OPTIMIZING STOCK TRADING STRATEGY WITH K-MEANS CLUSTERING

Mini Project

Big Data Analytics[BDA]

Department of Computer Engineering
TERNA ENGINEERING COLLEGE
Nerul (W), Navi Mumbai 400706

FINAL PRESENTATION

Group Members:

Amey Thakur	**B-50**
Hasan Rizvi	**B-51**
Mega Satish	**B-58**

Under the Guidance of :
Prof. D. M. Bavkar

INDEX

ABSTRACT

- ➔ We propose to determine a single exchange-traded fund (SPY) investing strategy that will maximise our total wealth.

- ➔ We intend to forecast future price fluctuations for a specific stock.

- ➔ We develop a predictor for multiple firms using these trained RL models that forecasts the every day close stock prices.

- ➔ The goal of this project is to learn and get hands-on experience in Data Analytics and Machine Learning.

PROBLEM STATEMENT

- ➔ The main emphasis and objective of our project is to analyse given raw data and do exploratory data analysis in order to fully comprehend and identify patterns.

- ➔ Then, using a Neural Network approach, construct a model and train it to get the desired outcomes.

- ➔ Finally, it will be deployed as a web application.

LITERATURE SURVEY

To get a better understanding of K-Means Clustering and its uses for the stock trading strategy, we referred to the following papers:

- ➔ “Fast K-Means Clustering for Very Large Datasets Based on MapReduce Combined with a New Cutting Method”, by Duong Van Hieu and Phayung Meesad. This paper gives an overview of the importance of the K-Means Clustering Algorithm for Big Data. This paper presents a new approach for reducing the number of iterations of the K-Means algorithm which can be applied to very large dataset clustering.

- ➔ “Enhancing stock prediction clustering using K-means with genetic algorithm”, by E. N. Desokey, A. Badr and A. F. Hegazy. The main objectives of this paper are to optimize the clustering of stock market prediction and to examine the impact of applying genetic algorithm optimization with the k-means clustering algorithm. The evaluation shows that using genetic algorithm and k-means clustering algorithm with Chi-square similarity measure achieved the highest accuracy with the least sum of square distances.

INTRODUCTION

- We've always been captivated by the stock market's seeming unpredictability. So, for a day trader, deciding what to trade is the first and most important step.

- Day traders should look for equities with plenty of liquidity, moderate to high volatility, and a large number of followers. Isolating the present market trend from any surrounding noise and then capitalising on that trend is the key to finding the appropriate stocks for intraday trading.

- It's critical to devise a well-balanced, low-risk plan that will benefit the majority of individuals. One such technique that we propose uses K-means clustering to generate automated trading strategies based on previous data.

METHODOLOGY

K-Means Clustering

- To process the learning data, the K-means algorithm in data mining starts with the first group of randomly selected centroids, which are used as the beginning points for every cluster, and then performs iterative (repetitive) calculations to optimize the positions of the centroids.

- It halts creating and optimizing clusters when either:
  - The centroids have stabilized — there is no change in their values because the clustering has been successful.
  - The defined number of iterations has been achieved.

MODEL DEPLOYMENT

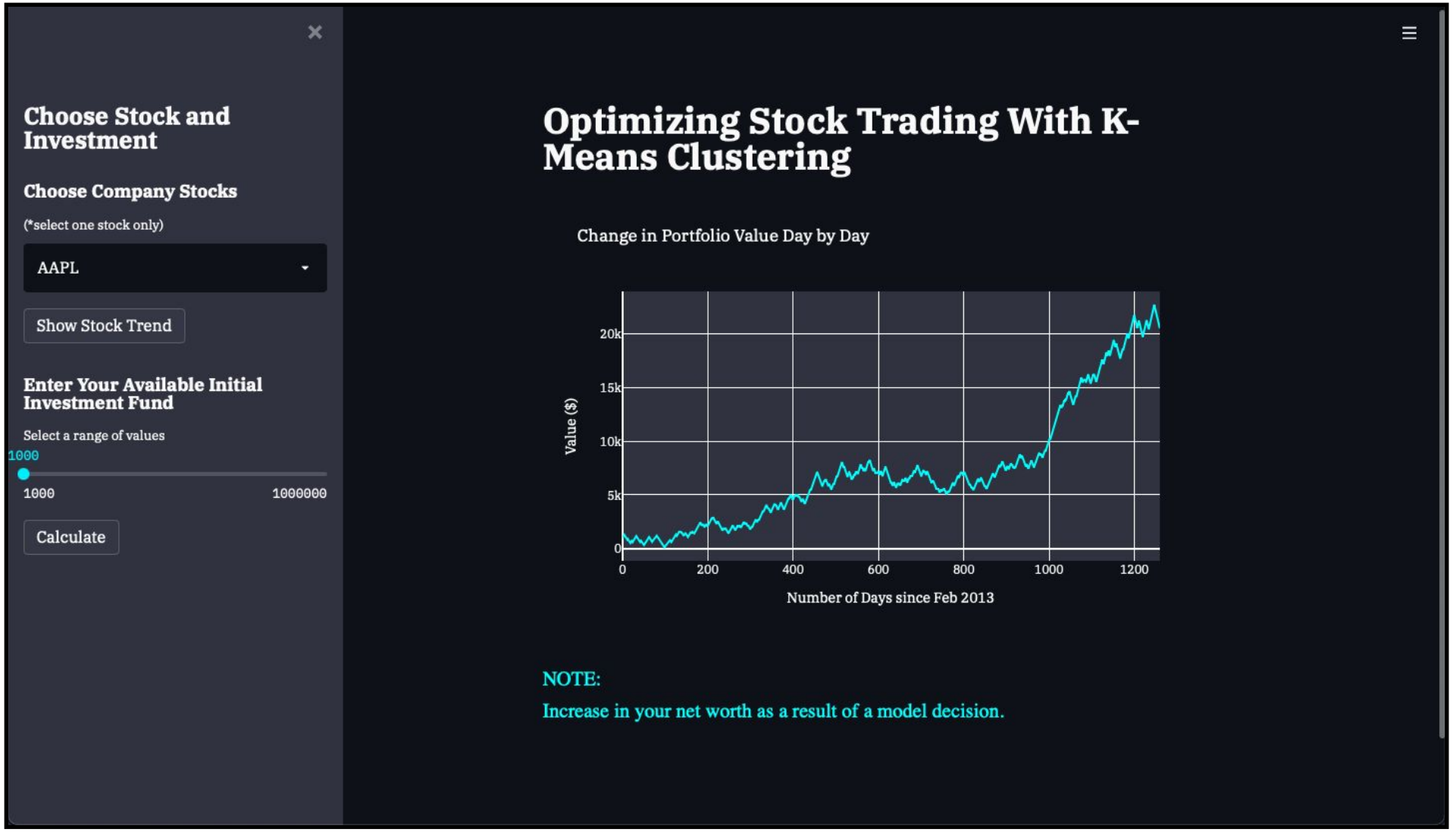
The necessary files for model deployment are the following:

1. app.py
2. requirement.txt - Contains a list of all the dependencies that your code requires in order to function properly.
3. Procfile - In an app, a Procfile is a list of process types.
4. setup.sh

SNAPSHOTS


Choose Stock and Investment
Choose Company Stocks
(*select one stock only)
AAPL
Show Stock Trend
Enter Your Available Initial Investment Fund
Select a range of values
1000
1000
1000000
Calculate
Optimizing Stock Trading With K-Means Clustering
Stock Trend of AAPL
Price ($)
180
160
140
120
100
80
60
2014
2015
2016
2017
2018
Date
NOTE:
Stock is on a solid upward trend. Investing here might be profitable.

SNAPSHOTS


Choose Stock and Investment
Choose Company Stocks
(*select one stock only)
AAPL
Show Stock Trend
Enter Your Available Initial Investment Fund
Select a range of values
1000
1000
1000000
Calculate
Optimizing Stock Trading With K-Means Clustering
Change in Portfolio Value Day by Day
Value ($)
20k
15k
10k
5k
0
0
200
400
600
800
1000
1200
Number of Days since Feb 2013
NOTE:
Increase in your net worth as a result of a model decision.

REFERENCES

1. Segmentation of stock trading customers according to potential value (2004) https://doi.org/10.1016/j.eswa.2003.12.002

2. Customer Profitability Analysis, Cost System Purposes and Decision Making Process: A Research Framework

3. https://www.researchgate.net/publication/325092899_Customer_Profitability_Analysis_Cost_System_Purposes_and_Decision_Making_Process_A_Research_Framework

4. Van Hieu D., Meesad P. (2015) Fast K-Means Clustering for Very Large Datasets Based on MapReduce Combined with a New Cutting Method. In: Nguyen VH., Le AC., Huynh VN. (eds) Knowledge and Systems Engineering. *Advances in Intelligent Systems and Computing*, vol 326. Springer, Cham. https://doi.org/10.1007/978-3-319-11680-8_23

5. E. N. Desokey, A. Badr and A. F. Hegazy (2017) "Enhancing stock prediction clustering using K-means with genetic algorithm," *2017 13th International Computer Engineering Conference (ICENCO)*, pp. 256-261, doi: 10.1109/ICENCO.2017.8289797.

THANK YOU