Debnath, Soumya Sen. [2]. S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017. [3]. Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015. [4]. Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, "An innovative neural network approach for stock market prediction", 2018. [5]. Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning. [6]. Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction. [7]. Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 -

Stock Market Prediction Using Machine Learning. [8]. V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India - Stock Market Prediction Using Machine Learning.



CHANDIGARH
UNIVERSITY

Discover. Learn. Empower.