**Stock Market Forecasting Report**

**1. Introduction**

**Objective**

The objective of this project is to develop a predictive model for stock price movements with an accuracy of at least 80%. Using machine learning techniques, we analyze historical stock data, engineer relevant features, and optimize model performance.

**Dataset Overview**

* The dataset contains 600 rows of stock price data with the following features:
  + timestamp (datetime)
  + open\_price (float)
  + high\_price (float)
  + low\_price (float)
  + close\_price (float)
  + volume (integer)

**2. Data Preprocessing**

**Exploratory Data Analysis (EDA)**

* Checked for missing values and found none.
* Verified the dataset for duplicates and found none.
* Set the timestamp column as the index.
* Dropped the symbol column since it contained only a single value.

**Feature Engineering**

* Created technical indicators:
  + **Simple Moving Averages (SMA)**: sma\_5, sma\_10
  + **Other Features Considered**: ema\_5, ema\_10, rsi
* Selected the most important features based on feature importance analysis:
  + open\_price, high\_price, close\_price, sma\_5, sma\_10

**3. Model Selection & Implementation**

**Machine Learning Models Used**

We experimented with the following models:

1. **Logistic Regression**
2. **Support Vector Machine (SVM)**
3. **Decision Tree Classifier**
4. **Random Forest Classifier**
5. **XGBoost Classifier**
6. **K-Nearest Neighbors (KNN)**

**Hyperparameter Tuning**

For each model, we performed hyperparameter tuning using GridSearchCV to optimize accuracy.

**4. Performance Evaluation**

**Best Model: SVM**

After testing multiple models, **SVM (Support Vector Machine)** achieved the highest accuracy:

**SVM Accuracy: 88.98%**

| **Metric** | **Class 0** | **Class 1** | **Macro Avg** | **Weighted Avg** |
| --- | --- | --- | --- | --- |
| Precision | 0.85 | 0.93 | 0.89 | 0.89 |
| Recall | 0.93 | 0.85 | 0.89 | 0.89 |
| F1-Score | 0.89 | 0.89 | 0.89 | 0.89 |

* **Confusion Matrix Analysis:** The model correctly classified most stock movement trends.
* **Feature Importance Analysis:** The most influential features were close\_price, open\_price, and sma\_5.

**5. Model Saving & Deployment**

To ensure reusability, we saved the trained SVM model using joblib:

import joblib

joblib.dump(best\_svm, 'svm\_model.pkl')

**6. Conclusion & Future Work**

**Key Takeaways**

* **SVM was the most effective model**, achieving **88.98% accuracy**.
* **Feature Engineering played a crucial role**, with sma\_5 and sma\_10 significantly improving performance.
* **Hyperparameter tuning** was essential in optimizing model accuracy.

**Future Improvements**

* Incorporating **deep learning models** like LSTMs to capture time-series patterns.
* Testing **ensemble models** combining multiple ML techniques.
* Expanding dataset coverage to include **market sentiment analysis** and additional indicators.