**Predictive Analytics of Crypto Market**

**Project Proposal**

Date: 10/27/2020

Connect with Power BI

Connect with Excel

Visualize in Power BI with live data

Clean the data in Excel and Data will refresh every 1 minute with new live data

Get live data from marketwatch.com

**Version – 2**

Date: 10/28/2020

Get live data from marketwatch.com

Get live data from coincapmarket.com

Visualize in Web page

Connect with web page

Clean the data in Excel and Data will refresh every 1 minute with new live data

Connect with Excel

Predictive Analytics using ML

Visualize in Tableau or Power BI

1. **Data source(s)**
   1. We are considering data from two web sources which will be used to perform visualization and predictive analytics.
2. www.coincapmarket.com
3. www.marketwatch.com
4. **Data capturing**
   1. Data is extracted into an Excel sheet using the “extract web data” option available under “Data” tab on Excel.
   2. Paste the target URL and click on import.
   3. Select cell and data will imported into the sheet
   4. By using the option, the entire data available on the website is captured and stored in excel sheet in an unstructured format.
5. **Curation process**
   1. Data is extracted in an unstructured format.
   2. To clean the data, take a crypto/stock price into a new sheet using a cell formula. The same process can be repeated to take 20 different stocks/cryptocurrencies into a new sheet.
   3. To update the live website data in our excel sheet, we need to automate to refresh for every 1 minute by changing the preferences.
   4. New data will be appended to the existing data after every data refresh.
   5. To save the old prices and new prices of each stock every time the excel sheet is refreshed, write a VB script procedure and save it. VB script in Excel can be accessed by clicking alt+f11
   6. By following all the above steps, we can clean the data and get all stock prices in a structured format.
6. **Data processing**
7. **Data analyzing**
8. **Predictions** 
   1. For predicting the data, we are using linear regression model in this project.
   2. The linear regression model perfectly fits for this data because for predicting the data we need to provide train data and test data and based on the provided data it generates the predicted data.
   3. In this project, new data (stock price) will be updated for every minute. For the past stored data is given as train set.
   4. We will calculate the metrics to predict the accuracy of the predicted data and the predicted data is visualized using the matplotlib.
9. **Visualizing**
   1. **Visualization using Tableau :**
10. For visualizing the crypto currency price we used Tableau in this project.
11. Open the Tableau and import the data through Microsoft Excel file.
12. Drag your corresponding sheet in to the dashboard then you will see the data in the data source.
13. In our project we visualized the crypto currency rates for every minute using line graph by keeping the time in the column and crypto currencies in the row.
14. Change the automatic to the line and then you can see the line graph.
15. In this project we used parameters so that we can see the line graph for every crypto currencies.
    1. **Visualization through web using JavaScript:**
16. We are also visualizing the cleaned data which is coming from different websites through web by using JavaScript. We were using line chart to visualize the data.
17. To accomplish this task, we have used PapaParse and C3 libraries. Both are explained briefly below.

* **PapaParse:** PapaParse library is a fast and powerful CSV parser for the browser that supports web workers and streaming large files. It is easy to use and parse CSV files directly to local or over the network. We have used this library to extract the source data which is in excel and stores in the form of arrays.
* **C3:** C3 is a JavaScript library which builds on top of D3. It’s a great and simple tool to make a chart in just a few minutes. C3 makes it easy to generate D3-based charts by wrapping the code required to construct the entire chart. We don't need to write D3 code any more.

1. With the help of these two libraries, we are visualizing some of the crypto currencies and stock prices through line chart.

**Tasks:**

|  |  |
| --- | --- |
| **Task Name** | **Assigned Team member** |
| Data capturing and cleaning in Excel | Deepak Malempati, Pavan Sai kumar |
| Predictive Analytics | Pranay Allikanti |
| Visualizing in webpage | Saiprasad Bobbilla |
| Visualize in Tableau or PowerBI | Chaitanya Popuri |

References:

1. Data Cleaning - <https://medium.com/@victorleungtw/getting-real-time-data-from-web-to-excel-467913abe61a>
2. Data Visualization tutorials - <https://www.youtube.com/watch?v=9TXdFxmYlAc&feature=emb_logo&ab_channel=StudentLife>
3. JavaScript - <https://www.sitepoint.com/interactive-javascript-charts-using-data-from-google-sheets/>

Homework

1. Create a github repo for the team
2. Create list of references
3. Divide your tasks based on architecture diagram
4. Document the every single step.

**Version – 3**

Date: 11/15/2020

**Multi-dimensional analysis of data streaming**

Get live data from coincapmarket.com

Get live data from coinmarketcap.com

Get live data from cryptoprices.com

Get live data from marketwatch.com

Visualize in Web page

Connect with web page

Clean the data in Excel and Data will refresh every 1 minute with new live data

Connect with Excel

Store in MS Access Database

Predictive Analytics using ML

Visualize in Tableau or Power BI

**Version 4**

Date: 11/22/2020

Get live data from coincapmarket.com

Get live data from coinmarketcap.com

Get live data from cryptoprices.com

Get live data from marketwatch.com

Connect with Excel

Load the data into excel

Data refresh every 1 minute

Write excel formula to extract crypto/stock prices into a new sheet

Write VB script to append new values for every 1 minute in exce1

Store in MS Access Database

Perform aggregation and computations to find max price and time of the instance

Using c3.js

Predictive Analytics using ML

Visualize in Tableau or Power BI

Visualize in Web page

**Flow Diagram for data extraction and transformation:**

Extract required attributes



Time stamp

HH:MM:SS

Append &

Compute

Merge

Aggregation (max)



Aggregate results

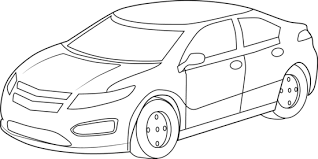
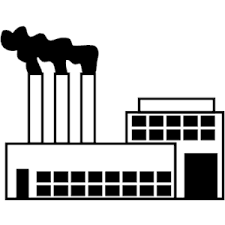
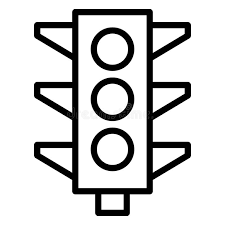
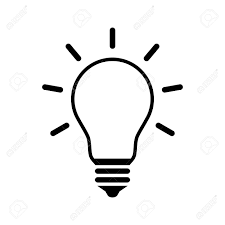
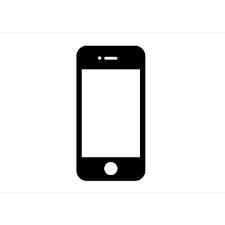
[coincapmarket](coincapmarket.com)

[marketwatch](marketwatch.com)

[cryptoprices](cryptoprices.com)

[coinmarketcap](coinmarketcap.com)

**Architecture Diagram:**



NewData

Speed Layer

Macro-enabled Excel(append)

Real time views

Batch layer

Complete Data

Service layer

Visualization

Tableau

Webpage

Predictive Analysis

**Steps to perform data extraction and transformation**

1. Identify and choose multiple data sources of your choice. In our project we are working with 4 different sources of with crypto currencies and stocks price live data.
2. Open excel and create 4 new sheets
3. In sheet 1, click on Data tab -> select “From web” option.
4. A pop up will be displayed. Enter the website url from where you are planning to extract data
5. Click import. This will import all the raw data of the webpage into excel sheet.
6. Repeat steps 3, 4, 5 for all the other sheets taking different website urls.
7. By the end of performing above steps, you should have all the raw data into all the 4 sheets.
8. Create another sheet which will be sheet5. This sheet5 will be used as transformed data sheet and we will be performing all the transformation techniques in this sheet.
9. Choose the crypto/stock price that you want to work on in all the sheets.
10. Now, using below excel formula get the values into sheet5 in A column

**Example:** =Sheet1!D155

=Sheet2!B80

1. Now B Column in sheet 5 will be time stamp. To get the timestamp and append it for every one minute, you must write a VBA script in excel
2. In order to write a VBA script, click alt+F11 and VBA script pop up will be opened.
3. Select sheet 5 in left top corner.
4. Copy & paste the code that is provided in code section of this document.
5. Click on file->save and it should ask to change the type of excel sheet into macro enabled workbook. Click on save as -> in the drop down choose macro enabled workbook and click save.
6. VBA script enabled workbooks must be saved as macros enabled
7. Finally, click on Data tab -> click refresh all -> connection properties -> change auto refresh to 1 minute and save
8. Now your excel should be refreshed every 1 minute and new data will be appended into sheet5 every 1 minute.
9. To perform aggregations write excel formula by selecting a entire column of crypto currency price and use desired aggregate function.
   1. **Example: =MAX(C2:C250)**