**Mini Project Title: IBM-AIML**

**Team Details:**

**1. Sathvik Reddy (22102A040043)**

**2. N.VinithaShri(22101A010314)**

**3. K.M.H.K.V.S.PAVAN(22101A010482)**



**AGENDA-REVIEW 1 (12.02.2024)**

* Abstract
* Introduction
* Problem Statement
* Objectives
* Literature Survey
* Limitations of Existing System
* Workflow of the Proposed System
* References

**ABSTRACT**

**Stock Price Predictor using Linear Regression: A Method for Accurate Forecasting**

* The problem at hand is to develop a stock price predictor to forecast future stock prices, which is crucial for investors and traders to make informed financial decisions.
* The chosen method for solving this problem is linear regression, a statistical modelling technique that analyses historical stock price data and identifies patterns and relationships to make predictions.
* The process begins with collecting and pre-processing relevant data, including historical stock prices and potentially correlated variables such as market indices, financial indicators, and news sentiment.
* Linear regression is applied to establish a mathematical relationship between the input variables and the target variable, which is the stock price.
* The model is trained using historical data to learn the patterns and relationships between the variables, enabling it to make accurate predictions.
* Once the model is trained, it can be used to forecast future stock prices by inputting new or unseen data points.

The purpose of this project is to provide investors and traders with a reliable tool for stock price prediction , allowing them to anticipate market trends and make well-informed investment

decisions.

**INTRODUCTION**

Stock market investments have become a crucial aspect for companies to acquire capital and expand their businesses, as well as for investors to make profitable returns.However, the rapidly changing stock prices and the dynamic nature of the stock market make it challenging for investors to predict the future value of a stock.This is where the need of accurate stock price prediction becomes essential.

Linear regression is a widely used supervised machine learning algorithm in predictive analysis and has been used in several studies for stock price prediction.The algorithm establishes linear relationships between independent and dependent variables and is compatible with continuous mathematical variable values.The stock market prediction model using linear regression trains with a training dataset and then makes predictions.

This project aims to develop a stock price predictor using regression, specifically linear regression. The project will involve cleaning, preprocessing, and preparing the dataset, training the linear regression model, and evaluating the accuracy of the model. The performance of the linear regression model will be compared to other machine learning algorithms to determine its suitability for stock price prediction

**PROBLEM STATEMENT**

# The problem number of the project is 53 as specified by IBM.

* + **Problem Statement is Stock price prediction.The method used to solve the problem statement is linear regression.Developing a model that analyses historical stock price data and establishes a mathematical relationship using linear regression to accurately predict future stock prices.**

**OBJECTIVES**

* Even though there are many types of regression the best, efficient and suitable type of regression is linear regression which gives an accurate prediction.
* Linear regression:Linear regression is a supervised machine learning algorithm used for predictive analysis, specifically for predicting a continuous target variable based on one or more independent variables. The algorithm establishes a linear relationship between the independent variables and the target variable, where the target variable is the dependent variable and the independent variables are the predictors.
* The linear regression model can be represented mathematically as:
* y = b0 + b1\*x1 + b2\*x2 + ... + bn\*xn + e, where:
* y is the target variable (dependent variable).
* x1, x2, ..., xn are the independent variables.
* b0, b1, b2, ..., bn are the coefficients of the model.
* e is the error term.
* The goal of linear regression is to find the values of the coefficients that best fit the data, such that the sum of the squared errors is minimised. Once the coefficients are determined, the model can be used to make predictions on new data.
* It assumes a linear relationship between the independent variables and the target variable and may not be suitable for complex, non-linear relationships.

**LITERATURE SURVEY**

* + Note: Specify links for datasets if available. Mention atleast 10 references. First reference should be your base paper. You can include more than 1 base paper as reference)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Paper Title** | **Journal / Conference details** | **Methods Proposed** | **Datasets Used** | **Limitations** |
| 1. | An improved | In Machine |  |  |  |
|  | multi-output | Learning for |
|  | gaussian process | Healthcare |
|  | rnn with real-time | Conference |
|  | validation for |  |
|  | early sepsis |  |
|  | detection |  |

**LIMITATIONS OF EXISTING SYSTEM**

# While linear regression is a simple and widely used algorithm in stock price prediction, it has several limitations that may affect its accuracy and suitability for complex problems. Here are some of the limitations of linear regression for stock price prediction:

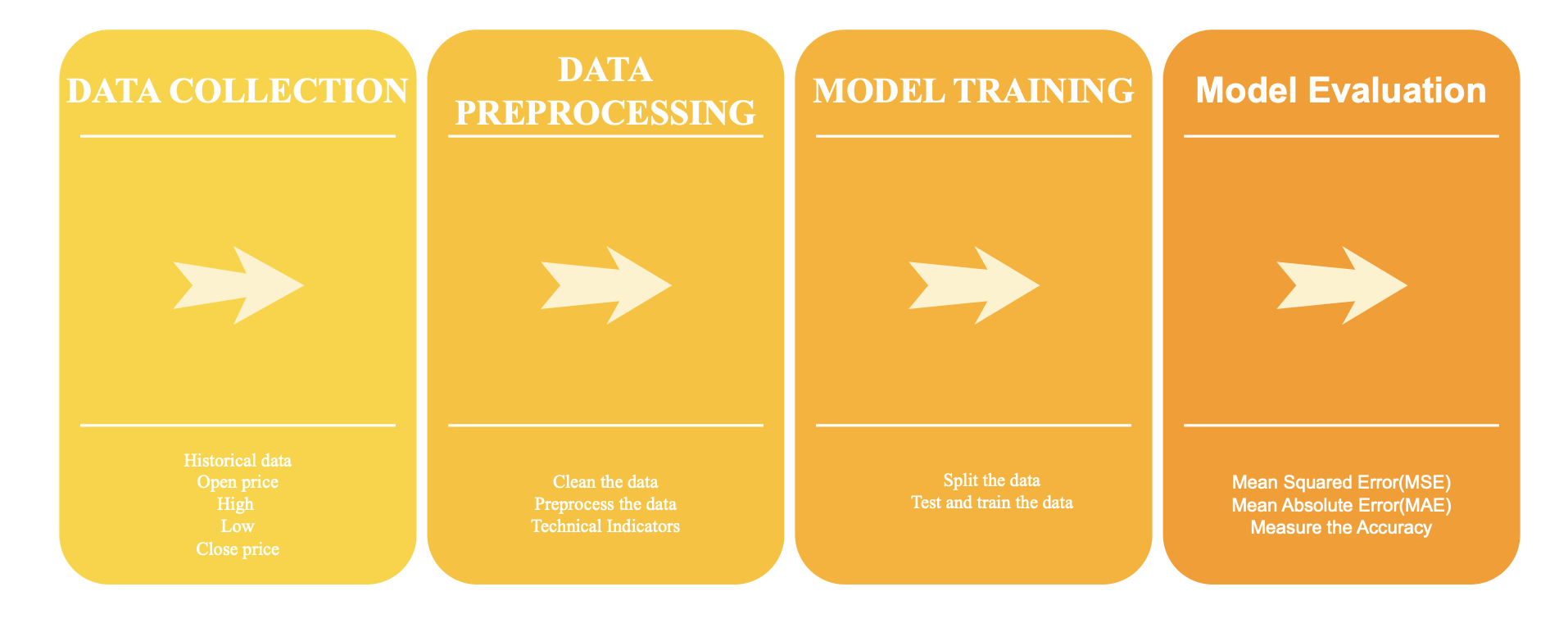
* + **Linear Relationship Assumption: Linear regression assumes a linear relationship between the independent variables and the target variable. However, stock prices are influenced by many factors, and their relationship may not always be linear.**

# Linear regression can suffer from overfitting, especially when the number of independent variables is large. Overfitting occurs when the model fits the training data too closely and fails to generalise to new data.

* + **Non-Stationary Data: Stock prices are non-stationary, which means that their distribution and behaviour change over time. Linear regression assumes stationary data, which can lead to poor predictions.**

# Non-Linear Relationships: Linear regression may not be suitable for predicting stock prices when there are non-linear relationships between the independent variables and the target variable.

**WORKFLOW OF PROPOSED SYSTEM**



**REFERENCES**

* + **Cakra, Y. E., & Trisedya, B. D. (2015, October). Stock price prediction using linear regression based on sentiment analysis. In 2015 international conference on advanced computer science and information systems (ICACSIS) (pp. 147-154). IEEE.**
  + **Groda, B., & Vrbka, J. (2017). Prediction of stock price developments using the Box-Jenkins method. In SHS Web of Conferences (Vol. 39, p. 01007). EDP Sciences.**

