PROJECT 4

CSE 535

INFORMATION RETRIEVAL

By- Latent\_

**Requirement 1** : Influencer Score is calculated in this phase of the requirement for every “tweet” collected in the dataset.

The idea to give weightage to every field of twitter based on its value to its corresponding tweet brought us the solution calculate IS based on this formula-

IS = 0.35 \* No. of retweets + 0.1 \* likes + 0.35 \*followers count + 0.1 \*sentiment count + 0.1 \* verified account

Using the above formula, the Influencer Score is calculated for every tweet and are ranked on the basis of that.

**Requirement 2 :**

**Requirement 3 :** We have provided insights based on the analysis of above requirements for-

1. Sentiment Analysis: It was done using “vaderSentimentAnalysis” package and a score was obtained for every tweet(during loading). For context in every language was translated to English using “googletranslate” and sentiment score was calculated.

For score > 0.2 , it is labelled as positive

For score < -0.2, it is labelled as negative and

-0.2 to 0.2 as neutral.

1. News Articles: News articles were fetched using “newsapi”, as they are related to search with 60 days limit.
2. Visualization:
3. Mvrov

**Requirement 4:** Multi faceted search: This on choice search was implemented where a user can search for any specific country or any specific POI or any combination of the two to get the desired results of that particular query.