SWE573

Tugcan Olgun
2016719210

Cryptocial

May, 2018

Purpose of the project

This project has been carried for the class SWE573 Software Development Practise in Bogazici University under Prof. Dr. Suzan USKUDARLI’s supervision.

Github repository: https://github.com/tugcanolgun/SWE573/

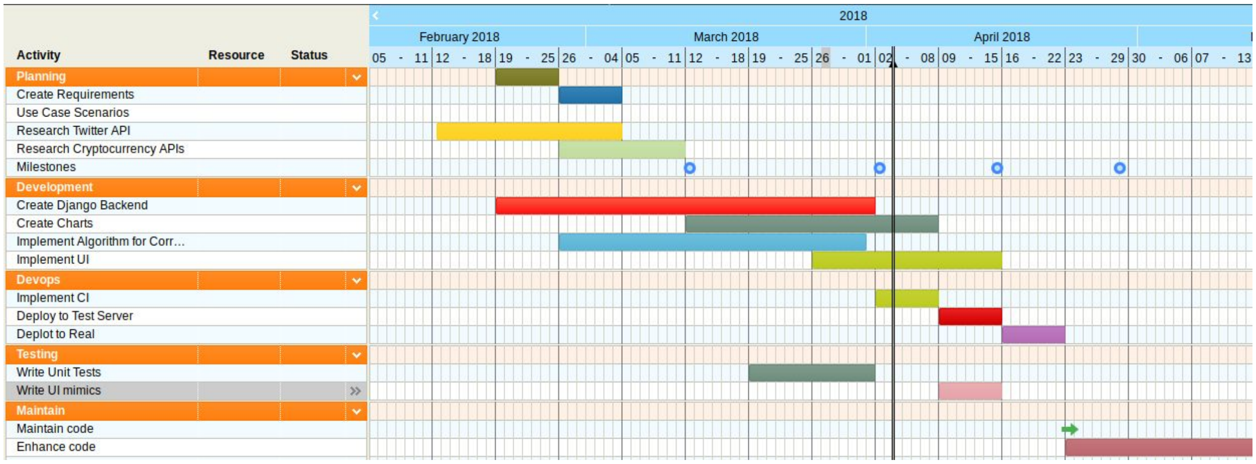
Live deployment: https://cryptocial.ml/

Introduction

Cryptocial is an experimental cryptocurrency-tweet analyzer. It correlates tweets based on their sentiment analysis and correlates with high-low real currency values. It fetches hourly tweets on certain currencies and do sentiment analysis. Then it fetches the high and low prices for that particular currency and do linear regression analysis. It both shows the result of regression and writes a simple readable sentence to read.

Planning

The project has started with the planning phase. The details of what must be done in which milestone tried to be defined very clearly to keep up with the phase of the project. Planning are divided into sections to distinguish the activities. The sections are Planning, Development, Devops, Testing and Maintain. Some activities overlap others by their nature.



Requirements

Requirements are the bones of the project. I defined in detail as much as possible to make good use of it while doing the project.

Functional Requirements

The following are the functional requirements. All of the requirements are fulfilled in this project.

1. Twitter API will be used to fetch twitter messages.

    1.1. Users will be able to choose a cryptocurrency on the list, upon choosing a portion of the messages related to the cryptocurrency will be shown.

2. CryptoCompare API will be used to fetch cryptocurrency prices.

    2.1. Users will be able to choose a cryptocurrency on the list, upon choosing the prices will be fetched.

2.2. Users will only be able to see the frequency correlated with the existing tweets and analysis in the database.

3. A script must be run indefinitely to gather new tweets.

3.1. Popular cryptocurrencies must be fetched from coinmarketcap API and written to the database.

3.2. New tweets must be fetched from newly added cryptocurrencies and written to the database.

3.3. When each cycle is over, sedimental analysis of newly added tweets must be executed.

4. Cryptocurrency names will be added to the database manually depending on trust levels of sediment analysis.

5. The system must analyse the historical prices with the twitter messages along with the new ones.

6.The frontend must display the results in a user friendly manner.

7. The system must store the used data in a database for not to exceed API fetching limits.

Nonfunctional Requirements

The following are the nonfunctional requirements. All of the libraries that are used in this project are open libraries.

1. Project must use only free libraries.

2. Project must include licenses of the libraries as well as its own licenses on production.

3. UI must be scalable.

Design

After requirements process the design phase has started. Writing use case scenarios, creation of mockups, sequence diagrams and class diagrams are done in this phase.

Use case scenarios

First scenario

Our first persona is woman named **Arzu Bal***. She works as a data scientist in a big data related software company. She has recently been asked to do correlated data analysis on cryptocurrency. She wants to check what sort of solutions available on the web and googles "cryptocurrency twitter". She reaches the website.

1. She picks a currency from the home page.
2. Sees the correlation data between cryptocurrency and twitter sentimental analysis.
3. She gets excited as it is a very good use of data analysis.
4. She picks one currency after another to see what other possible outcomes there are

Second scenario

Our second persona is man named **Kriko Ahmet**. He works in a car repair shop and recently is interested in cryptocurrency which he has heard from his friend. As he has no idea about which cryptocurrency to choose from, he wants to play safe and go with the popular cryptocurrency.

1. Sees there are cryptocurrency banners to choose from
2. He picks on one of them

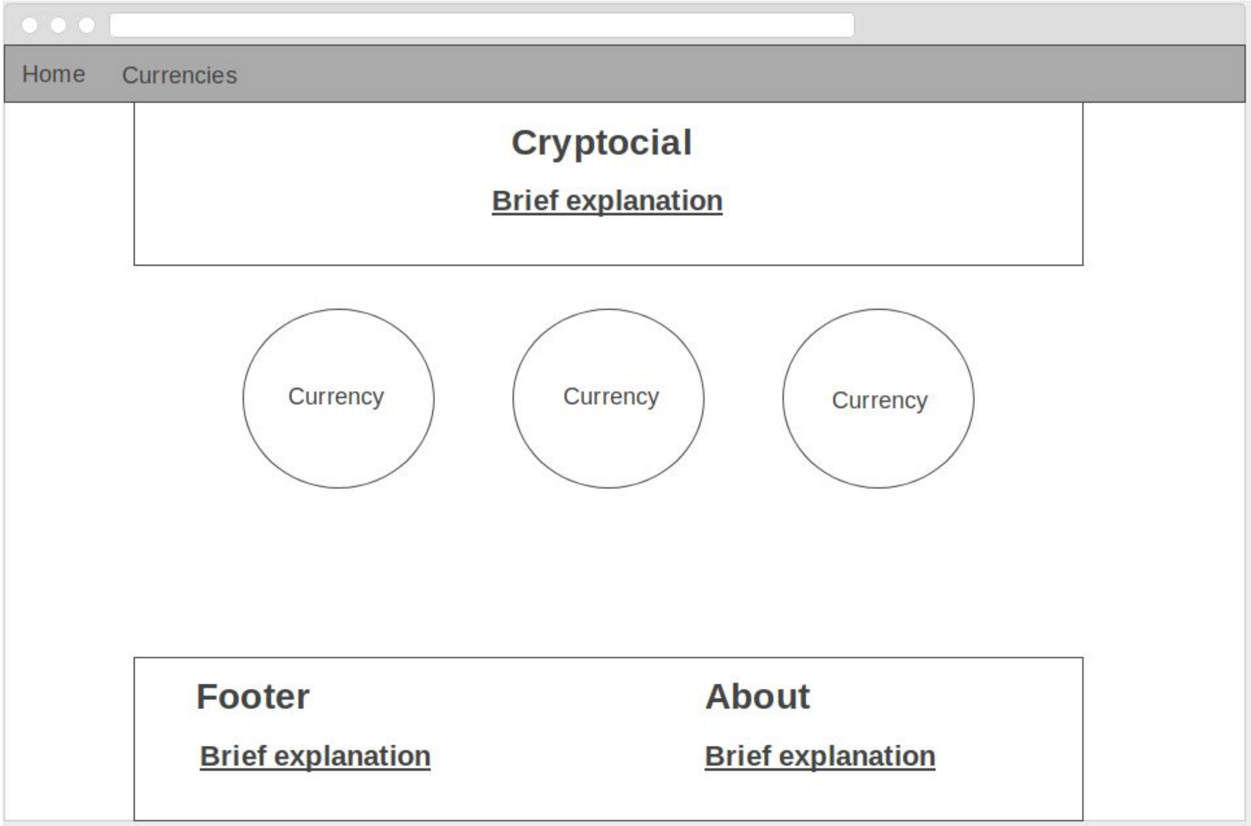
3. He is presented with a graph of the selected cryptocurrency and the correlation with the number of recent twitter chatter
4. He see there is a twitter lead correlation and buys that currency to get rich!
5. He wants to check the other currencies to check about their recent trends
6. He clicks the home page and picks on another currency
7. He repeats these steps until he is satisfied with the information
8. He closes the website

Mockups

The site was designed to have two functional pages. Home and detail view page. Both pages are designed to be very clear for users and during designing any extra buttons, links or texts that could be perceived as clutter tried to be avoided.

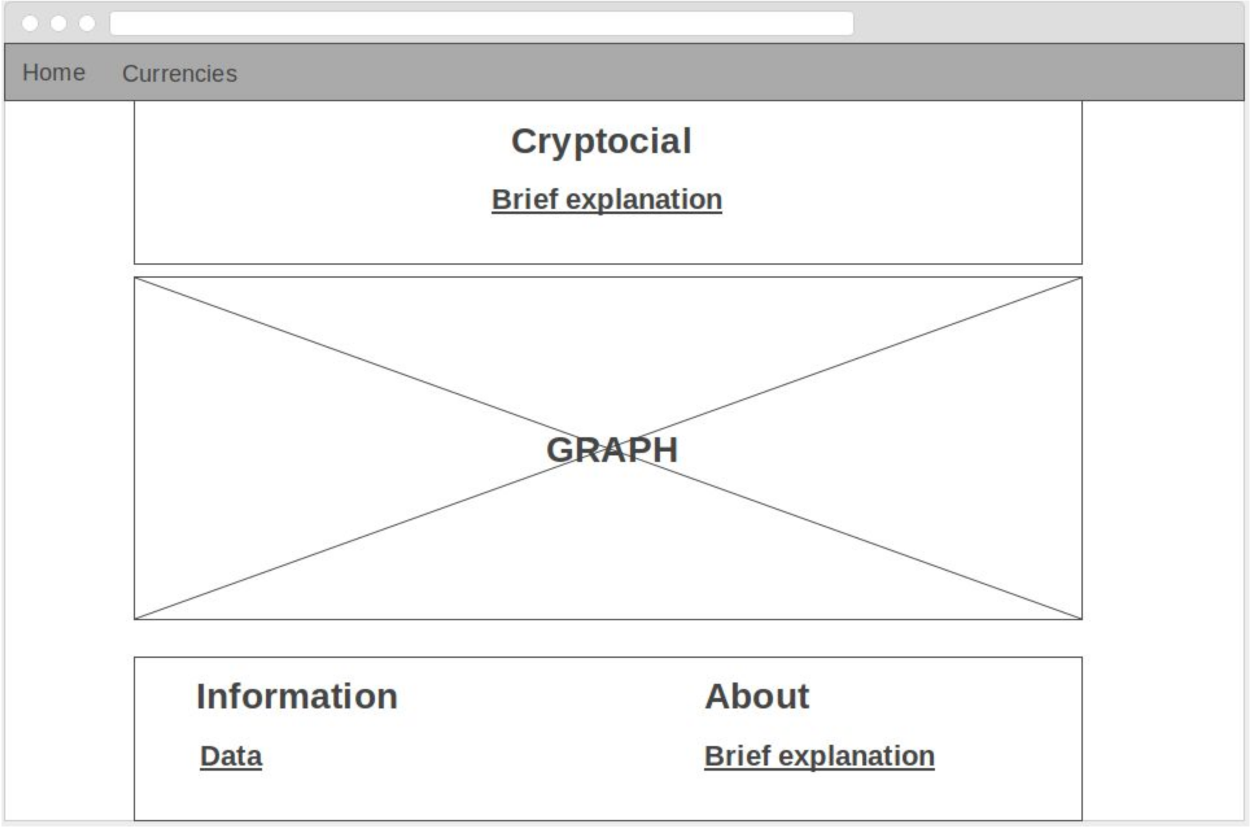
Home screen

Homescreen has a menu bar at top and a very brief explanation of what the site is about. In the middle it has big logos for different currencies that are already been analyzed. And at the bottom, there is a footer which includes the explanation of what the actual purpose of the site as well as attribution links to the elements that are used in the site.



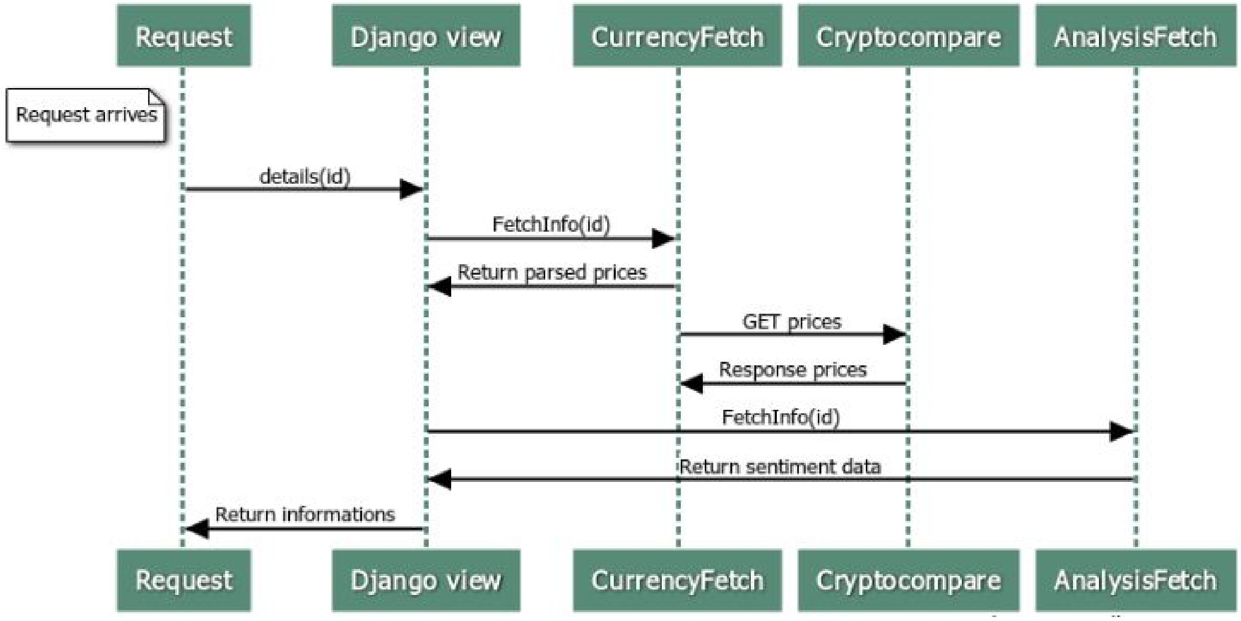
Detail screen

Detail screen is designed to have more or less same template to have consistency. It has a javascript graph which shows different attributes and a details section just underneath for those who would want more details in terms of the data that are being shown in the graph.

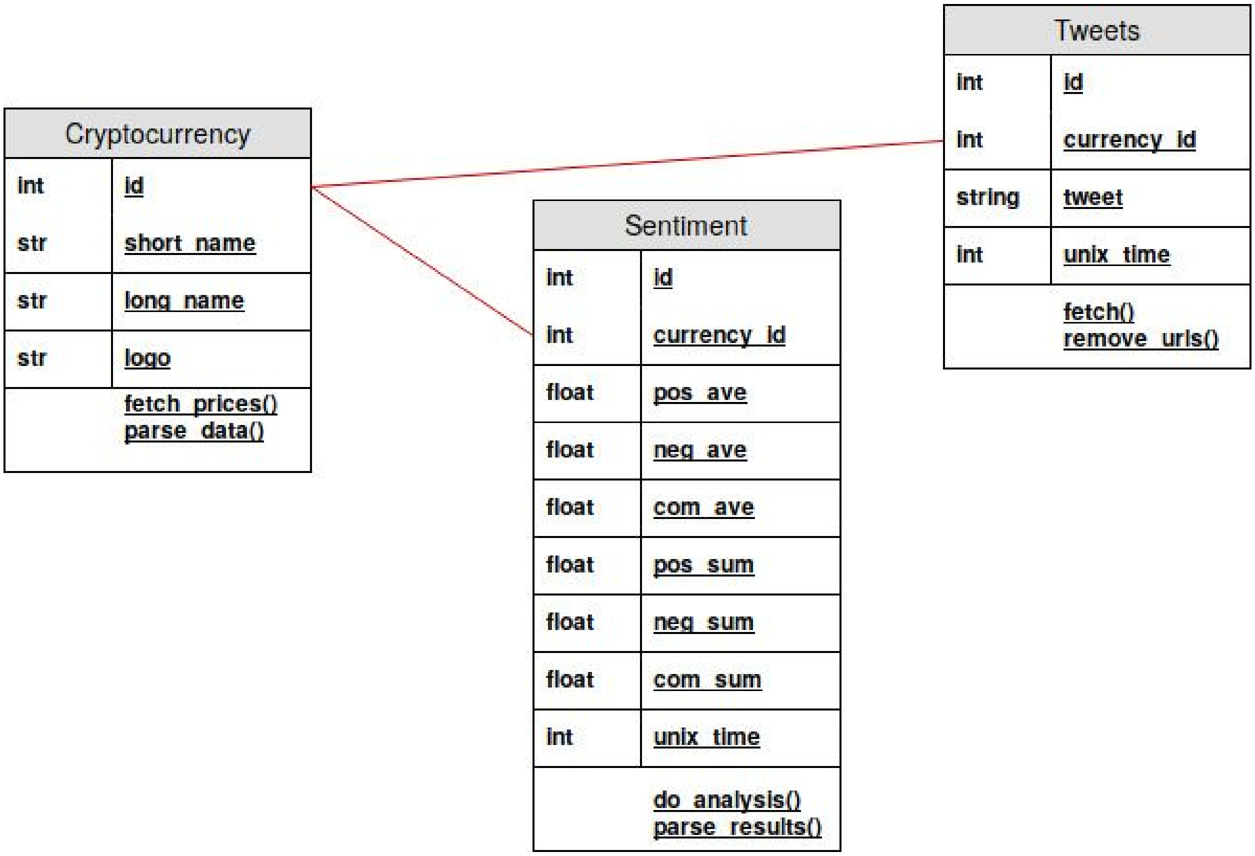


Sequence diagram

Details page



Class diagram



Development

Development phase started with (using python3.6) experimenting with Twitter API. API has its own limitations. To deal with these limitations, tweets have to be fetched on calculated intervals. After fetching tweets based on hashtags excluding retweets finished, sentiment analytic tool has started to be experimented on. nltk with vader package is a very good sentiment analytics tool. To combine both tweets and sentiment analysis, an sql database has set up. As doing these calculations take time and as the purpose of the project is to get the correlation data between tweets and the currency prices not only the present time but also previous, an ongoing script has started to be written to get the most popular tweets from coinmarketcap.com's API. After getting the currency names, tweets are fetched, cleaned up with their links, saved to the database. After each saving, the sentiment analytic script is being run to get the analysis done and save to the database with their appropriate information such as the name of the currency and their timestamp along with the analytic information.

While this ongoing script is running, the django website gets the information from the database that the script is running on. Upon user selection from the home page, only instant API fetch is from cryptocompare's.

Not all currencies are displayed on the homescreen. Only the ones that are allowed on the Django admin panel are shown.

The following is a list of libraries that are used with this project that are not in the standard library of Python.

- django
- numpy
- requests
- tweepy
- nltk
- pytest

Testing

For testing purposes. The pytest library is used. With addition of integration with Travis platform which is a distributed continuous integration service used to build and test software projects hosted at github, testing is always carried with each git push.

Deployment

The project is deployed in a Ubuntu 16.04 server hosted at digitalocean. The django template which runs in enclosed python environment is served with gunicorn with concurrent workers behind nginx. To enhance the security, certbot, formerly known as letsencrypt is used to provide an SSL certificate. The domain in which is used in this project retrieved from freenom.com. A live cryptocial project can currently be found on https://cryptocial.ml