To Identify the Person behind a Tweet

Vishal Marhatta
Manoj Baviskar
Akmal Ziyad

Functional Specification

- Upload the dataset to predict(tweets).
- User will enter the anonymous tweet to predict the actual author behind it.
- User will get to know the percentage based upon which user will decide that the tweet belongs to which author.

Technical Specification

- **Processor**: Quad Core
- **RAM**:2GB
- **Tweet format**:

  [Twitter username] [timestamp] [Tweet id] {
  [Tweet message (multiline allowed)]
  #POS [POS Tag data] #POS
  }
  Example:-
  author_nickname 2016-03-28 20:52:53 091294878667731987 {
  @Friend Im gonna bring beans
  #POS I L V V#POS
  }
- **Programing Languages** : Python , javaScript
- **Frameworks** :ExpressJS,Mongoose
- **Libraries** : ReactJS
- **Runtime Environment** : nodeJS
- **Database** : MongoDB

Design Specification

- **Front End**
  - Uploading dataset using button click
  - Display Results

- **Backend**
  - Dataset Preprocessing
  - Filter Retweets
  - Remove irrelevant data
  - Classification

External Interface Specification

Website

- Upload button to browse and upload the data
- Training time will be displayed
- Notification when Training completes
- Text box to enter tweet of person to be checked
- Display the results in intuitive manner

Persona

- Mukesh is a forensic investigator and is 30 years old.Mukesh is given a job to find out the author behind an anonymous tweet .Mukesh is confused between some authors.So he decides to use XYZ system in order to find out the suspect .He inputs the tweet to the system and the system replied with the percentage match with suspected authors.Mukesh found that the tweet matches 70% with the author A’s tweets.He found the suspect and solved the case.

Sequence Diagram