Sunday, September 15, 2019

User Story + Agile + Modular Mini Project: Build a library (preferable in python) that analyzes twitter feeds: sentiment of text twitter feed.

Sprint1

Define the User stories

(1) Initial Topic: Interview Dressing

- Starting Point: maybe we can do something different
- Shortcoming: it is hard to use the quantitative methods.
- Example: we cannot find a blogger who write about the interview dressing only, also the brand of dress is hard to find in English, even after using the advanced research

(2) Adjusted Topic: Vogue Magazine

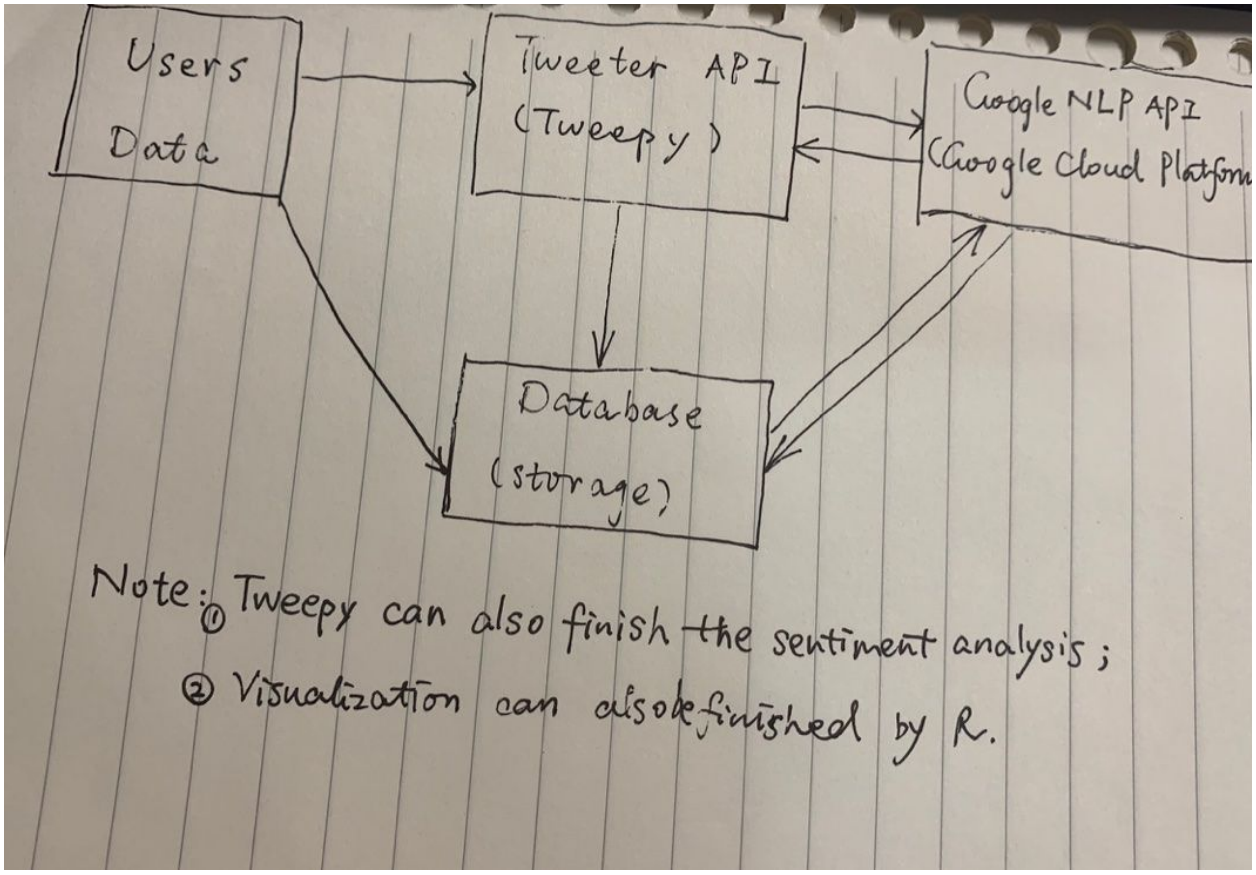
- Starting Point: we searched for the best brand for twitter in 2019 through the Internet

  https://heartifb.com/the-fashion-bloggers-guide-of-who-to-follow-on-twitter/

- Reason: a traditional magazine has responsibility to promote their works to the reader via all of the popular social media
- User stories in details:

Vogue Magazine try to lead the world culture to the global readers via the paper/digital magazine. Unlike Elle, Vogue is always in the No.1 for all charts. Only one magazine, covers 23 countries.

Architecture



Define the MVP: Study the Twitter and Google frameworks

(1) For Twitter: understanding the different Twitter APIs. For example, there are few Twitter python libraries.

Twitter API should not be selected randomly, we need a library which is easy to start. Also, it should fit our case. In this project, we start from texts, not images. I don't think changing from images into texts is a good way, because twython can fulfill the requirement.

Library	**Comments**
Python-twitter	a pure interface, not well organized
Tweepy	**official documentation provided and examples**
TweetPony	limited examples
Python Twitter Tools	good, with many contributors
twitter-goject	only 1 contributor
TwitterSearch	have a search operation
twython	can handle images
TwitterAPI	need a setup step
Birdy	lack of example

Before using the tweepy, we need to get the credential first.

(2) Google Natural Language APIs

There are two types of API: (1) AutoML Natural Language; (2) Natural Language API. We will use the latter one for sentiment analysis.

Before using the tGoogle Natural Language API, we need to get the free trial first.

Write examples to use at least to get twitter feed

and get results from one of the Google modules.

Natural Language API demo

Try the API

Google, headquartered in Mountain View (1600 Amphitheatre Pkwy, Mountain View, CA 940430), unveiled the new Android phone for $799 at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.

RESET

See supported languages

Entities | Sentiment | Syntax | Categories

⟨Google⟩1 , headquartered in ⟨Mountain View⟩2 ⟨Mountain View (1600 Amphitheatre Pkwy⟩12 ⟨1600⟩16 ⟨Amphitheatre Pkwy⟩7 , ⟨Mountain View⟩2 , ⟨CA 940430⟩8 ⟨940430⟩15), unveiled the new ⟨Android⟩3 ⟨phone⟩5 for ⟨$799⟩13 ⟨799⟩14 at the ⟨Consumer Electronic Show⟩11 . ⟨Sundar Pichai⟩4 said in his ⟨keynote⟩9 that ⟨users⟩6 love their new ⟨Android⟩3 ⟨phones⟩10 .

1. **Google** — ORGANIZATION
Wikipedia Article
Salience: 0.19

2. **Mountain View** — LOCATION
Wikipedia Article
Salience: 0.18

Sprint 2

Module 1: Get i text feeds from twitter for a twitter handle

I got a hard time learn the twitter API libraries. If I can start from the class example, it will be easier for me. The good thing is that I know where the libraries are and what I can do in the future. I even learned about the libraries in Java, C#.

Module 2: Sentiment and keyword analysis of tweets

I tried to set up virtual machine from the Google Cloud platform, but I cannot get my SSH keys on mac. At the same time, we the platform seems to update during the process.

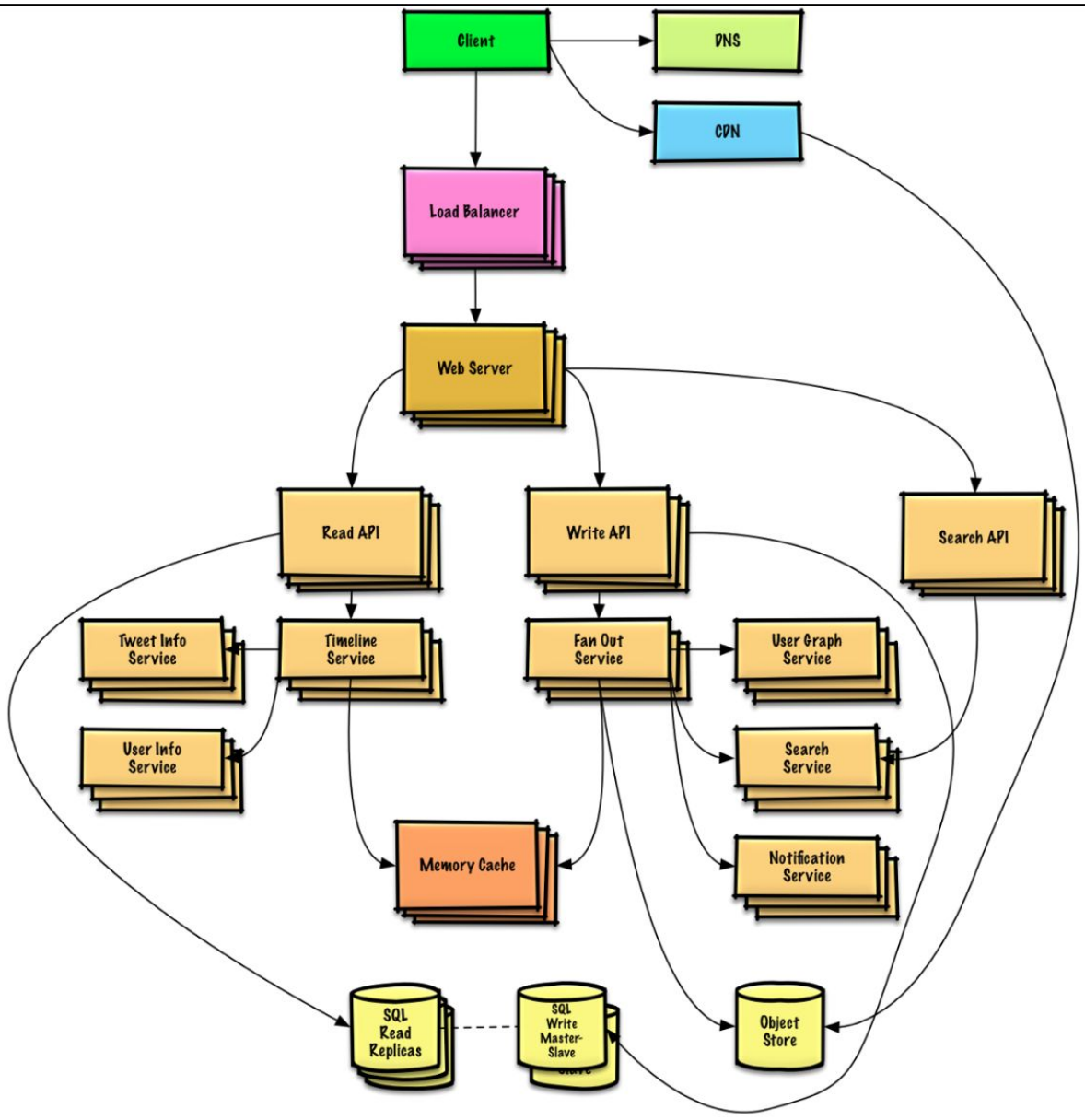
Moreover, it took me sometime to understand the Google credentials. Is there anything differences between the twitter API and the Google API?

For more information, see the github webpage.

Sprint 3

Tested integrated system

The most important thing is to have a clear understanding of the system design. It might not go into CS/BS, but we still need to have a clear mind of the integrated system.



For more information, see the github webpage(including the full report):