**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 data 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 Bitcoin rates for every minute through line graph by keeping the time in the column and Bitcoin in the row.
14. Change the automatic to thee line and the n you can see the line graph.
    1. **Visualization through Web :**
15. We have also visualizing the data through web by using JavaScript.
16. We are visualizing line chart for the excel data in the web.

**Tasks:**

|  |  |
| --- | --- |
| **Task Name** | **Assigned Team member** |
| Data capturing and cleaning in Excel | Deepak Malempati, Pavan Sai kumar |
| Predictive Analytics | Pranay Allikanti |
| Visualizing in webpage | Sai Prasad Bobbila |
| Visualize in PowerBI and Tableau | 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.