**T.C.**

**BAHÇEŞEHİR UNIVERSITY**

**FACULTY OF ENGINEERING AND NATURAL SCIENCES**

**DEPARTMENT OF COMPUTER ENGINEERING**

**PROJECT PROPOSAL**

**SENTIMENT ANALYSIS**

**Capstone Project**

**Yunus Emre Aydar**

**Cem Konuk**

**Advisor: Ph.D Tevfik Aytekin**

**İSTANBUL, February 2017**

# TABLE OF CONTENTS

TABLE OF CONTENTS ii

LIST OF TABLES iii

1. OVERVIEW 1

1.1. Description of the Project 1

1.2. Literature Review 1

1.3. Goals 2

2. TECHNICAL CONSIDERATIONS 2

2.1. Technical Limitations 2

2.2. Facilities and Components 2

2.3. Product Design and Safety 2

2.4. Technical Problems 3

2.5 Product Specification and Verification 3

3. WORK PLAN 3

3.1 Deliverables and Division of Tasks 3

3.2 Tasks and Time Line 4

3.3 Cost Proposal 4

REFERENCES 5

# LIST OF TABLES

Table 1. The project Gantt chart. 4

# 1. OVERVIEW

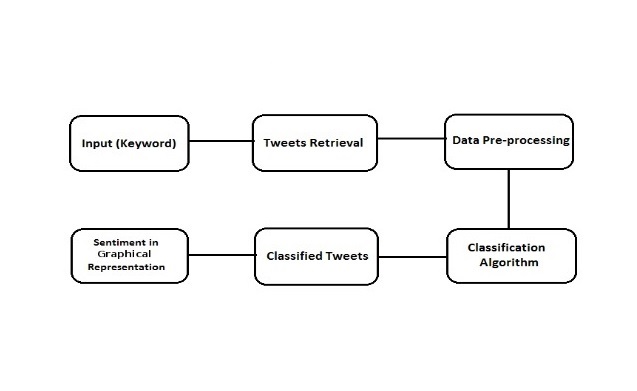
## 1.1. Description of the Project

Sentiment analysis is the process of analysing natural language text for discovering

the opinions or emotions (e.g., positive, negative, or neutral) of the users on specific topics or

products. In this project the aim to collect tweets from the Twitter stream (some other datasets

will also be used) and do basic sentiment analysis.



## 1.2. Literature Review

Sentiment analysis refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. Natural language processing is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between computers and human language. There are many examples of sentiment analysis being used in a variety of fields such as online retail, to all forms of blogging to find out the opinion of people on a specific topic. Differently on our product, we are planning to use the locations of tweets to learn predominant regions of negative opinions, and create a match system to compare two similar words and compute the ratio between positive, and negative opinion.

**1.3. Goals**

Our product obtains data of people's opinion in order to help clients for learning a product's review.

# 2. TECHNICAL CONSIDERATIONS

## 2.1. Technical Limitations

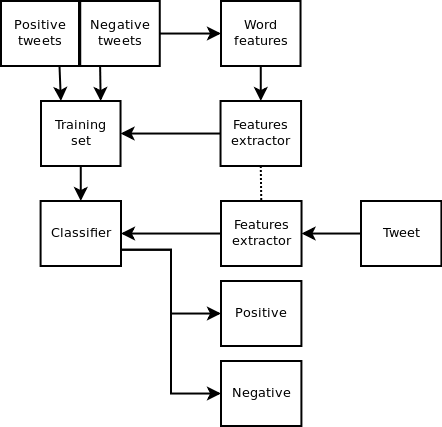
Problem of synchronization, incompatibility between different program, and learning different programing languages might limitiations that affect the goals of our project.

## 2.2. Facilities and Components

We will be using Natural Language Toolkit, programming languages such as C++, Python, Java , and a software that retrieve data from Twitter.

## 2.3. Product Design and Safety

Firstly, we will take tweets using open source Twitter Libraries. Then, the program will decide whether positive or negative words used mostly by classifier.After processing every tweet, we will show the overall positive and negative scores with a graphical user interface.



## 

## 2.4. Technical Problems

Problem of synchronization between programs and toolkits, incompatibility between different programming languages might be the technical difficulties that we could be facing in this project.

## 2.5 Product Specification and Verification

With the help of the language toolkit, program will probably succeed to pass classification phase and separete negative and positive words.

# 3. WORK PLAN

## 3.1 Deliverables and Division of Tasks

Deliverables will be divided into three main stages:

1. Research

Research will be about language kits, getting data from Twitter and most approppriate programing language for this project.

2. Design and Coding

Design and Coding will include discovering desing techniques, discovering new libraries, assembling systems, and integrating all systems.

3. Testing and Documentation

Testing and Documentation will involve preparing proposal, verification, preparing report, and presentation parts.

## 

## 3.2 Tasks and Time Line

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | WEEKS | | | | | | | | | | | | | | | | |
| TASK LIST | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** |
| **Research** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Define the goals, and perform a literature search. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Searching languages and preparing the proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Design and Coding** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Assemble system A |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Assemble system B |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Preparing Software |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Integrate all systems |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Testing and Documentation** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Prepare the proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Verification |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Prepare the report |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Presentation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

## 3.3 Cost Proposal

Cost Proposal was not designated.

**REFENRENCES**

1. Wikipedia, Link : <https://en.wikipedia.org/wiki/Natural_language_processing>
2. SlideShare, Link: <https://www.slideshare.net/sumit786raj/sentiment-analysis-of-twitter-data>
3. Laurent Luce’s Blog, Link: <http://www.laurentluce.com/posts/twitter-sentiment-analysis-using-python-and-nltk/>
4. F. Javier Alba Blog, Link: <http://fjavieralba.com/basic-sentiment-analysis-with-python.html>