Stock Market Prediction

Purva Bundela - bundela.p@husky.neu.edu

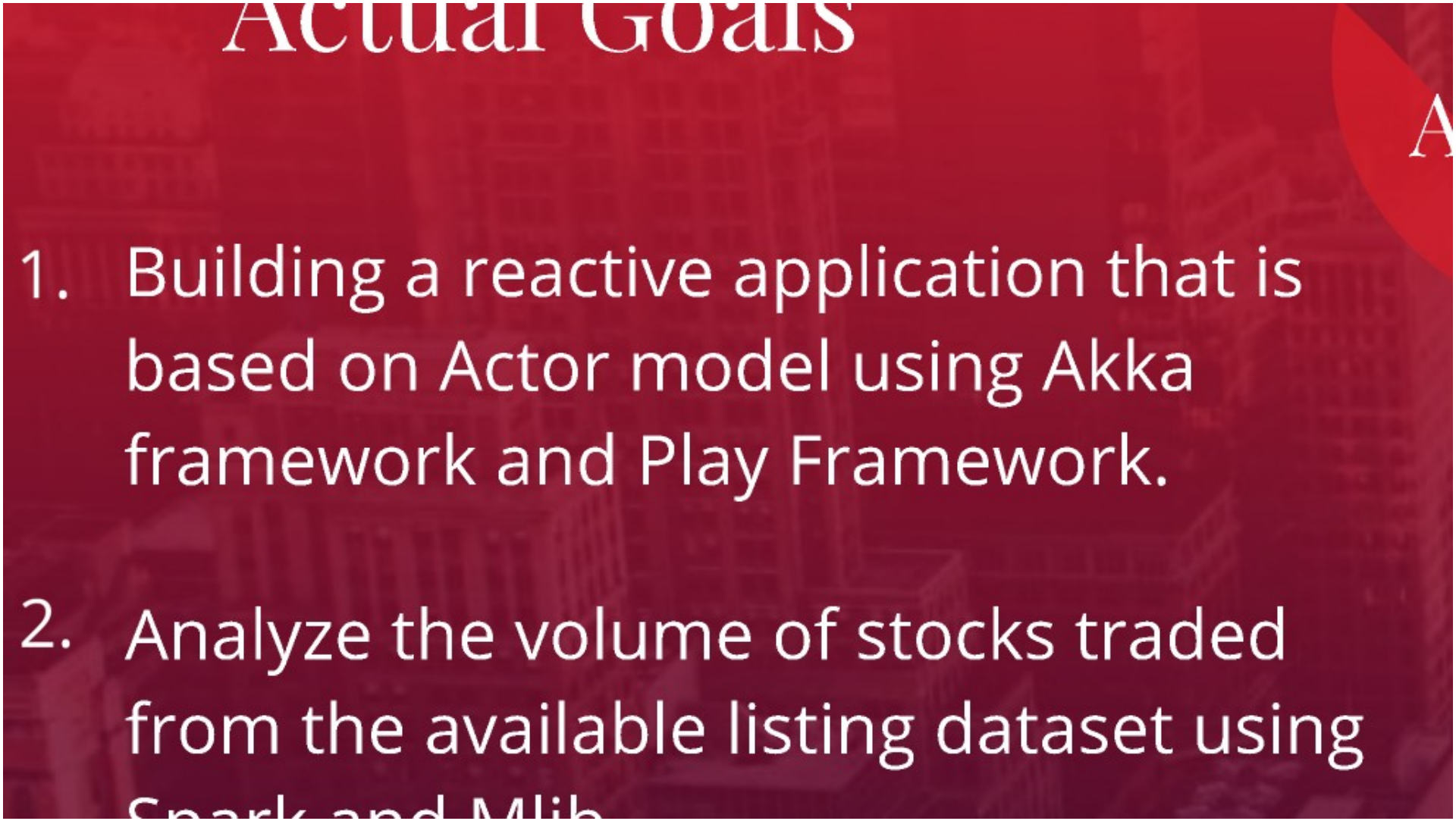
Dishank Shah - shah.di@husky.neu.edu

Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark


Actual Goals
Goals
Achieved


Actual Goals
Goals Achieved
1. Building a reactive application that is based on Actor model using Akka framework and Play Framework.


Actual Goals
1. Building a reactive application that is based on Actor model using Akka framework and Play Framework.


Actual Goals
1. Building a reactive application that is based on Actor model using Akka framework and Play Framework.
2. Analyze the volume of stocks traded from the available listing dataset using

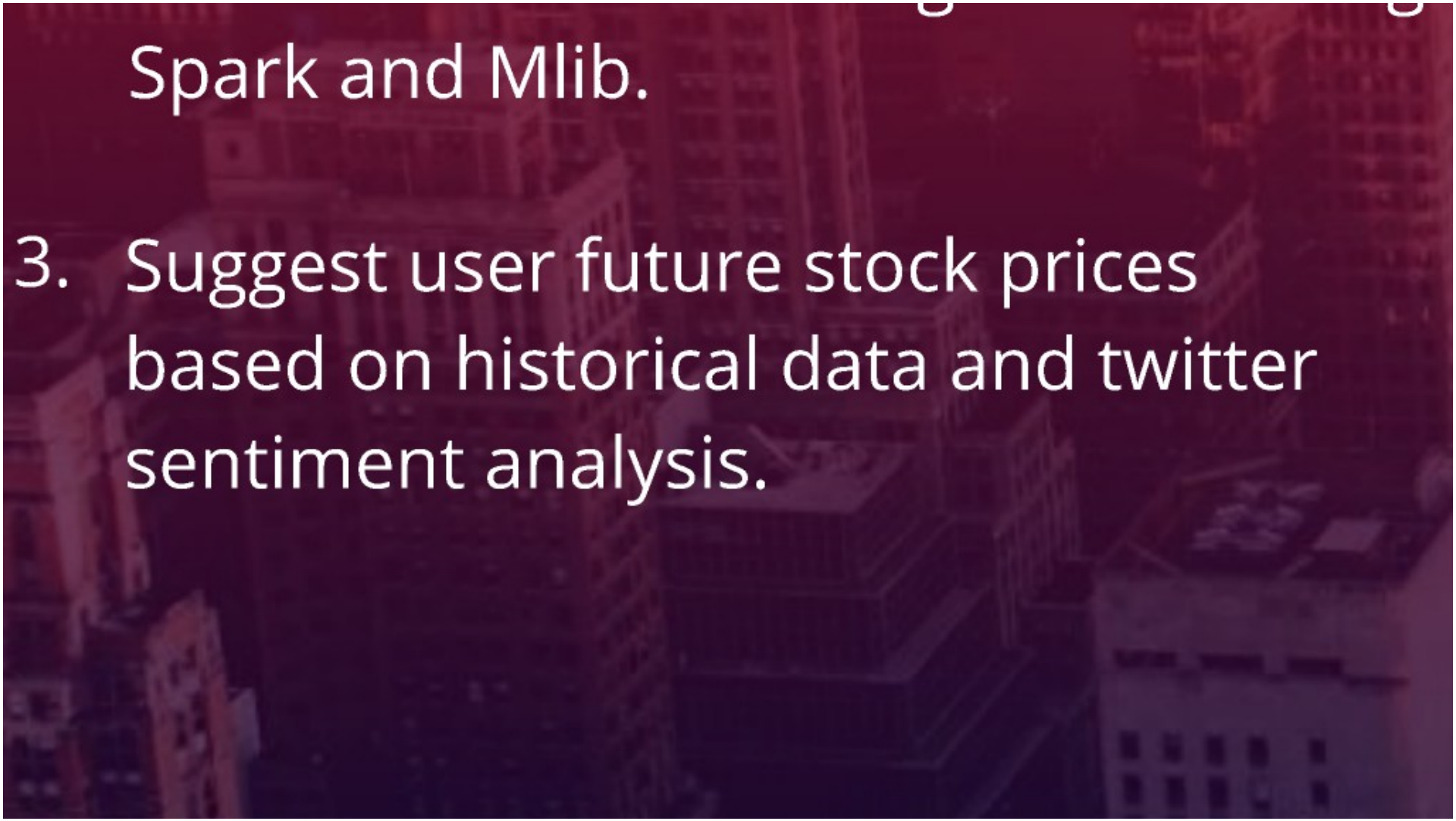
based on Actor model using Akka framework and Play Framework.

2. Analyze the volume of stocks traded from the available listing dataset using Spark and Mlib.

based on Actor model using Akka framework and Play Framework.

2. Analyze the volume of stocks traded from the available listing dataset using Spark and Mlib.

3. Suggest user future stock prices based on historical data and twitter sentiment analysis


Spark and Mlib.
3. Suggest user future stock prices based on historical data and twitter sentiment analysis.


Goals Achieved
Reactive application
Our model takes 12 secs to predict the future values of stock.
Also, the application can handle 1000 requests.
Analysis of price based on dates
Our model is based on timeseries so it can forecast for any number of days.
Suggest Future price
Our model forecast values for 30 days. Also, it will show company stock status based on twitter sentiment analysis.


Reactive application
Our model takes 12 secs to predict the future values of stock.
Also, the application can handle 1000 requests.
Analysis of price based on dates
Our model is based on timeseries so it can forecast for any number of days.
Suggest Future price
Our model forecast values for 30 days. Also, it will show company stock status based on twitter sentiment analysis.

Stock Market Prediction

Purva Bundela - bundela.p@husky.neu.edu

Dishank Shah - shah.di@husky.neu.edu

Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark


Technologies used to perform these use cases
1. Play Framework
2. Spark Timeseries Algorithm
3. Twitter Sentiment Analysis
4. Akka Framework
5. Scala
6. D3.Js
Use Case 1
Use Case 2

Use case 1

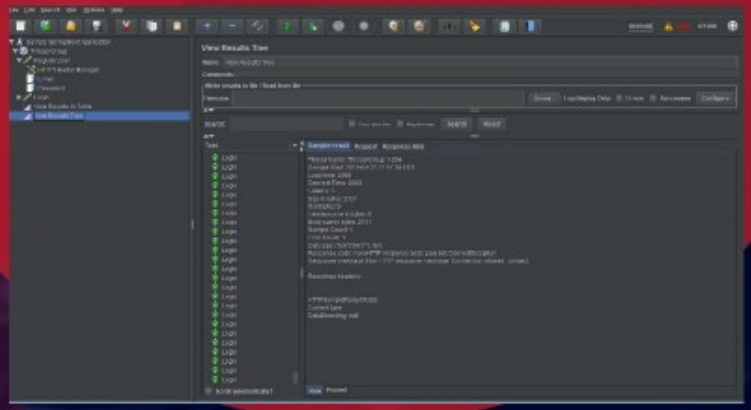
As a

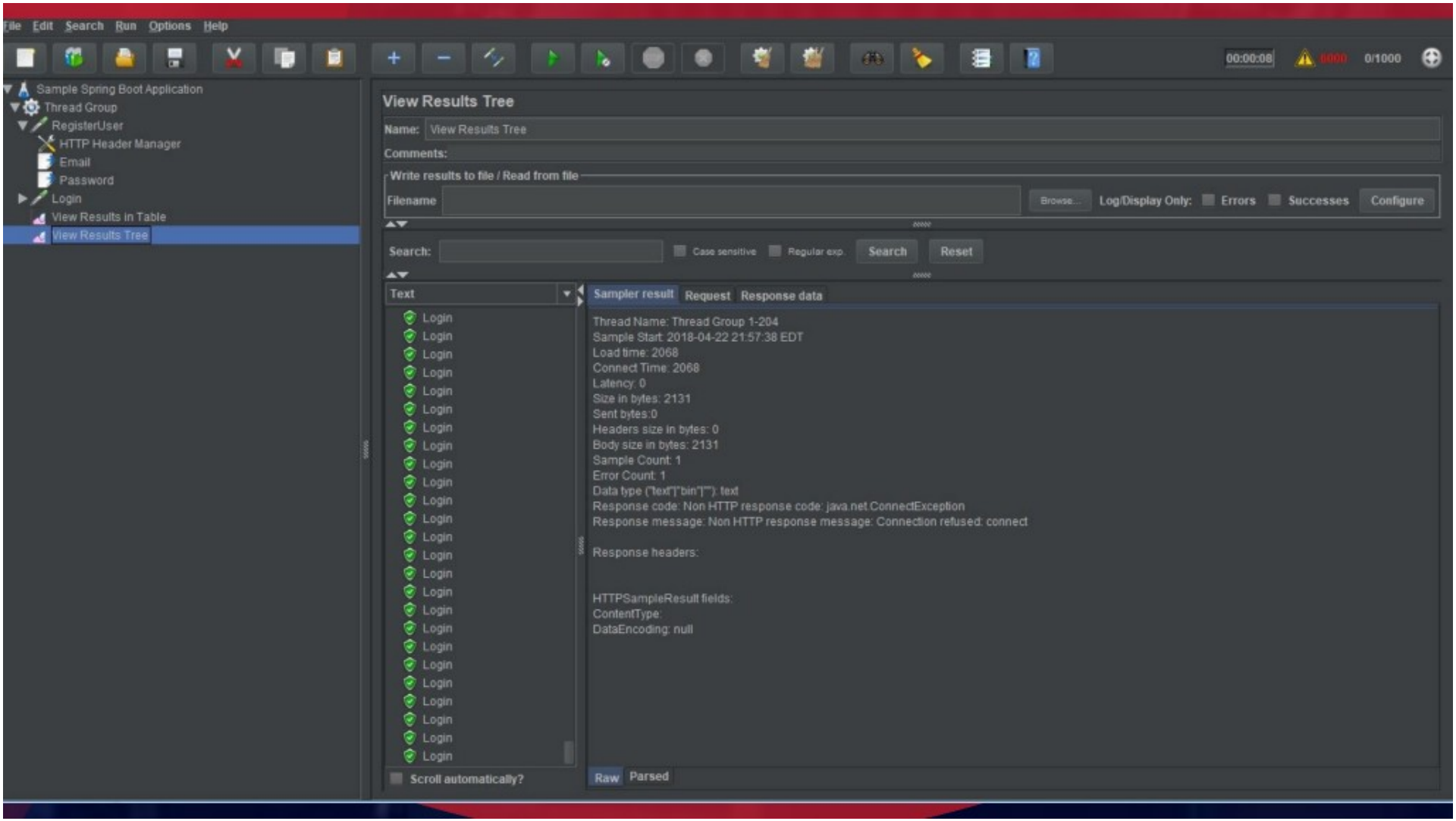
Select the company name from the given list of companies and system will return the prediction.

Use Case 2

As flexible and scalable design, the application will be responsive to high traffic

The application is able to handle 1000 requests at a time



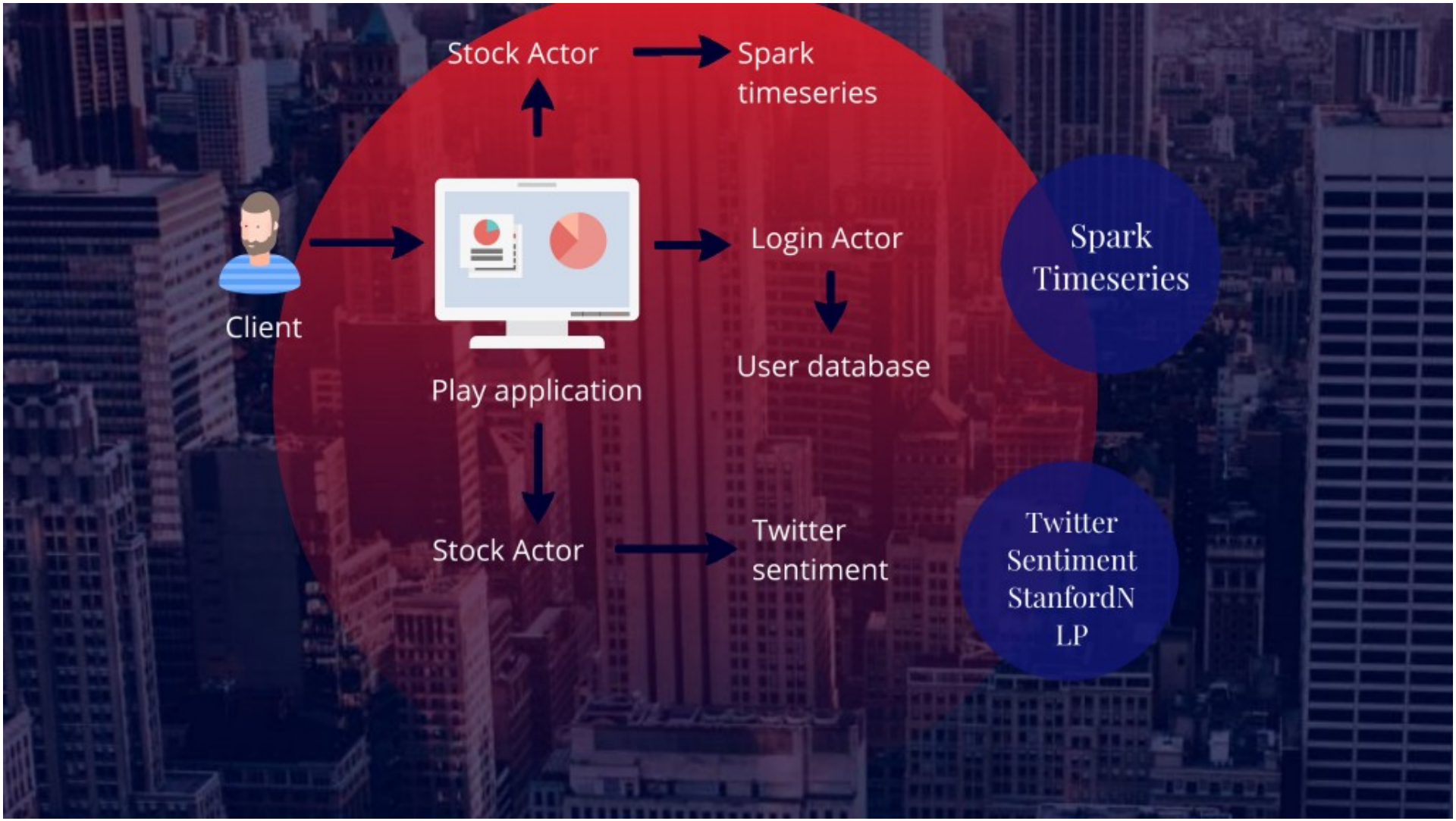

File Edit Search Run Options Help
00:00:08
0/1000
Sample Spring Boot Application
Thread Group
RegisterUser
HTTP Header Manager
Email
Password
Login
View Results in Table
View Results Tree
View Results Tree
Name: View Results Tree
Comments:
Write results to file / Read from file
Filename
Browse...
Log/Display Only: Errors Successes
Configure
Search:
Case sensitive
Regular exp.
Search
Reset
Text
Login
Scroll automatically?
Sampler result
Request
Response data
Thread Name: Thread Group 1-204
Sample Start: 2018-04-22 21:57:38 EDT
Load time: 2068
Connect Time: 2068
Latency: 0
Size in bytes: 2131
Sent bytes:0
Headers size in bytes: 0
Body size in bytes: 2131
Sample Count: 1
Error Count: 1
Data type ("text"|"bin"|""): text
Response code: Non HTTP response code: java.net.ConnectException
Response message: Non HTTP response message: Connection refused: connect
Response headers:
HTTPSampleResult fields:
ContentType:
DataEncoding: null
Raw
Parsed

Stock Market Prediction

Purva Bundela - bundela.p@husky.neu.edu

Dishank Shah - shah.di@husky.neu.edu

Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark


Stock Actor
Spark timeseries
Client
Play application
Login Actor
User database
Spark Timeseries
Stock Actor
Twitter sentiment
Twitter Sentiment StanfordNLP

Twitter Sentiment Analysis

We Followed Stanford NLP to get the sentiments of the sentences

Collecting the Tweets
Using Keywords

Filter the Tweets
Recieved

set up the pipeline, and get the sentiment score for each sentence as integer

0 -> very negative
1 -> negative
2 -> neutral
3 -> positive
4 -> very positive

Spark Timeseries Methodology

A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time

Preprocessing and cleaning the data

Using Scala

Feature extraction and load into RDD

Using Scala

Training the data using ARIMA model and forecasting the output

Using ARIMA and spark

Parse data into desired format and visualize the data

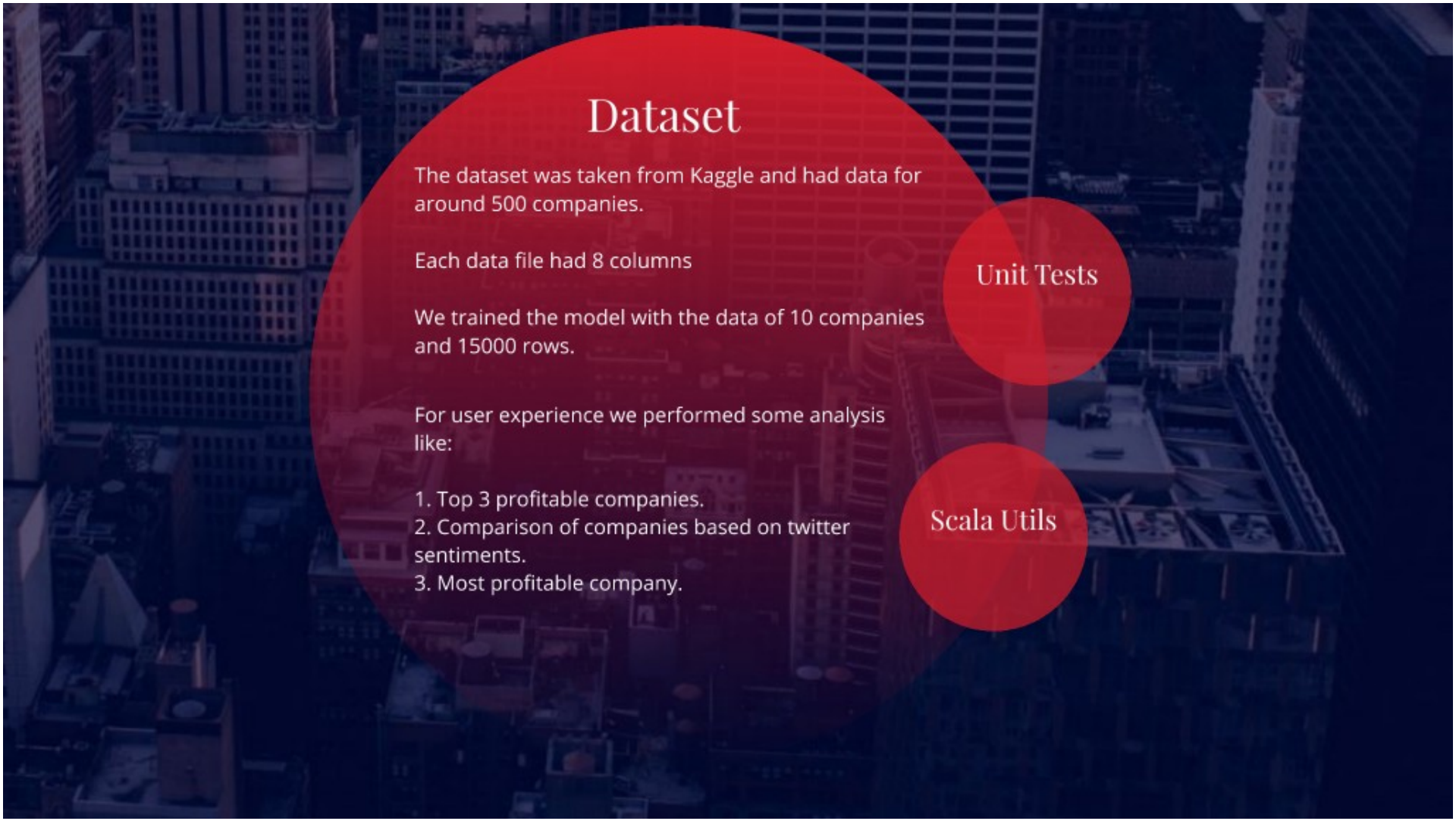
Using D3.js

Stock Market Prediction

Purva Bundela - bundela.p@husky.neu.edu

Dishank Shah - shah.di@husky.neu.edu

Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark


Dataset
The dataset was taken from Kaggle and had data for around 500 companies.
Each data file had 8 columns
We trained the model with the data of 10 companies and 15000 rows.
For user experience we performed some analysis like:
1. Top 3 profitable companies.
2. Comparison of companies based on twitter sentiments.
3. Most profitable company.
Unit Tests
Scala Utils

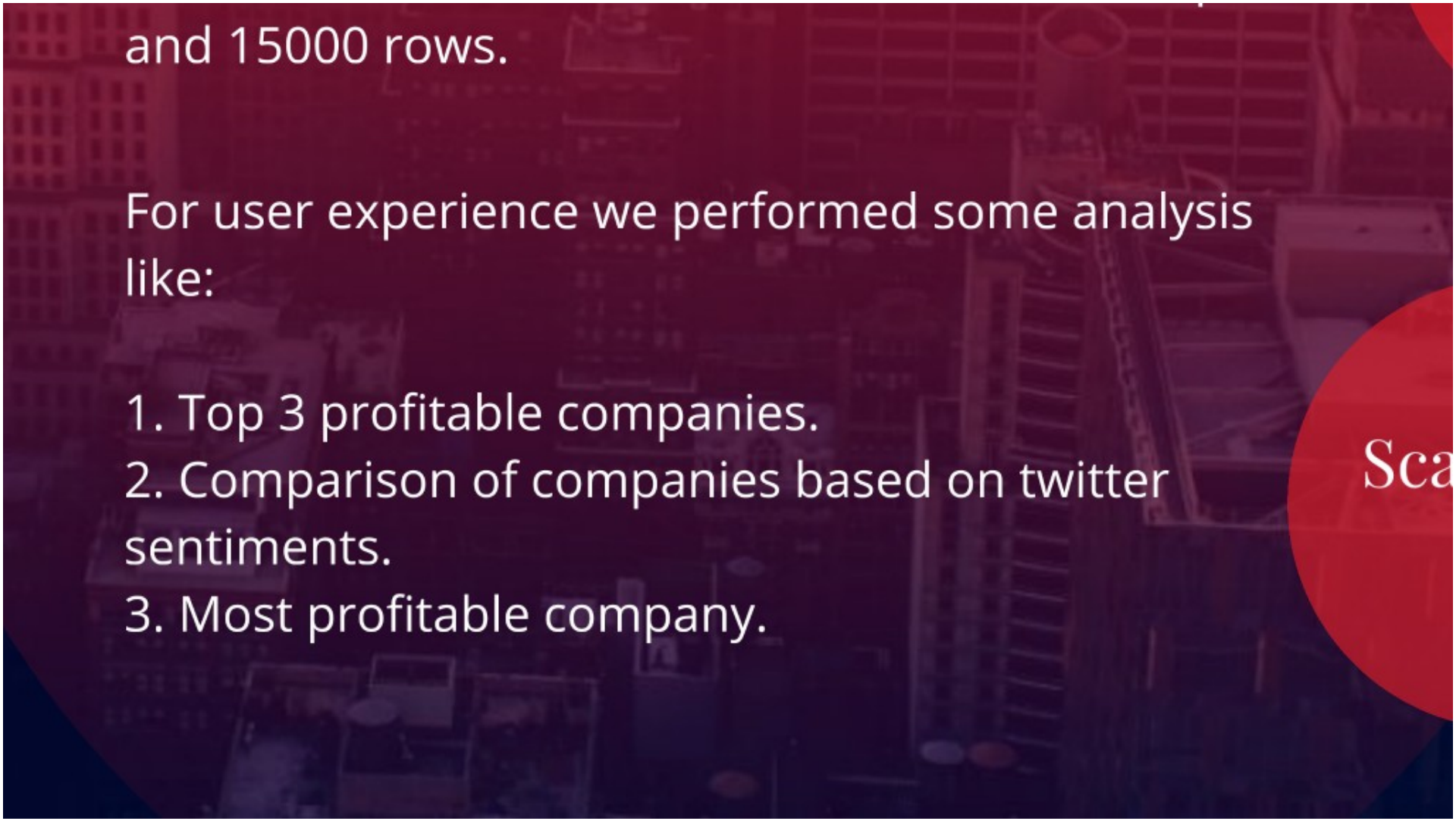
Dataset

The dataset was taken from Kaggle and had data for around 500 companies.

Each data file had 8 columns

We trained the model with the data of 10 companies and 15000 rows.

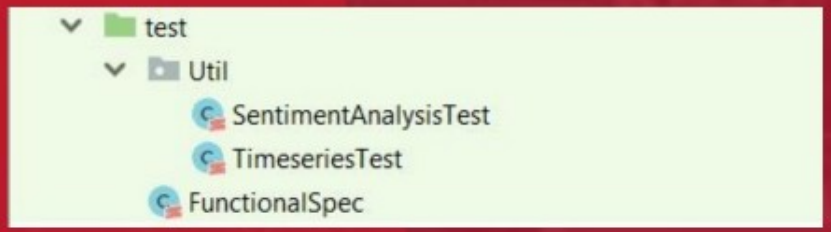
For user experience we performed some analysis like:


and 15000 rows.
For user experience we performed some analysis like:
1. Top 3 profitable companies.
2. Comparison of companies based on twitter sentiments.
3. Most profitable company.
Sca


Scala Utils
Timeseries.scala
Ingest.scala
SentimentAnalysis.scala
Tweet.scala
TwitterRequest.scala
Usecases.scala
Util
Ingest
SentimentAnalysis
Timeseries
Tweet
TwitterRequest
Usecases

Unit Test Coverage

Total 15 test cases.
Around 40% of functions were covered.

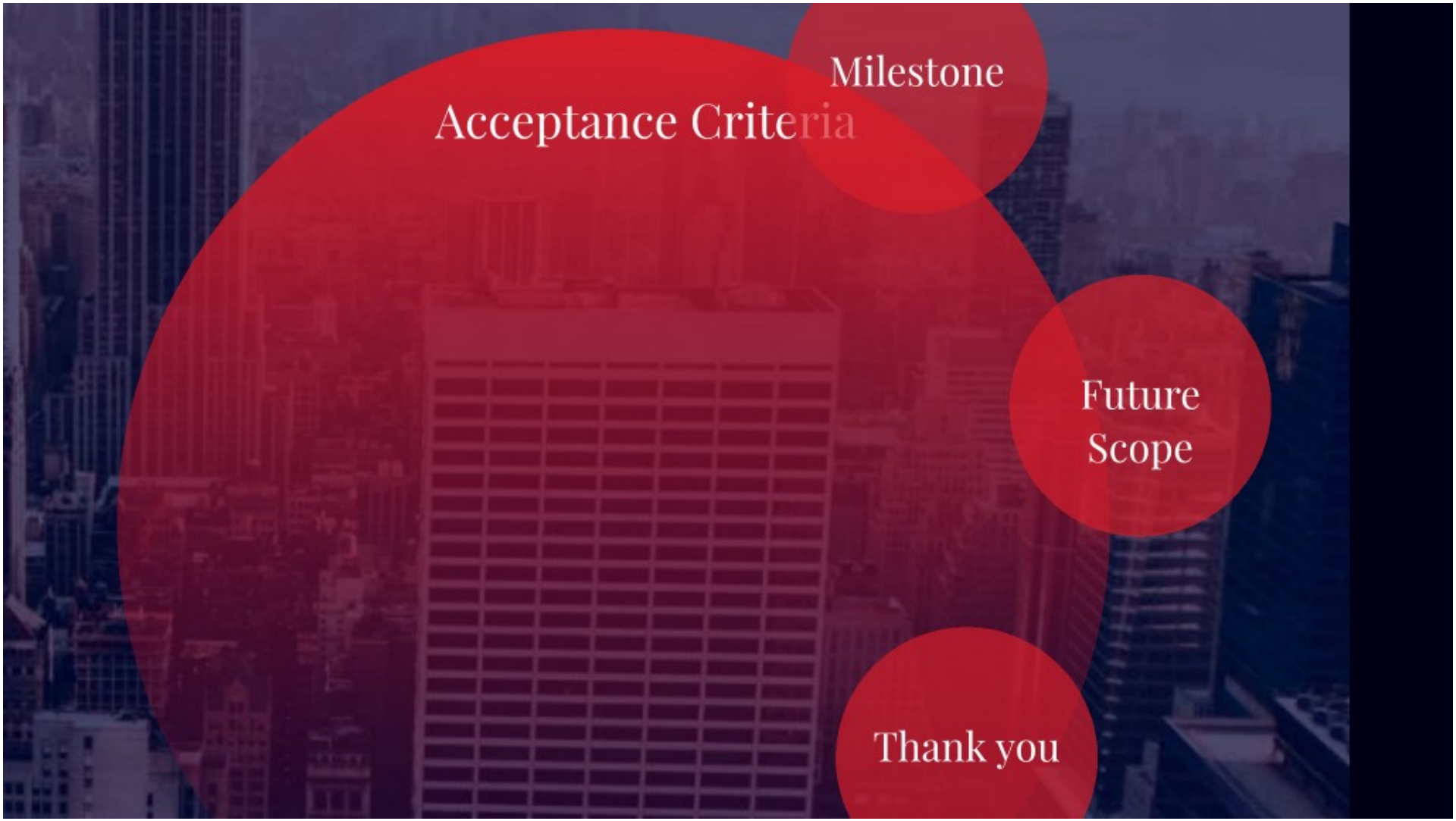


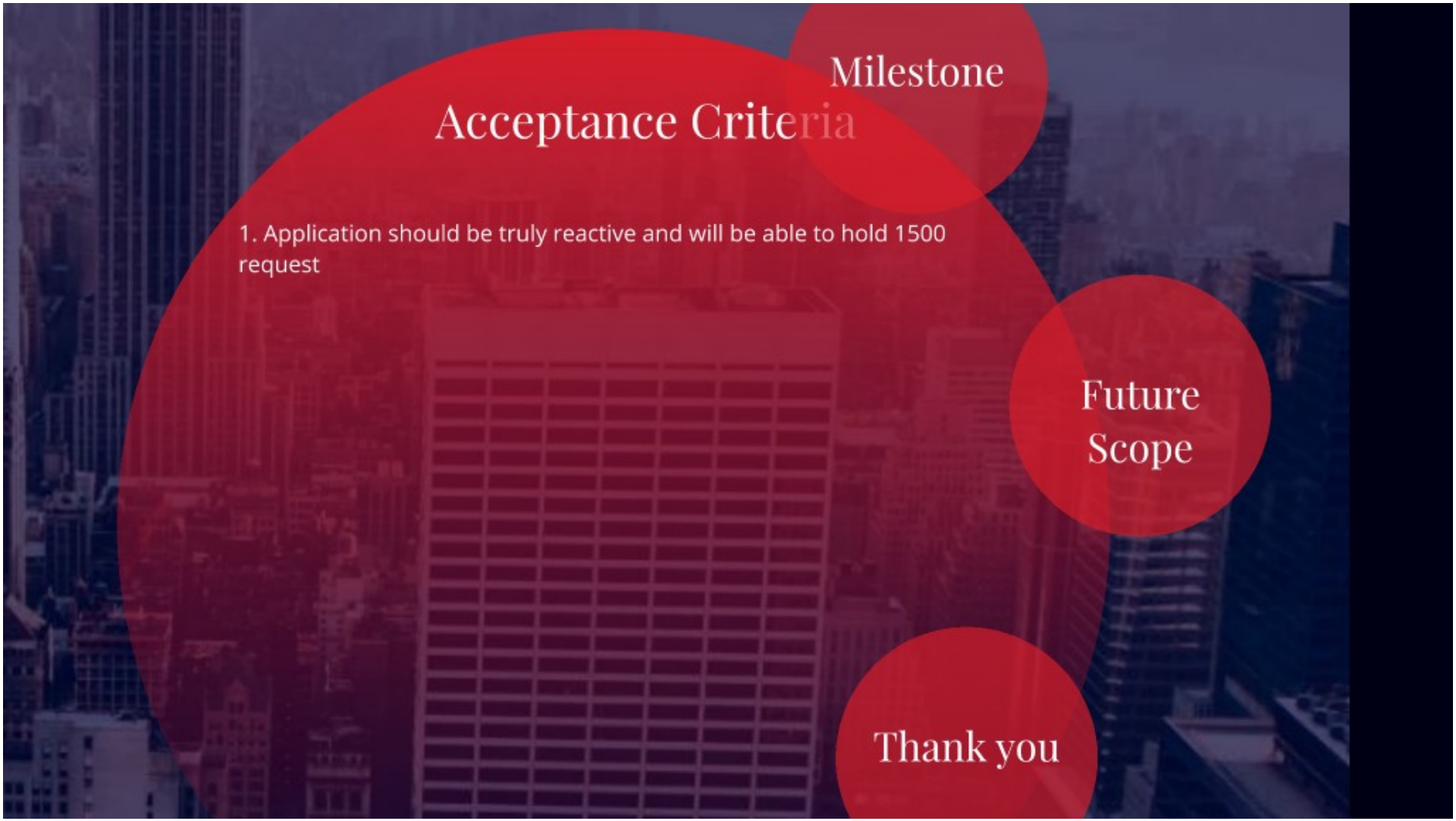
Stock Market Prediction

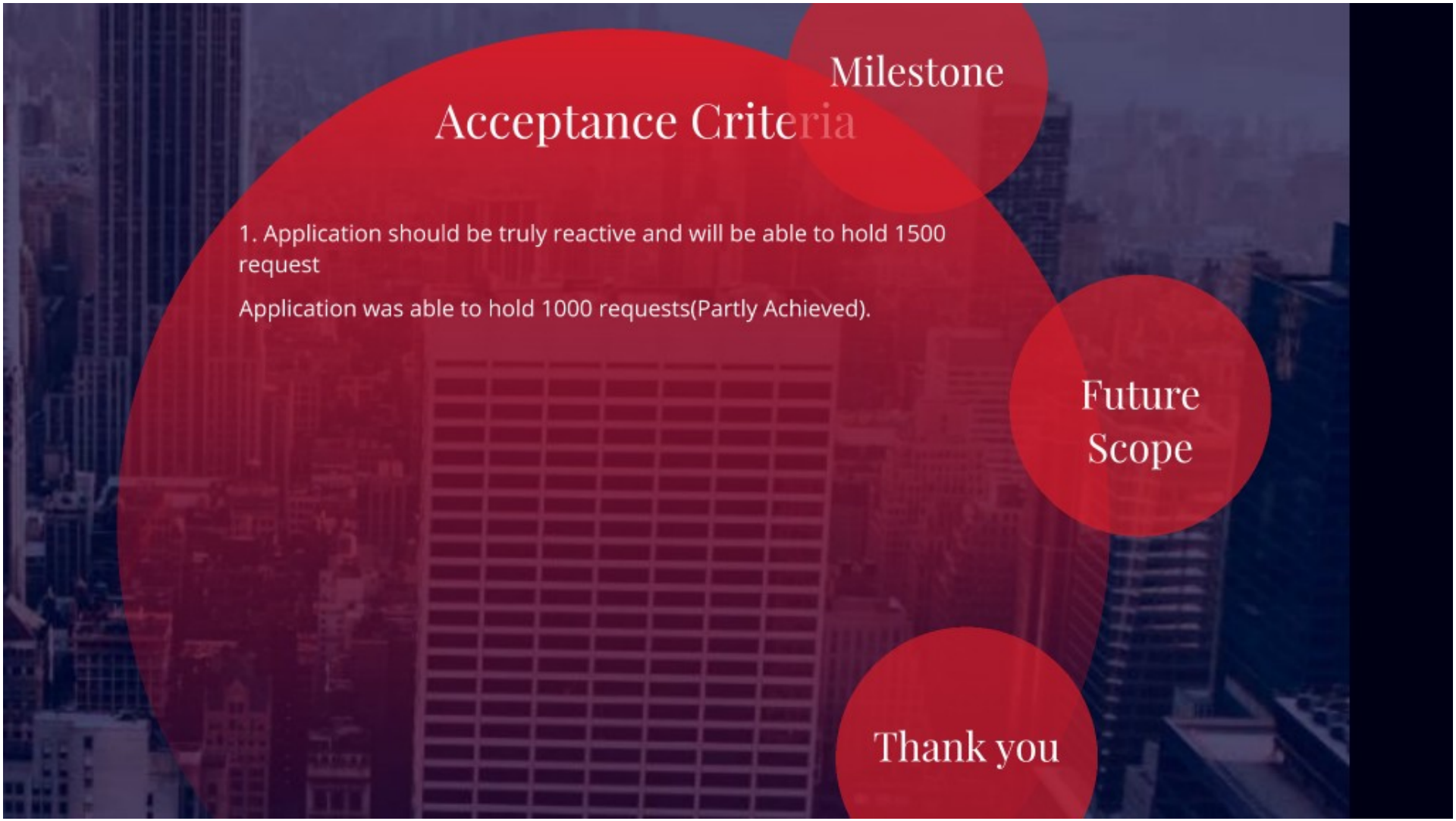
Purva Bundela - bundela.p@husky.neu.edu

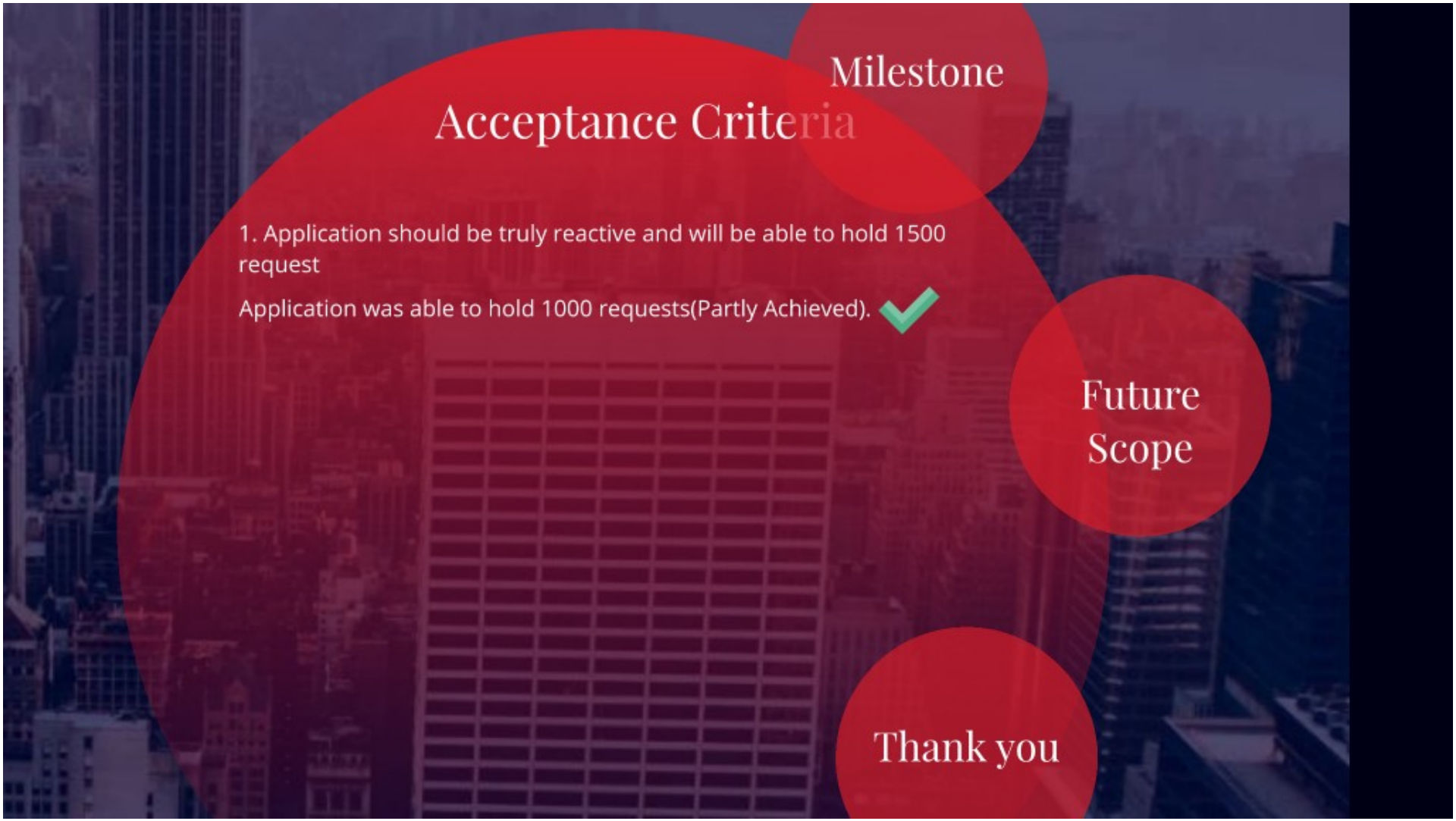
Dishank Shah - shah.di@husky.neu.edu

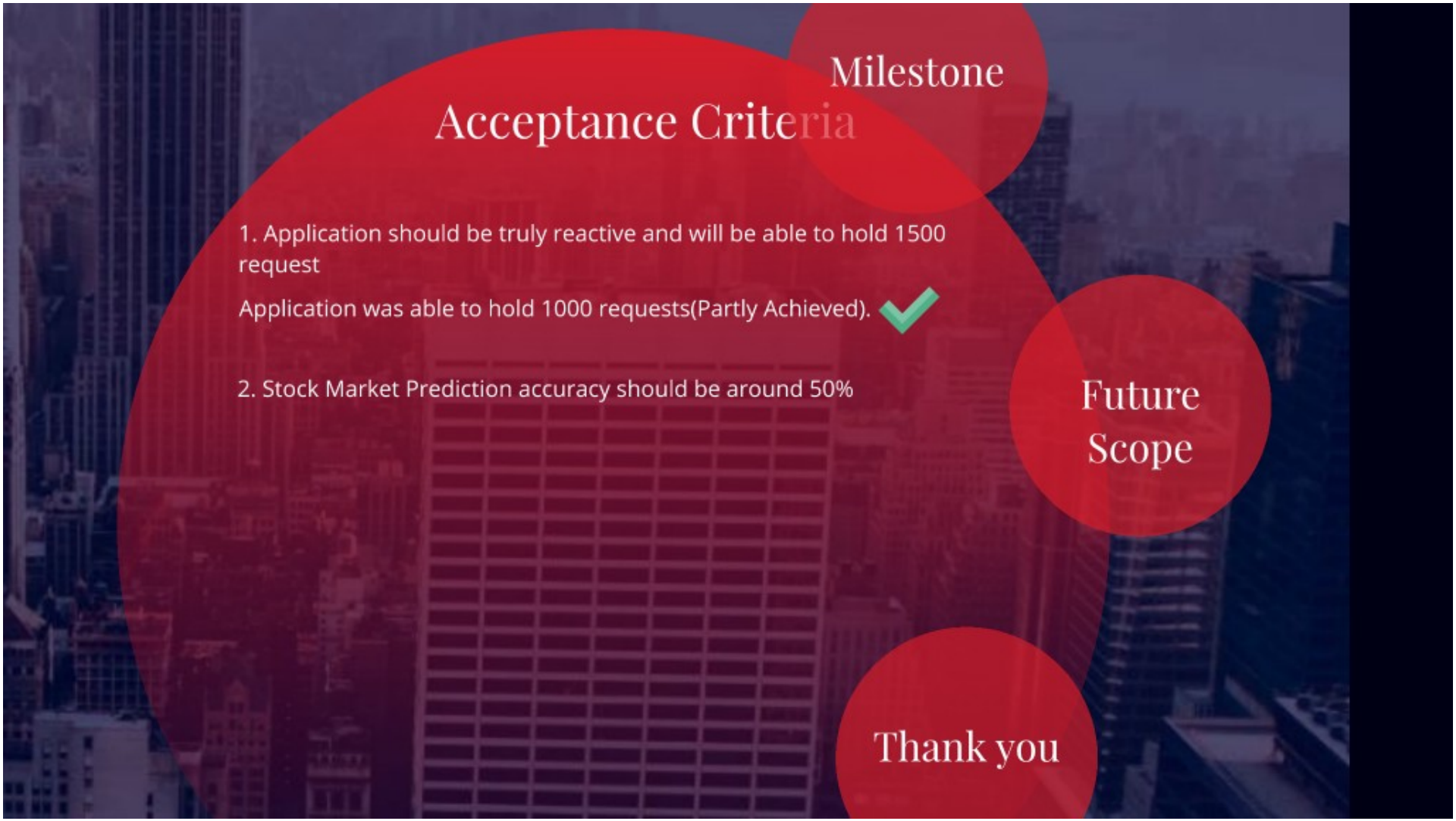
Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark

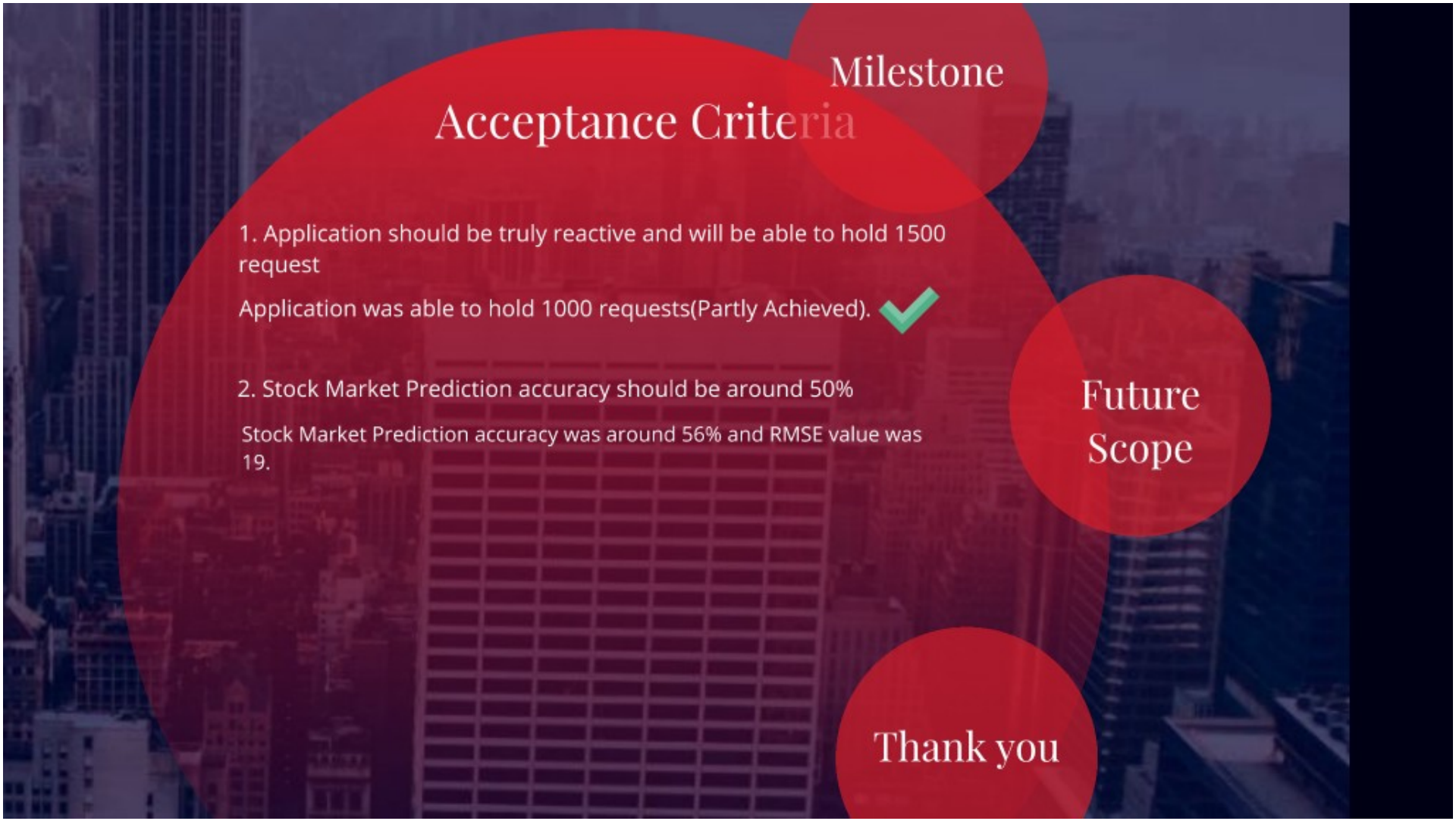

Milestone
Acceptance Criteria
Future
Scope
Thank you

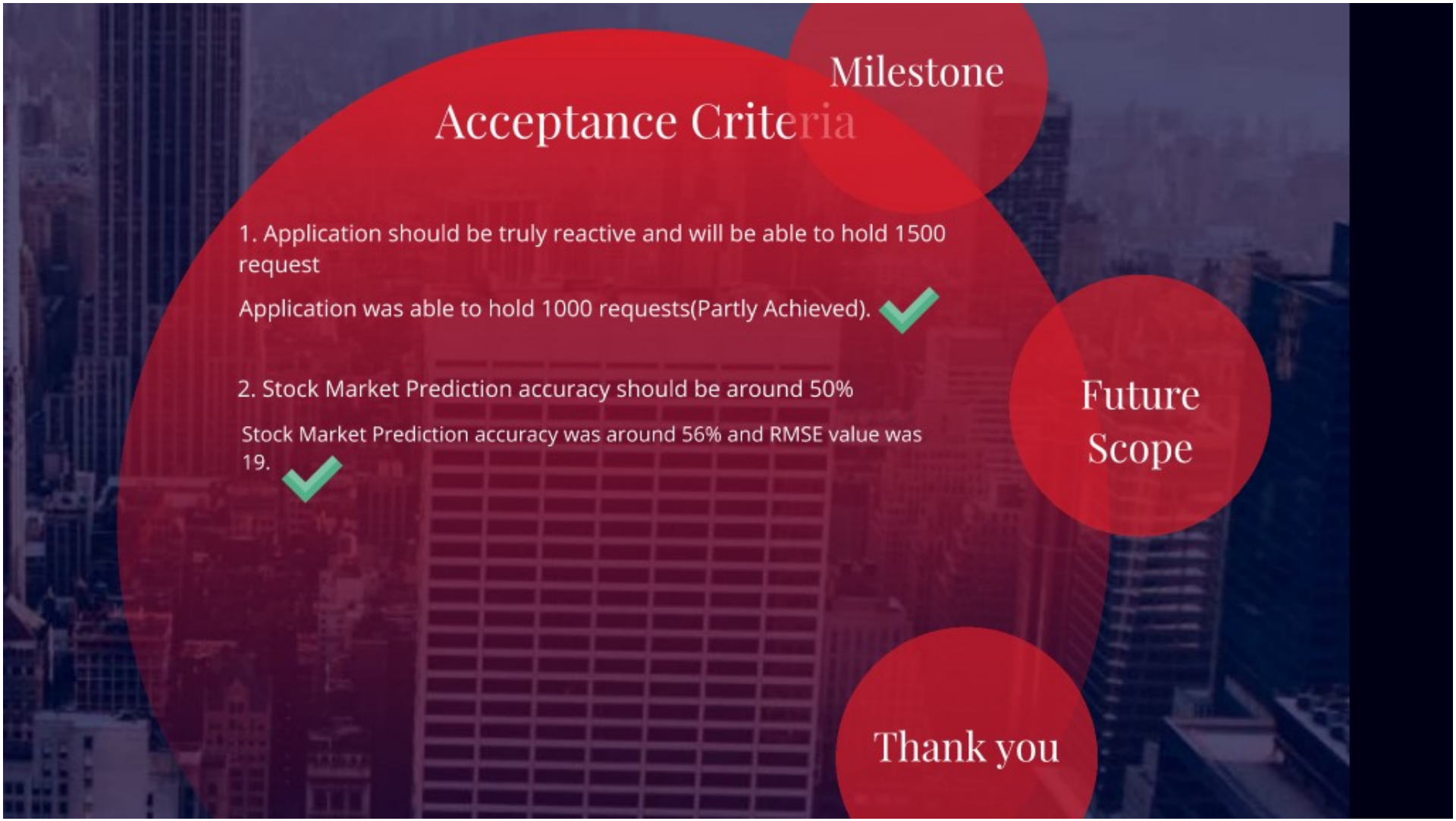

Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Future Scope
Thank you


Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Application was able to hold 1000 requests(Partly Achieved).
Future Scope
Thank you


Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Application was able to hold 1000 requests(Partly Achieved).
Future Scope
Thank you


Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Application was able to hold 1000 requests(Partly Achieved).
2. Stock Market Prediction accuracy should be around 50%
Future Scope
Thank you


Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Application was able to hold 1000 requests(Partly Achieved).
2. Stock Market Prediction accuracy should be around 50%
Stock Market Prediction accuracy was around 56% and RMSE value was 19.
Future Scope
Thank you

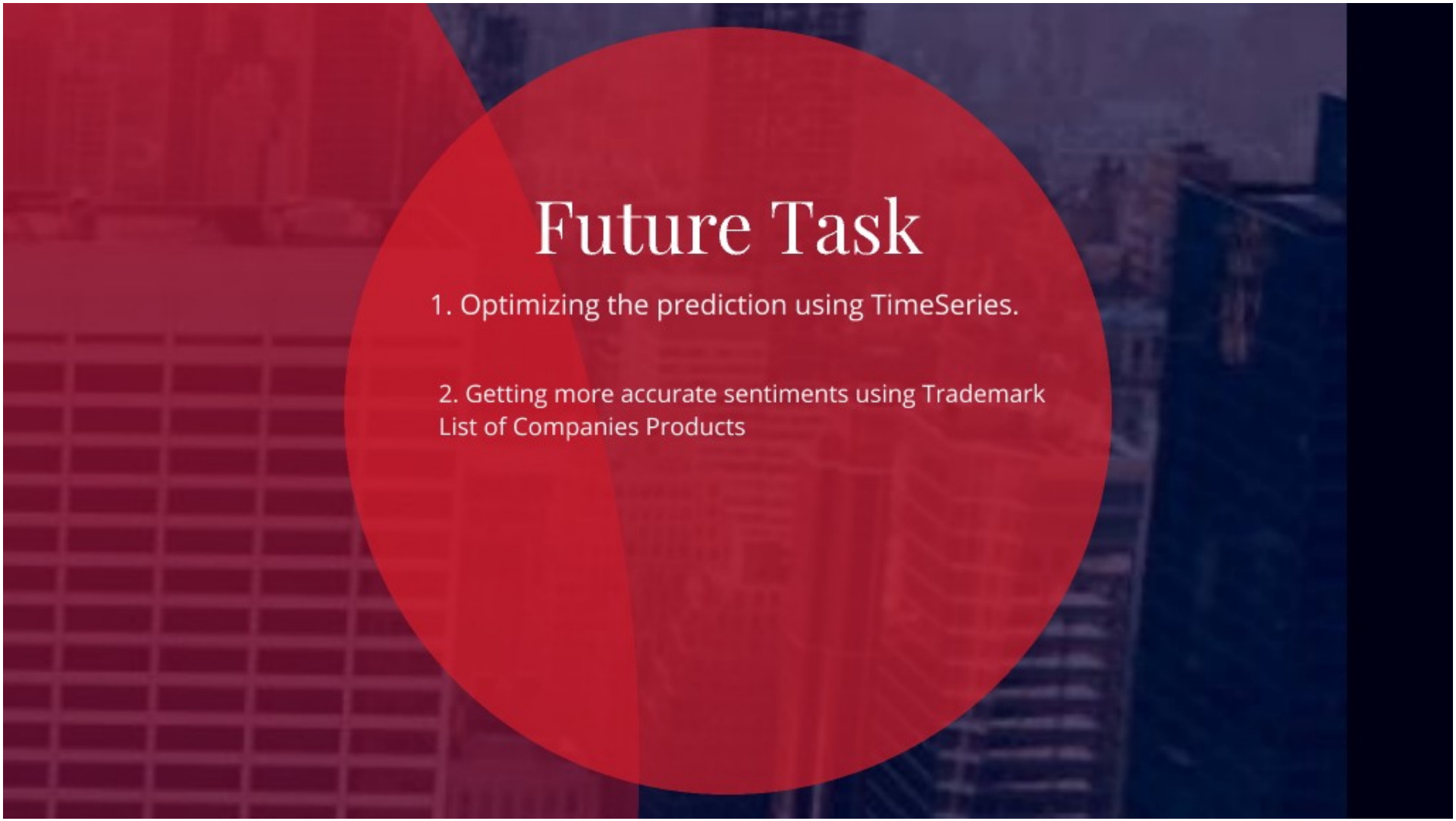

Milestone
Acceptance Criteria
1. Application should be truly reactive and will be able to hold 1500 request
Application was able to hold 1000 requests(Partly Achieved).
2. Stock Market Prediction accuracy should be around 50%
Stock Market Prediction accuracy was around 56% and RMSE value was 19.
Future Scope
Thank you

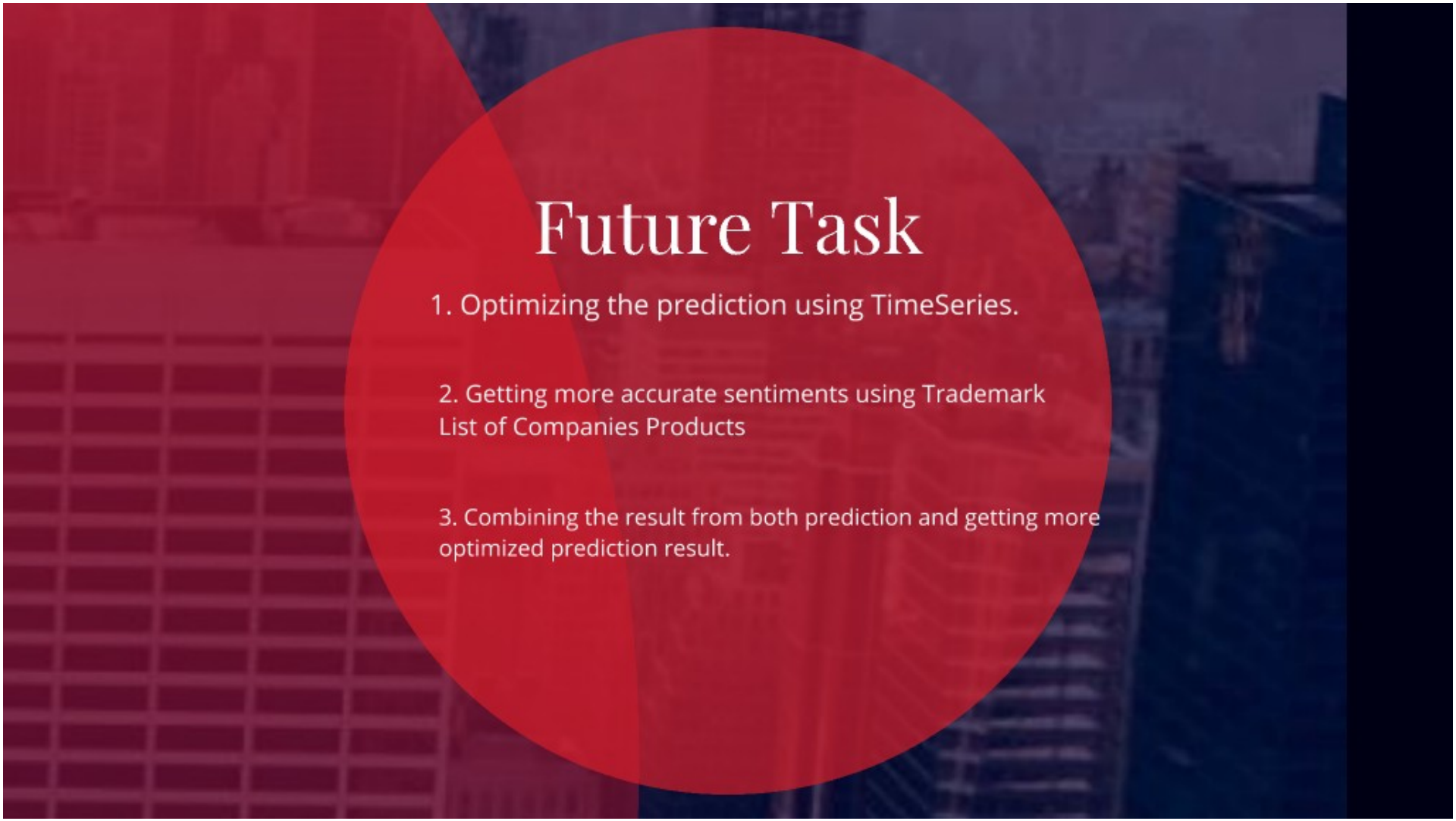
Milestone

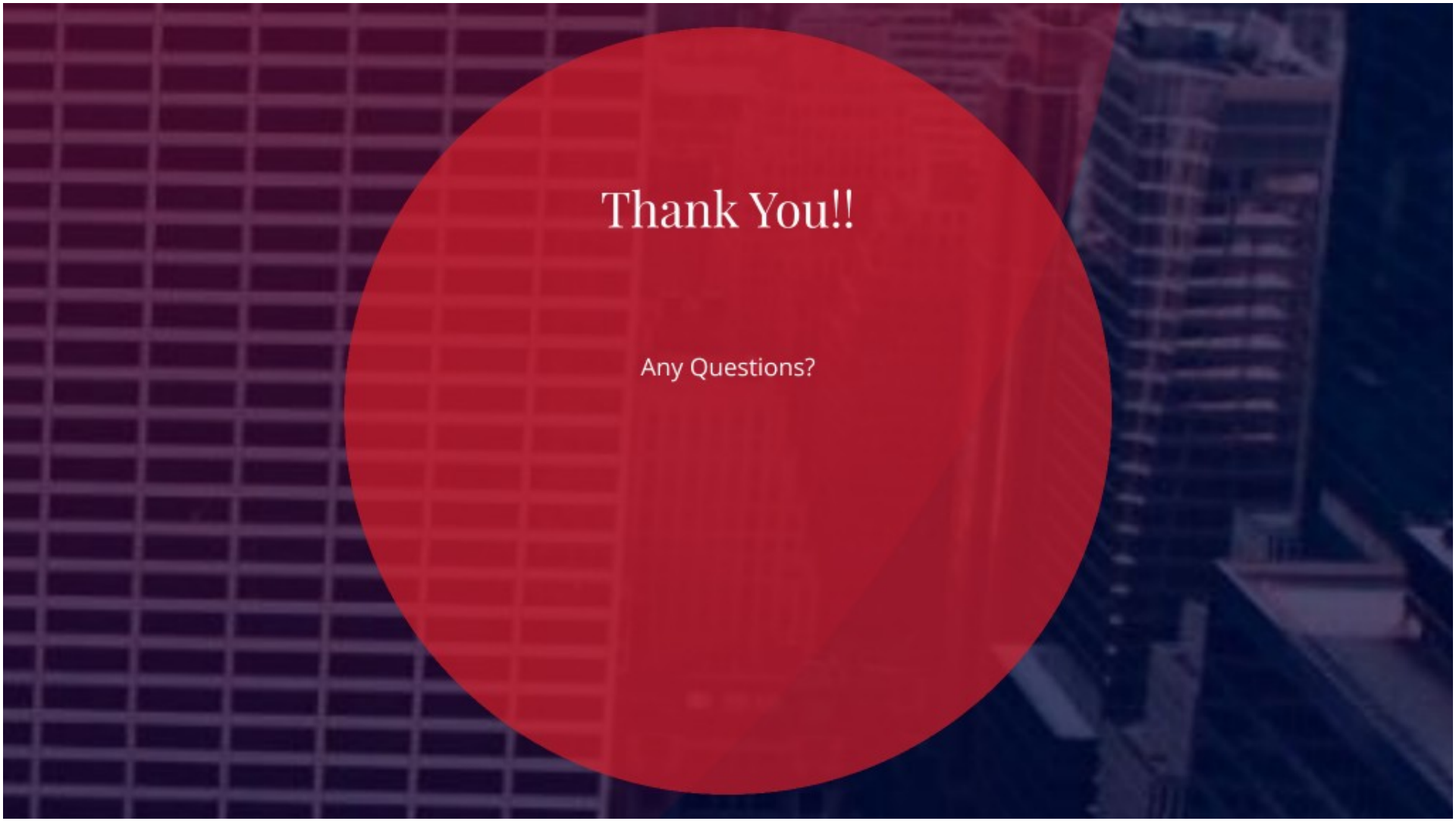
Task	Time
Come up with the Prototype for our reactive System and data cleaning	4/01
Exploring and implementing Actor model and spark and unit tests	4/14
Integrating Spark and Spring	4/20
Functional and load testing /self acceptance	4/22


Future Task


Future Task
1. Optimizing the prediction using TimeSeries.


Future Task
1. Optimizing the prediction using TimeSeries.
2. Getting more accurate sentiments using Trademark List of Companies Products


Future Task
1. Optimizing the prediction using TimeSeries.
2. Getting more accurate sentiments using Trademark List of Companies Products
3. Combining the result from both prediction and getting more optimized prediction result.


Thank You!!
Any Questions?

Stock Market Prediction

Purva Bundela - bundela.p@husky.neu.edu

Dishank Shah - shah.di@husky.neu.edu

Repository Link: https://github.com/PurvaBundela/Stock-Market-Prediction-Using-Scala-and-Spark