**TESLA SHARE PREDICTION PROJECT DOCUMENTATION**

## ****1. Introduction****

The **Tesla Stock Prediction Project** aims to analyze Tesla Inc.’s stock market behavior using data-driven techniques, including statistical analysis, machine learning models, and visualization tools. The objective is to develop predictive models for stock price movements and evaluate their effectiveness.

### ****1.1 Motivation & Scope****

Stock market predictions are crucial for investors and analysts. Tesla Inc., a leading company in the EV sector, has highly volatile stock prices, making predictions valuable for decision-making. Using machine learning models enhances traditional forecasting methods by identifying complex patterns within the dataset.

## ****2. Dataset Overview****

The dataset consists of Tesla Inc. (TSLA.csv) historical stock prices, with attributes:

* **Date**: Trading date
* **Open**: Opening stock price
* **High**: Highest stock price of the day
* **Low**: Lowest stock price of the day
* **Close**: Closing stock price
* **Adj Close**: Adjusted closing price after corporate actions
* **Volume**: Number of shares traded

### ****2.1 Data Preprocessing****

To ensure accuracy and completeness, the following preprocessing steps were implemented:

* **Handling Missing Values:** Used .fillna() and .mean(numeric\_only=True) to manage missing data.
* **Feature Engineering:** Added daily returns (pct\_change), moving averages (rolling mean), and volatility measures.
* **Scaling & Standardization:** Applied normalization for consistency across models.

## ****3. Data Visualization & Exploratory Data Analysis****

### ****3.1 Key Visualizations****

* **Stock Price Trends:** A line plot of Tesla’s closing prices to observe long-term trends.
* **Daily Returns Analysis:** A graph demonstrating volatility over time.
* **Distribution of Daily Returns:** Histogram representing stock return patterns.
* **Moving Averages (20-Day & 50-Day):** Identifying short-term vs. long-term trends.
* **Correlation Heatmap:** Understanding relationships among stock attributes.

### ****3.2 Observations****

* Tesla’s stock exhibits significant volatility.
* Moving averages provide trend indicators for potential price movements.
* Correlation analysis highlights dependencies between variables such as Open, High, and Close prices.

## ****4. Machine Learning Models Implemented****

### ****4.1 Decision Tree Classifier****

* **Objective:** Classifies whether stock prices increased or decreased.
* **Features Used:** Open, High, Low, Volume.
* **Results:** Accuracy of **50.67%**, moderate precision-recall balance.

### ****4.2 K-Nearest Neighbors (KNN) Classifier****

* **Objective:** Classifies stock trends based on nearest data points.
* **Results:** Accuracy **44.08%**, reflecting moderate predictability.

### ****4.3 Naive Bayes Classifier****

* **Objective:** Probabilistic model for stock price movement classification.
* **Results:** Accuracy **49.32%**, indicating mixed efficiency.

### ****4.4 Random Forest Classifier****

* **Objective:** Uses multiple decision trees for improved prediction.
* **Results:** **64.69% accuracy**, proving effective.

### ****4.5 Principal Component Analysis (PCA)****

* **Objective:** Reduces dimensionality while retaining essential stock data insights.

### ****4.6 Linear Regression Model****

* **Objective:** Predicts stock closing price based on historical values.
* **Results:** **RMSE = 2.04**, **R² Score = 99.96%**, demonstrating strong performance.

## ****5. Model Evaluation & Comparisons****

A comparison of the models based on their accuracy and prediction performance:

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy** | **Best Used For** |
| Decision Tree Classifier | 50.67% | Stock movement classification |
| KNN Classifier | 44.08% | Pattern-based classification |
| Naive Bayes Classifier | 49.32% | Probabilistic predictions |
| Random Forest Classifier | 64.69% | Robust classification |
| Linear Regression | 99.96% | Numerical stock price prediction |

### ****5.1 Observations****

* **Random Forest Classifier** had the highest classification accuracy.
* **Linear Regression** performed exceptionally well in predicting stock prices.
* Further tuning of hyperparameters might enhance the performance of KNN and Naive Bayes models.

## ****6. Conclusion & Future Scope****

### ****6.1 Summary****

This project provides valuable insights into Tesla’s stock behavior using machine learning approaches. The **Random Forest Classifier** showed the highest accuracy in classification, while **Linear Regression** demonstrated strong predictive capabilities for price forecasting.

### ****6.2 Future Scope****

Potential future improvements include:

1. **Deep Learning Approaches:** Incorporating LSTM models for better sequential forecasting.
2. **Sentiment Analysis:** Using news and social media sentiment to enhance predictions.
3. **Real-Time Stock API Integration:** Accessing live stock data for dynamic forecasting.
4. **Multi-Stock Analysis:** Comparing Tesla stock trends with competitors like Rivian and Lucid Motors.

## ****7. References****

* **Tesla stock dataset (**TSLA.csv**)**
* **Scikit-learn, pandas, matplotlib, seaborn libraries**
* **Financial market research studies**