

Stock Market Prediction

COURSE: BIG DATA SYSTEM ENGINEERING USING SCALA

PROFESSOR: ROBIN HILLYARD

TEAM-8: DISHANK, PURVA

GIT HUB LINK: [HTTPS://GITHUB.COM/PURVABUNDELA/STOCK-MARKET-PREDICTION-USING-SCALA-AND-SPARK](https://github.com/Purvabundela/STOCK-MARKET-PREDICTION-USING-SCALA-AND-SPARK)

Goals of the project

- ▶ Building a reactive application that is based on Actor model using Akka framework and Play Framework.
- ▶ Analyze the volume of stocks traded from the available listing dataset using Spark and MLlib .
- ▶ Suggest user future stock prices based on historical data and twitter sentiment analysis.

Use case

- ▶ User will enter the company name for prediction and model will provide the future stock prediction of the mentioned company.
- ▶ The admin will have access to a dashboard showcasing the summary of all the stock prices.
- ▶ As flexible and scalable design, the application will be responsive to high traffic

Data sources

- ▶ The data source we used was taken from Kaggle which has about 600000 rows for 505 Companies over the period of five years.

kaggle Search kaggle  Competitions Datasets Kernels Discussion Jobs ... Sign In

Reviewed Dataset

S&P 500 stock data
Historical stock data for all current S&P 500 companies

Cam Nugent • last updated 7 months ago

Overview Data Kernels Discussion Activity

Download (21 MB)

Tags finance medium featured

Top Contributors >

Rank	User	Contribution
1st	KshitijaPansare	S&P 500 stock data - Time S... run 6 months ago
2nd	Scotty Kay	S&P 500 Plotting, with detail... run 6 months ago
3rd	SreeJay	S&P Stock Analysis for ABT (... run 20 days ago)

Kernels >

Post	Votes
S&P 500 stock data - Time S...	16
S&P 500 Plotting, with detail...	13
S&P Stock Analysis for ABT (...	8

Discussion >

Post	Votes	Replies
Adjusted Close Price vs Clos...	1	1
R package : Easy loading of...	3	3
data set bug	2	2

3 Files (62.69 MB)

 [all_stocks_1yr.csv](#)

 [all_stocks_5yr.csv](#)

 [individual_stocks...](#)

 [A_data.csv](#)

 [AAL_data.csv](#)

 [AAP_data.csv](#)

 [AAPL_data.csv](#)

 [ABBV_data.csv](#)

 [ABC_data.csv](#)

 [ABT_data.csv](#)

 [ACN_data.csv](#)

 [ADBE_data.csv](#)

 [ADI_data.csv](#)

(505 total)

all_stocks_1yr.csv
768.2 KB • Updated 7 months ago

About this file
All stocks in the S&P 500 index, and their historical prices for the past year

Preview Column Metadata

An error occurred while retrieving the preview.

Methodology

- ▶ Build an reactive application using Play framework and Actor Model
- ▶ Use of Spark and Spark MLlib to come up with the prediction of stock for the user

Milestone

Task	Time
Come up with the Prototype for our reactive System and data cleaning	4/01
Exploring and implementing Actor model and spark and unit tests	4/14
Integrating Spark and Spring	4/20
Functional and load testing /self acceptance	4/22

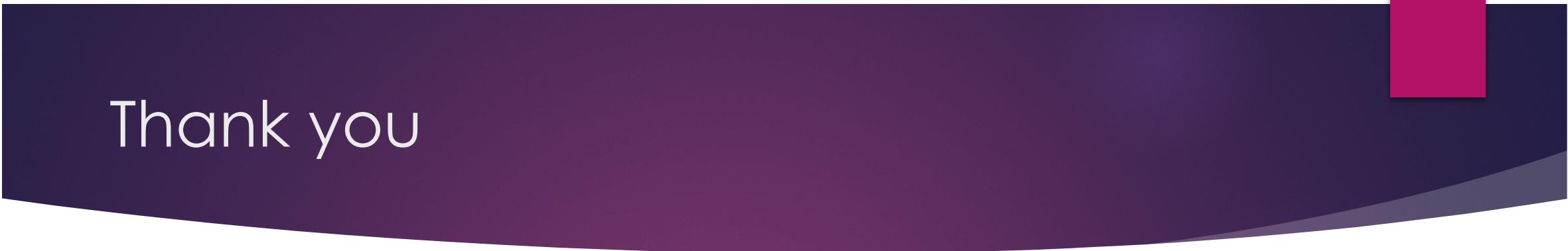
Scala Programming

- ▶ Everything in Scala

- ▶ Akka
- ▶ Spark
- ▶ Spark MLlib
- ▶ Play

Acceptance Criteria

- ▶ Application should be truly reactive and will be able to hold 1000 request, will respond to the result around 10 sec
- ▶ Stock Market Prediction accuracy will provide 50% accuracy of 90% of data



Thank you