




STOCK PRICE PREDICTION USING RECURRENT NEURAL NETWORKS (RNN)

Bulls & bears bandits – Capstone Project




Business Overview

- Stock price prediction plays an important role in financial market and accurate forecasts can have significant implications for investors and financial institutions.
 - **Objective:** Develop and deploy a stock price prediction model using Recurrent Neural Networks (RNN) that can be used by investors and financial institutions to make profit.
- 



Problem Statement



Traditional stock price prediction methods such as **technical analysis** which is based on historical stock price data and **fundamental analysis** which is based on a company's financial performance and other factors, have not been able to accurately predict the the future stock prices hence need to develop an algorithm that can assist in solving this.



Objectives

Main Objective


- Develop a robust stock price prediction model using recurrent neural networks(RNN).

Specific Objectives

- Build and implement RNN model for stock price prediction then improve the model with LSTM.
- Evaluate the performance of the model using F1 score and accuracy of the models using MAE.
- Use LSTM to forecast stock prices.
- Create a user-friendly dashboard for stakeholders to access predictions.

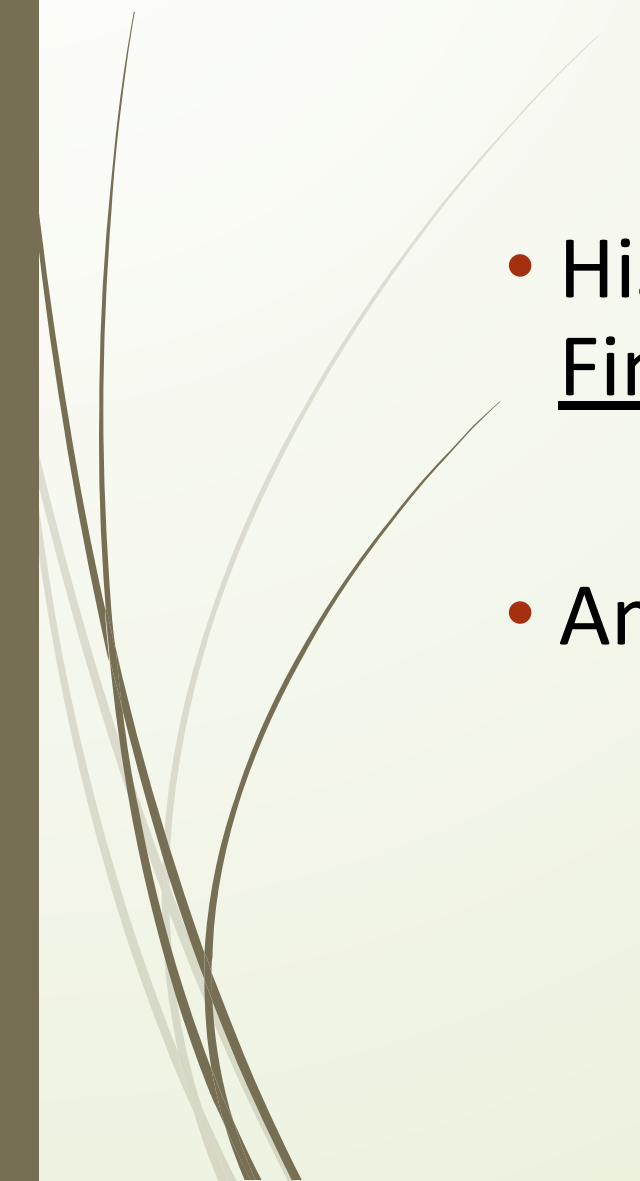


Success Criteria

- ☐ Achieve a low MAE and MSE errors and accurately forecast stock prices.
 - ☐ Ability of the LSTM to capture long-term dependencies in stock price data.
 - ☐ Have the RNN model perform better than our baseline model which uses traditional time-series forecasting method (SARIMAX).
 - ☐ Usability of the models by investors and financial analysts.
- 




Data Inventory

- Historical stock price dataset from Yahoo Finance and Wall Street Journal(WSJ)
 - Amazon (AMZN) or KCB Bank(KCB:NAI).
- 




Assumptions

- ❑ Historical stock price dataset conforms to the data quality characteristics.
 - ❑ LSTM model will be able to learn the long-term dependencies in the stock price data.
 - ❑ Models will be able to predict future stock prices with a high degree of accuracy even on the event of market volatility.
- 



Constraints

- ❑ **Number of stocks analysed** – the project will be limited to analysis of one stock commodity which is AMZ due to time required to complete the project.
 - ❑ **Computational resources** for training and testing complex models. – the project will be implemented using Python and Google Colab.
 - ❑ **Regulatory and ethical considerations** regarding the use of financial data.
- 



Data Mining Goals

- Extract useful information from the historical stock price dataset to train and evaluate RNN model.
- **Specific data mining tasks:**
 - **Feature selection:** Selecting the most important features from the dataset to use for training the models.
 - **Data preprocessing:** Cleaning and transforming the data to make it suitable for training the models.
 - **Model training:** Training the LSTM and RNN models on the preprocessed data.
 - **Model evaluation:** Evaluating the performance of the models on a held-out test dataset.



Data Understanding



□ Data features:

- Date
- Open price
- High price
- Low price
- Close price
- Adjusted Close price
- Volume

- Data will be collected from Yahoo Finance and cleaned and preprocessed to make it suitable for training and evaluation using SARIMAX and RNN models.
- Data will be organized into time series format, with each data point representing a specific trading day or month.
- All data except the date will be in numerical data type format.



Model Development



- Our baseline model (SARIMAX) and RNN model will be implemented using python and appropriate libraries used.
- The models will be tuned to optimize their performance.
- Models will be trained on historical stock price dataset and evaluated on a held-out test dataset.



Model Evaluation

- Models will be evaluated using the following metrics:
 - Mean Absolute Error (MAE) - low.
 - Mean Squared Error (MSE) - low.
 - F1 Score
- The RNN model will be compared to traditional time-series forecasting method (SARIMAX) to determine the efficiency of the new model.



Conclusion



- ☐ Bulls and bears bandits are confident that the project will be completed within the specified time and with the available resources.
- ☐ Bulls and bears bandits are confident that the project will be successful in achieving its objectives.