**< Stock Prices Prediction>**

**Submitted for**

**DATA VISULIZATION AND DASHBOARD**

Submitted by:

**(E23MCAG0037) DEEPAK GARG**

**(E23MCAG0020) NISHKA MEHLAWAT**

**(E23MCAG0091) AMISHU NEGI**

Submitted to-

**Dr. Nitin Arvind Shelke**

**July-Dec 2023**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

A close-up of a logo

Description automatically generated

**INDEX**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **S.no.** | **Title** | **Page No.** | | **1.** | **Abstract** | **1** | | **2.** | **Introduction and Related work** | **1** | | **3.** | **Software used** | **4** | | **4.** | **Methodology** | **5** | | **5.** | **Experimental Result** | **6** | | **6.** | **Conclusion** | **11** | | **7.** | **References** | **12** | |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **LIST OF FIGURES**   |  |  |  | | --- | --- | --- | | **Figure No.** | **Title** | **Page No.** | | **Fig 1.** | **Methodology** | 5 | | **Fig 2.** | **Pair plot of stock price return of different banks** | **7** | | **Fig 3.** | **Morgan Stanley return using distplot** | **7** | | **Fig 4.** | **Citigroups stock price return using distplot** | **8** | | **Fig 5.** | **BAC and GS stock price return** | **8** | | **Fig 6.** | **Stock price return of each bank using line plot** | **9** | | **Fig 7.** | **30 days average stock price of BAC using lineplot** | **9** | | **Fig 8.** | **Correlation heatmap of different banks** | **10** | | **Fig 9.** | **Clustermap of different bank stock price** | **10** |     **LIST OF TABLES**   |  |  |  | | --- | --- | --- | | **Table No.** | **Title** | **Page No.** | | **1.1** | Summarization table for literature review | 3 |  1. **ABSTRACT**   Building a comprehensive tool for investors, analysts, and financial professionals requires a complex process that integrates multiple components, such as an advanced dashboard and data visualization of stock prices. The methodical process of collecting data, which involves gathering historical or current stock price information from reliable financial sources, is an essential first step. The subtleties of stock price movements are best communicated using visualization, which uses tools like heatmaps, pair plots, distplots, and subplots to provide a thorough picture of market trends. The analysis is made more thorough using technical indicators like Bollinger Bands, relative strength index (RSI), and moving averages. All these graphic components work together to provide a complete understanding of the dynamics of stock prices. |  | |
| 1. **INTRODUCTION AND RELATED WORK**   The market value of publicly traded companies is reflected in stock prices, which represent the perceived value of each individual share within that company. These prices are ultimately set by the complex dance of supply and demand in financial markets. Good expectations for a company's future drive demand for its stock, which in turn drives up the price of the stock. On the other hand, negative attitudes may cause demand to decrease, which in turn may cause the stock price to fall.  By making shares available to the public, companies can raise capital through the stock market. In return, investors could profit from dividends and capital growth as well as the company's success. Because stock prices are dynamic, they constantly change during trading hours in response to new information and activity in the market. For market participants, real-time stock prices, historical data, and thorough visualizations are essential resources that help them navigate the complex world of financial markets and make well-informed investment decisions.  (Sharma et al., 2008) proposed a framework examine time-series data and investigate every possibility to make an accurate stock price prediction. We discovered in our research that Recurrent Neural Nets (RNN) are specifically utilized for learning patterns and sequences. Because they are networks with loops, the information can stay in them and be accurately memorized. Because of their vanishing gradient descent issue, recurrent neural nets are unable to learn from historical data as would be predicted. The solution to this issue was found in Long-ShortTerm Memory Networks, or LSTMs for short. These RNNs can recognize long-term dependencies.  (Islam et al., n.d.) introduce a framework the study also contrasts the long short-term memory (LSTM) model, convolutional neural network, auto-regressive integrated moving average (ARIMA), and moving average in terms of stock price prediction. ARIMA is found to perform better than the other algorithms. To evaluate the effectiveness of algorithms, a series of experiments were carried out. Several metrics, including mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE), were used to analyze the algorithms.  (Li et al., 2023) In this paper, we propose a clustering-enhanced deep learning framework to predict stock prices using three well-established deep learning forecasting models, such as Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU), with the goal of optimizing the accuracy of stock price prediction. The clustering is regarded by the proposed framework as a forecasting pre-processing that can enhance the training model quality.  (Vijh et al., 2020) Promising outcomes have also been observed in certain neural network-based techniques, including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Deep Neural Networks (LSTM). (Modi et al., n.d.) An R and Hadoop-based big data framework is used for the processing. and the RMSE values of stock items are used to calculate the accuracy. (Alshammari et al., 2022) When the actual direction rate of return is stable and predicts stability, the confusion matrix shows the highest value; when it is stable and predicts an upward direction, it shows the lowest value. 52.73 is the accuracy of the polynomial kernel function of SVM's confusion matrix.  (K, 2022) The results showed that the decision tree model was not as accurate as generalized linear regression, random forest, and linear regression. When naive Bayes and logistic regression are applied to the data texture, the accuracy ratios range from 77% to 80%. For further research, we advise applying LSTM to deep learning models. Table 1, shows summarization of our literature review and related work.   |  |  |  |  | | --- | --- | --- | --- | | **Author** | **Dataset used** | **Techniques used** | **Model used** | | (Sharma et al., 2008) | Google stock price | Modelling Analysis | ANN | | (Islam et al., n.d.) | k-electric Karachi stock price | Moving average and auto regression | LSTM, CNN,  RMSE, MAE | | (Li et al., 2023) | Us stock price | Clustering enhanced deep learning | RNN, GRU,  WDTW | | (Vijh et al., 2020) | Stock price of 5 companies. | Machine and deep learning | Random forest, ANN | | (Modi et al., n.d.) | Stock Market data | ML and data Mining | SVM | | (Alshammari et al., 2022) | Kuwait stock market | classification | Decision Tree, SVM, logistic regression, Random Forest | | (K, 2022) | Stock prices of various companies. | Machine learning, big data | Linear regression, Decision tree, Random Forest. |   Table1. Summarization of literature review   1. **SOFTWARE USED**   **Excel:** For data analysis and visualization, including stock price visualization, a lot of financial professionals use Microsoft Excel. It can handle historical data and has several charting options.  **Python (Matplotlib, Seaborn, Plotly):** Plotly, Seaborn, and Matplotlib are three popular libraries for data analysis and visualization in Python. Data manipulation can be done with the Pandas library, and a variety of plots and charts can be made with visualization libraries.  **Power BI:** Users can visualize and share insights from their data with this business analytics tool from Microsoft. It is appropriate for analyzing and visualizing data from the stock market because it supports a variety of data sources.   1. **METHODOLOGY**   Developing a solid process to visualize stock price data entails a methodical approach with the goal of producing precise and informative depictions. The process starts with a precise definition of the goals and moves on to find trustworthy data sources. Fig1, shows methodology of our data visualization on stock price of different banks.    **Fig1. Methodology**   1. **Objective:** Clearly state the purpose of the stock price data visualization, including whether it is to compare stocks or to obtain more comprehensive market insights into various businesses. Finding trends, patterns, or abnormalities in stock prices is part of this. 2. **Data Collection and Cleaning:** To ensure the integrity and dependability of the dataset, gather data from Kaggle stock price datasets of different companies and preprocess it by handling missing values and outliers. 3. **Select Visualization tools:** Select the right visualization tools considering programming languages like Python with Matplotlib and Seaborn as well as platforms like Power BI. The choice should consider the team's experience level as well as the intricacy of the visualization specifications. 4. **Visualization techniques:** Depending on the goals, choose visualization techniques. For comparing stocks and learning more about the market, methods like pair plots, distplots, and subplots can be useful. While subplots enable side-by-side comparisons, pair plots highlight relationships between variables and distplots display the distribution of individual stock prices. 5. **Correlation Analysis:** To investigate correlations between various stocks, indices, or economic indicators, use scatter plots or correlation matrices. An overview of variable movements that shows dependencies and interactions is provided by correlation matrices. Correlations are visually represented by scatter plots, which highlight possible patterns. To illustrate the direction and strength of correlations using color gradients, think about utilizing heatmaps. 6. **Interactive features:** Use tools such as Plotly and Power BI to incorporate interactive features that will increase user engagement. Incorporate tooltips to offer more details, filters to concentrate on data points, and drill-down capabilities for in-depth analyses of patterns or periods. By enabling users to extract customized insights, these features guarantee an intuitive user experience. 7. **EXPERIMENTAL RESULTS**   A visual depiction of the data distribution and pairwise correlations through the investigation of relationships between the daily returns of various stocks in the dataset. To suit your analysis, change the date range and tickers as necessary. Fig2, shows pairplot of returns in different companies.    **Fig2. Shows pair plot of stock price return in different banks.**  the distribution of Morgan Stanley stock's daily returns over the given time, offering information on the stock's volatility and return trends. Fig3, shows Morgan Stanley return using distplot.    **Fig3. Morgan Stanley return using distplot**  providing insights into the volatility and return patterns during that period by visualizing the distribution of daily returns for Citigroup stock. Fig4, shows Citi group return using distplot.    **Fig4. Shows stock price return of Citi group using distplot**  Using subplots, this visualization insight illustrates a return of BAC and GS. Fig5, shows return stock price return of BAC and GS using subplots.    **Fig5. Shows stock price return of BAC ad GS**  It displays the close stock price for each bank over the course of the index of time using a line plot.Fig6, shows stock price of each bank using line plot.    **Fig6. Shows stock price return of each bank using line plot**  the closing prices for Bank of America (BAC) stock over the 30-day rolling average for the given period (January 1, 2008, to January 1, 2009). Fig7, shows 30 day average stock price of BAC using line plot.    **Fig7. Shows 30 days average stock price of BAC using lineplot.**  The correlation between the closing prices of several banks is shown visually in the heatmap. The strength and direction of these correlations can be understood with the aid of the color intensity and the annotated values. Fig8, shows correlation heatmap of different banks.    **Fig8. Shows correlation heatmap of different banks.**  In addition to visualizing the correlations, the cluster map groups the banks according to similarity in price movements. Fig9. Shows cluster map of different bank based on their stock prices.    **Fig9. Shows clustermap of different bank based on stock prices.**   1. **CONCLUSION**   The financial market's complex dynamics have been better understood thanks to the data visualization of stock prices for different banks. The heatmap, clustermap, and individual stock trends included in the visual analysis have revealed interesting patterns and correlations between the closing prices of various banks. The correlation analysis has clarified the interdependencies between banks, guiding potential portfolio diversification strategies through heatmap and cluster map visualizations. Potential groupings based on shared market influences have been further highlighted by the hierarchical clustering. In-depth analyses of individual stock trends, made possible by line plots, have shown differences in volatility, growth periods, and notable outliers, offering a sophisticated insight into the unique market behavior of each bank.   1. **REFERENCES**   Alshammari, B. M., Aldhmour, F., AlQenaei, Z. M., & Almohri, H. (2022). Stock market prediction by applying big data mining. *Arab Gulf Journal of Scientific Research*, *40*(2), 139–152. https://doi.org/10.1108/AGJSR-05-2022-0053  Islam, N., Afzal, M., Wali, M. A., & Shakeel, H. (n.d.). Data Analysis, Visualization and Prediction of Stock Market Prices of K-Electric. In *Pakistan Journal of Engineering and Technology* (Issue 5).  K, Dr. C. K. (2022). STOCK MARKET PREDICTION USING BIG DATA. *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, *06*(12). https://doi.org/10.55041/IJSREM16782  Li, M., Zhu, Y., Shen, Y., & Angelova, M. (2023). Clustering-enhanced stock price prediction using deep learning. *World Wide Web*, *26*(1), 207–232. https://doi.org/10.1007/s11280-021-01003-0  Modi, P., Shah, S., & Shah, H. (n.d.). *Big Data Analysis in Stock Market Prediction*. www.ijert.org  Sharma, A., Modak, S., & Sridhar, E. (2008). Data Visualization and Stock Market and Prediction. *International Research Journal of Engineering and Technology*, 2037. www.irjet.net  Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. *Procedia Computer Science*, *167*, 599–606. https://doi.org/10.1016/j.procs.2020.03.326 |  |
|  |  |