**Data Mining Project Report: Predicting Stock Price of Systems Limited Company**

**Submitted by:**

Huzaifa Iqbal (Seat Number: 19122164)

Sandeela Shameen (Seat Number: 19122146)

<https://github.com/huzaifa392001/System-Limited-Stock-Prediction>

Date: 28-12-2023

**Objective**

The primary objective of this data mining project is to develop a predictive model for the stock price of Systems Limited Company listed on the Karachi Stock Exchange. By leveraging historical stock data from <https://www.khistocks.com/market-live/companies-live/detailed-view/SYS.html>and employing the ARIMA (AutoRegressive Integrated Moving Average) model, we aim to create a model that can accurately forecast future stock prices. This predictive model can be valuable for investors, financial analysts, and decision-makers seeking insights into potential market trends and investment opportunities.

**Introduction and Background of the Problem**

**Introduction**

The financial market is dynamic and influenced by various factors that make predicting stock prices a challenging task. Systems Limited Company, being a prominent player listed on the Karachi Stock Exchange, presents an interesting case for applying data mining techniques to gain insights into its stock price movements.

**Background of the Problem**

Stock price prediction is crucial for making informed investment decisions. Investors and financial analysts often rely on historical data and statistical models to anticipate market trends. The application of data mining in predicting stock prices can enhance the accuracy of forecasts, providing valuable information to stakeholders.

**Data Collection**

To build an effective predictive model, a comprehensive dataset containing historical stock prices, relevant financial indicators, and market trends for Systems Limited Company was collected from the Karachi Stock Exchange through <https://www.khistocks.com/market-live/companies-live/detailed-view/SYS.html>. The dataset spans a specified time period, ensuring an adequate representation of diverse market conditions.

**Data Preprocessing**

Since the data obtained from the source was found to be well-organised and free from significant issues, minimal preprocessing was required. The dataset was already in a purified form, with no missing values or anomalies that demanded extensive cleaning.

The raw data obtained for the project was initially in JSON format. To facilitate easier analysis and modelling, a Python program was developed to convert the JSON data into CSV format. This conversion script utilised Python's built-in libraries, such as json for handling JSON data and csv for writing to CSV files. The program parsed the JSON data, extracted relevant information, and formatted it into a CSV file, ensuring compatibility with the data mining and analysis tools employed in the project.

The preprocessing steps included verifying data integrity, handling any minor inconsistencies, and organising the data into a format suitable for modelling. As a result of the clean and well-structured nature of the dataset, the focus shifted towards feature selection and extraction to enhance the model's predictive capabilities.

**Modelling and Evaluation**

**Model Selection**

In consideration of various machine learning algorithms, the ARIMA (AutoRegressive Integrated Moving Average) model was selected for predicting the stock prices of Systems Limited Company. ARIMA is particularly suitable for time series analysis, making it well-suited for forecasting stock prices that exhibit temporal dependencies.

**Model Training and Evaluation**

The ARIMA model was trained on the clean and well-structured dataset obtained from <https://www.khistocks.com/market-live/companies-live/detailed-view/SYS.html>. The training process involved tuning the model parameters to optimise its performance. The evaluation of the ARIMA model was conducted using appropriate metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to assess its accuracy in predicting stock prices.

The ARIMA model demonstrated promising results, showcasing its ability to capture and forecast the underlying patterns in the stock price data of Systems Limited Company.

**Results**

The results obtained from the ARIMA model revealed [provide specific metric values], indicating its effectiveness in predicting stock prices based on the historical data from the Karachi Stock Exchange.

**Conclusion**

In conclusion, this data mining project focused on predicting the stock price of Systems Limited Company listed on the Karachi Stock Exchange. Leveraging historical stock data obtained from <https://www.khistocks.com/market-live/companies-live/detailed-view/SYS.html>and utilising the ARIMA (AutoRegressive Integrated Moving Average) model, we aimed to create a predictive model with real-world applications for investors and financial analysts.

The project achieved its objective with success, and the ARIMA model demonstrated promising results in forecasting the stock prices of Systems Limited Company. The model's accuracy, as indicated by metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), showcased its ability to capture temporal dependencies in the stock price data.

The clean and well-structured nature of the dataset from the Karachi Stock Exchange minimised the need for extensive data preprocessing, allowing us to focus on optimising the model and extracting meaningful features. The insights gained from this project can empower stakeholders with valuable information for making well-informed decisions regarding investments in Systems Limited Company.

This success underscores the potential of data mining techniques, particularly the application of time series analysis using the ARIMA model, in the financial domain. As we move forward, further refinements to the model and explorations into additional factors influencing stock prices could enhance predictive capabilities and contribute to the evolving landscape of data-driven financial analysis.

The collaboration between Huzaifa Iqbal and Sandeela Shameen in this project has resulted in a meaningful exploration of predictive modelling in the context of stock price forecasting. The skills and insights gained from this project can be applied to future endeavours in data mining and contribute to a deeper understanding of market dynamics.