

AP-PROJECT

FACEBOOK LIKE PREDICTOR

Mansi Goel, 2014062, mansi14062@iiitd.ac.in

Vrinda Malhotra, 2014122, vrinda14122@iiitd.ac.in

Our task was to predict the number of likes on a facebook post of an user given some posts and their likes as training data set.

Using a series of algorithms, we can estimate the change in pattern of likes for each user depending upon the content of their posts.

- 1) **Using most common and relevant words** : By removing all the stop words from the posts of a user and making a hash table of most relevant words. We can link these words according the likes or preferences of each of their friends. So, we can predict the number of likes for a post with similar content as that of a post in our training data.
- 2) **Using Presence of Images or links in post** : Analysing this for every user, we can predict deviation from the usual number of likes in a post if it contains an image, a video, a link or coordinates of a map. So, we can predict the change in the pattern of likes if the post contains images or links or something other than normal text.
- 3) **Using trending words** : The words which are in the latest trend can be fetched easily from the browser. According to the observed pattern, number of likes are usually greater for the post which has more in trend words.
- 4) **Using close friends**: The friends who like more than 60% of our posts in the training set are more likely to also like the next post. Also, this is a part of the criteria that facebook's newsfeed is based upon.
- 5) **Using Presence of emoticons in post** : We use a lot of emoticons and hence relate with them better. So, presence of emoticons would mean more likes.
- 6) **Using facebook activity** : If a person is very socially active on facebook with recent updates and more than 400 number of friends, it is more than likely that the usual number of likes on his posts will increase with time.

Mathematical formulas Used:

We have used the basic hypothesis that if an user likes a post then it will also like a similar post containing similar keywords. The more similar posts liked by him concludes much higher probability of liking another similar post.

We have further used root mean square formula to predict the final number of averaged likes on a user post.