# Feature Extraction and Sentimental Analysis of text and emoticon using text mining

### Project Proposal

## 

## Supervisor

Sir Umar Khattak

## Submitted by

Tariq Aziz

{01-235161-067}

Omer Farooq

{01-235161-081}

**Department of Computer Science,**

Bahria University, Islamabad

# Introduction

In the modern era where everything is digitalized and social media has emerged to be the platform to raise user’s opinion it is important for the business now more than ever to start monitoring the opinions of their audiences. It is becoming an important part of market research as you can know what your audience thinks about your product and service. That’s why we are creating a system that will analyze the audience sentiments i.e. sentiments is the underlying feeling associated with the opinions and emotions i.e. whether the sentiments were negative positive or neutral and whether the emotions were anger, fear, guilt, affections, sadness etc.

Further it will also extracts the important terms to a related to the tweet and also name entity recognition to provide the business insight of the user sentiments.

# Objective

**“**To provide an insight of the user sentiments”.

# Problem Description

Incredible amount of data is generated every day on the internet i.e. 2.5 quintillion bytes of data to be precise out of which 80% is unstructured. Most of data which is regarding emails, social media, articles, survey, blogs etc. Data that doesn’t make sense on its own, but impacts the audience. Data that can be expensive and time-consuming to analyze. To make sense of all this unstructured and noisy data we need a system that will provide an insight of this unstructured data.

# Methodology

-We will be using incremental methodology in our project.

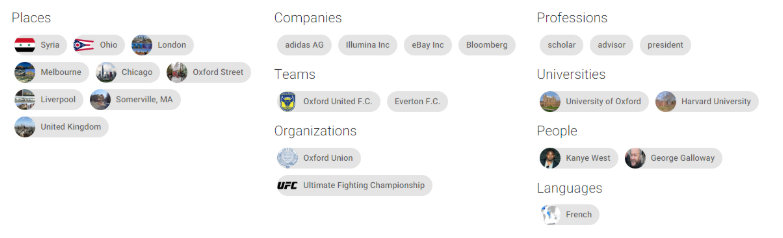
-We will be using NLP techniques like

* Data Processing
* Data Filtering
* Stemming
* Feature Extraction
* Sentimental Analysis

# Project Scope

* Data will be fetched from the twitter API.
* The System will process and filter the data.
* Text mining will be performed for the feature extraction and name entity extraction.
* Complete and word by word and emoticons analysis will be displayed to the user using data visualization.
* Location based Search will also be provided.

Feature Extraction of the most spoken words with Stanford:



# Feasibility Study

* + 1. **Risks Involved**:
* + 1. **Resource Requirement**

# Solution Application Areas

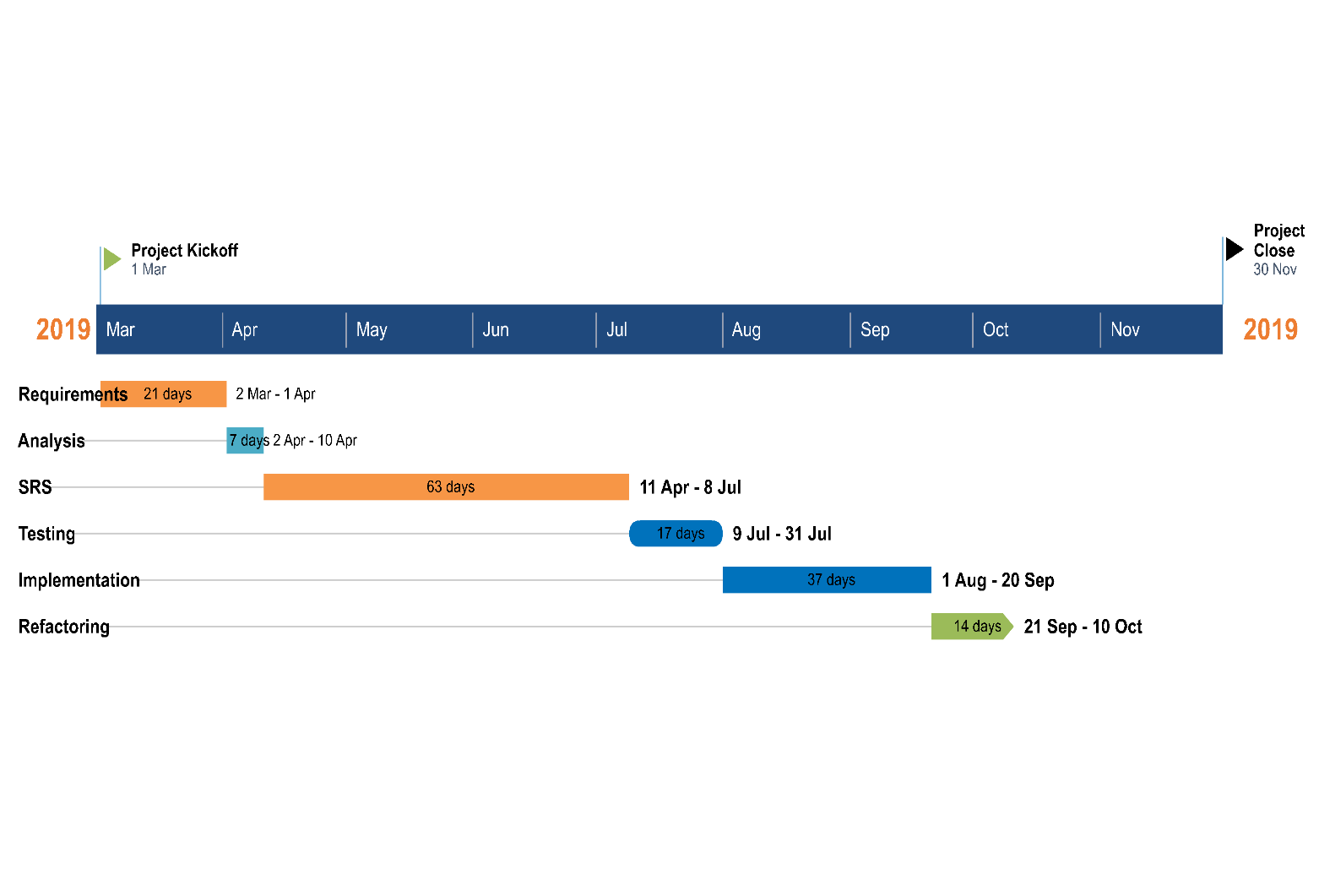
Sentimental Analysis has a wide of application in branding, customer service business intelligence etc.

# Tools/Technology

* Nodejs
* Angular
* Html,CSS,Bootstrap,Ajax
* Google Maps

# The Expertise of the Team Members:

# Milestones



# References

* Capture and share the world's moments. (2019, February 24). *Instagram*. Retrieved from Instagram: https://www.instagram.com/
* Hingorani, M. (2016, ). *mat.ucsb.* Retrieved February 24, 2019, from mat.ucsb: https://www.mat.ucsb.edu/Masters/Mohit\_Hingorani\_Web\_Thesis.pdf
* *https://web.500px.com/*. (2019, February 24). Retrieved from 500px: https://web.500px.com/
* Lars Kai Hansen, A. A. (2011). Good Friends, Bad News - Affect and Virality on Twitter. *FutureTech 2011. The 6th International Conference on Future Information Technology.* *185*, pp. 34-43. Loutraki, Greece: Springer Science+Business Media. Retrieved February 2019, from https://research.cbs.dk/en/publications/good-friends-bad-news-affect-and-virality-in-twitter
* *Localgrapher*. (2019, February 24). Retrieved from Localgrapher: https://www.localgrapher.com/
* The Global Network for Photographers. (2019, February 24). *500px*. Retrieved from 500px: https://web.500px.com/