

# Hashtag Battle Project

**Shuya Chu**

csy.ellie@gmail.com

## 1 Design

For this wepapp, my work flow as following:

A graph will be here

## 2 Analytical Potential

TO HONEY: Being asked to answer these questions.

1. If we were to build analytics for this feature, what analytics would potentially be useful for customers that we could build with the data we have?
2. How do you want to integrate this with our existing product from a product and business standpoint (not a technical standpoint)?
3. What type of analytics can you provide ?
4. Briefly explain the architecture.

If I have more time on this project, I will focus more on designing a proper feature space to cluster the user that posted twitter for a certain brand. To be more specific, while counting, we can also have access to the user's info, i.e. gender, location, favorite, # of followers, it would provide significant insights if we can cluster the users using bootstrap or simple k-means algorithm. These cluster results would tell us the main feature of people who concerns the brand and thus helps our company partner to have a better idea about marketing their products.

Also, we can extend the platform to get more useful information. Extending this feature to Facebook, Linkedin and so on will help the company understand which of these social media is the best plat-

form to advertise. For example when Facebook concerns more about everyday life, it would be a nice place advertising a new recruiting software products.

Moreover, this feature can also help Tint in some way. First, in my mind, tint collect information based on user's requirement. What if the input is not the most popular to describe their interests? Running a tagbattle between "MH17" and "MH" may provide a better describe suggestion when a user want to know news about the recent plane crash tragedy. Integrating this to existing product, user experience will be improved.

## 3 Future Work

:

There is a problem I didn't solve in this implementation. If frequency of one of the hashtag being mentioned is too high, some times the sqllite database will not be able to fulfill the required concurrency level. I have two ways to solve this problem but due to the time I am not able to implement that.

One way is to stop using SQLite, which is a lightweight database, and thus cant support a high level of concurrency.

Another way is when a update to database is needed, fire an asynchronous task using celery worker and parallel these tasks.

## 4 Main Working

I spend the almost whole seven days to work on this project. During this time, I used 1 day to learn django, understand its framework and design my project, 4 days to study Celery: Distributed Task

Queue. After understanding how celery works and how it interact with django and python, I used the rest time to finish the back end of this app. Also, since I am new to css/js, I use 1 day to learn bootstrap and embedded it to the project.