FUNCTIONAL SPECIFICATION

TweetMine – Twitter Sentiment Analysis Tool

KRZYSZTOF OBLAK
C00161361

Table of Contents

1. Introduction 1
   1.1. Purpose 1
   1.2. Scope 1
2. Vision 1
   2.1. Project Description 1
   2.2. Deliverables/Outcomes 1
   2.3. Functionality & Technology 1
   2.4. Audience 2
   2.5. Data, Information & Content 2
3. Use Case Diagram 2
4. Brief Use Cases 3
   4.1. General Search 3
   4.2. Search Cloud 3
   4.3. Recent Search 3
5. Detailed Use Cases 3
   5.1. General Search 3
   5.2. Search Cloud 4
   5.3. Recent Search 4
6. Supplementary Specification 5
   6.1. Usability 5
   6.2. Reliability 5
   6.3. Performance 5
   6.4. Supportability 5
   6.5. Security 5

1. Introduction

1.1. Purpose

The purpose of this document is to summarise the functional requirements of TweetMine a Twitter sentiment analysis tool. It is not a system solution, but a guideline of the required system functionality.

1.2. Scope

This manual provides details of the core functionalities of the project; it does not include any design aspects of the app.

2. Vision

2.1. Project Description

Sentiment analysis for social media (Twitter) is a web based application that will let its users to perform a keyword search to obtain information outlining opinions contained within messages (tweets) posted on Twitter. Those messages will be then categorised based on their nature (positive, negative, neutral) as a result of sentiment analysis and natural language processing techniques.

2.2. Deliverables/Outcomes

1. A web based service able to handle multiple users at the same time.
2. A search engine connected to Twitter API.
3. Sentiment analysis module able to process tweets and mine opinion.
4. A feedback page to present the results to the user.
5. Usage and trends statistics.

2.3. Functionality & Technology

The service will require each potential user to sign in with their respective Twitter account thus eliminating the need for a registration capabilities and associated process of validation in order to use the service as all should be handled by Twitter API. After signing into the service the user will be presented with a simple search engine created with the use of web technologies such as HTML5 and CSS3 along with possible use of jQuery running in the background. Upon submission of a search query the engine will translate its query onto Twitter API call and retrieve a number of tweets as a result. The result will be run through a natural language processing module which will scan through the messages along with associated tags (#'s) and evaluate based on keyword

occurrences, patterns and phrases the nature of the tweets. The processed result will be then presented to the user in a friendly format. An additional feature for the service would be a small statistics on the searches performed, trends and current user count presented to a user at all times after signing in by the use of jQuery framework.

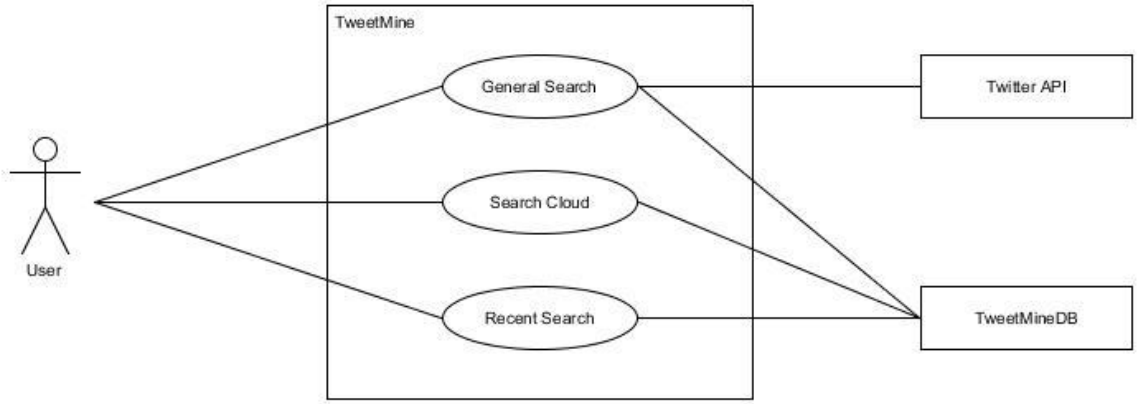
2.4. Audience

The service is intended to be used by both marketing industry professionals who need to know valuable feedback about their products and services as well as ordinary consumers who seek feedback from others.

2.5. Data, Information & Content

Twitter short messages commonly known as 'tweets' which contain the date, personal screen name along with a text message. Additional statistics to be presented on the page for quick search access as well as informative content.

3. Use Case Diagram



4. Brief Use Cases

4.1. General Search

Name: General Search
Actors: User, Twitter API, TweetMineDB
Description: When a user wishes to use the search functionality of the system, they must first provide one or more keywords as well as to pick a result limit from a dropdown up to max of 500 results (default is 100) to be returned. The user then clicks the search button to submit the form displayed on main application screen. Supplied keywords will be added or have their counters updated in the database. The API will perform the search query and return a list of tweets containing the specified keywords.

4.2. Search Cloud

Name: Search Cloud
Actors: User, TweetMineDB
Description: When a user wishes to view the entire list of search keywords used in the past, they must first click on the appropriate link on the left panel displayed on main application screen. The database will be contacted and a query for all keywords will be run. The set of results containing the keywords and their respective counter values will be returned and presented to the user. If a user clicks on any of the keywords presented a general search will be performed for that keyword.

4.3. Recent Search

Name: Recent Search
Actors: User, TweetMineDB
Description: If a user wishes to use one of the recent search keywords used in the past, they can do so by clicking on one out of five buttons located at the top-left section of the main application screen. Those buttons correspond to the five most recent keywords retrieved from the database based on their dates of storage. If a user clicks on any of the five keywords presented a general search will be performed for that keyword.

5. Detailed Use Cases

5.1. General Search

Name: Recent Search
Actors: User, Twitter API, TweetMineDB
Main Success Scenario:
1. The user enters one or more keywords into the search form.

2. The user will set the result limit.
3. The user will click the search button.
4. The system will contact the Twitter API in order to perform the search.
5. The Twitter API will return a set of result up to the specified limit.
6. The system will contact the database in order store the keywords or update its counter and date values.
7. The system will perform sentiment analysis on the data provided.
8. The system will display a result page to the user and the use case ends.

Alternatives:

1a. No keywords were supplied.
    1. The system will display a blank result page to the user and the use case ends.

5.2. Search Cloud

Name: Search Cloud
Actors: User, TweetMineDB
Main Success Scenario:

1. The user clicks on the search cloud button on the left panel.
2. The system will contact the database in order to retrieve all of the keywords and their counters.
3. The system will order the data using ascending order ignoring the case.
4. The system will display the formatted data to the user.

Alternatives:

2a. The system was unable to connect to the database.
    1. The system will display an alert outlining the issue and the use case ends.

5.3. Recent Search

Name: Recent Search
Actors: User, TweetMineDB
Main Success Scenario:

1. The system will contact the database in order to retrieve five most recent keywords.
2. The system will order the data based on the date attribute.
3. The system will display the formatted data to the user.

Alternatives:

1a. The system was unable to connect to the database.
    2. The system will display an alert outlining the issue and the use case ends.

6. Supplementary Specification

6.1. Usability

The GUI for the website must be easy to navigate around and provide simplistic design complying with related standards and also must be uniform throughout the site to allow the users to quickly familiarise with the website

6.2. Reliability

The system must be available 24 hours a day 7 days a week as it may be used at any time, throughout many various locations to access all available options with less than 5% downtime.

6.3. Performance

The system must be able to support multiple users. There's no initial target user base but the system should be able to support more than 100's of users and allow for expansion to different social networks such as Facebook & Google+. The response time of the system must allow for all actions to be performed in real time, the results must be available for review instantly.

6.4. Supportability

The website must be compatible with all major desktop browsers and mobile devices such as smartphones and tablets. The overall look and feel as well as usage must be done the same way across platforms.

6.5. Security

The website uses OAuth 2.0 protocol to contact and interact with the Twitter API using application-only authentication therefore user accounts and login facilities are not required.