AUTOMATED PROCESSING FOR SOCIAL MEDIA DATA IN A MASS EMERGENCY

Software Requirements Specification

Criticality analyzing component

Project ID: **18-007**

Perera P.A.D - IT14093210

B.Sc. Special (Honors) Degree in Information Technology Specializing in Software Engineering

Table of Content

1. Introduction

1.1 Purpose

The purpose of this SRS document is to outline the requirements and present a detailed description of the process needed for the system that named **Automated Processing for Social Media Data in a Mass Emergency.** The document contains the necessary requirements of the system, as well as the process to create and discover them. It will explain the functional and non-functional requirements, purpose and features of the component, the interfaces of the component, design constraints, project approach, what the component will do, the constraints under which it must operate and how the component will interact with other external applications. The information is organized in such a way that the developers will not only understand the boundaries within which they need to work, but also what functionality needs to be developed and in what order. This document is intended to be proposed to a user for approval and also this document is targeting the, designers, developers and other stakeholders as its audience. [3]

1.2 Scope

The component to be produced called "Criticality analyzer". This component predict the critical level of the situation by analyzing social media post. Basically in disaster management one of main problem is to identify and prioritize the people how need help. Because there may be a people who are in a very critical situation and they need help before others. The main objective of having this component in the system is to identify the people who need help first.

As per reason literature, almost all the disaster management systems used the sentiment analysis technique to show the emotions of the people who face to the disaster situation going on. But In this system we use sentiment analysis algorithm to predict that how critical the situation is.

For do that, we need to develop a new sentiment analysis algorithm which can predict the criticality of the situation going on. To develop this component I will be use the natural language processing techniques. The out come of the algorithm will be a percentage of the the criticality.

The main purpose of this component is Help supporting team to categorize and prioritize the people who need help first. it will really helpful for supporting teams to get the idea about criticality of the situation and to get the correct decisions.

1.3 Definitions, Acronyms, and Abbreviations

NLP	Natural Language Processing
UI	User Interface
DB	Data Base
API	Application Program Interface
PC	Personal Computer
RAM	Random Access Memory
JSON	Javascript Object Notation

1.4 References

[1] Bing Liu. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 2012.

[2] Yelena Mejova, Ingmar Weber, and Michael W Macy. Twitter: A Digital Socioscope. Cambridge University Press, 2015.

[3] Bruce R. Lindsay. Social Media and Disasters: Current Uses,Future Options, and Policy Considerations. Technical report, Congressional Research Service, September.

[4] Ahmed Nagy and Jeannie Stamberger. Crowd sentiment detection during disasters and crises. In Proceedings of the 9th International ISCRAM Conference, pages 1–9, 2012

1.5 Overview

The remainder of this SRS document includes three sections and appendixes. The second section provides an overall view of the component functionality and interaction with other components. This section also discusses the specific requirements such as functional and nonfunctional requirements, design constraints and various approaches. Furthermore, this section also mentions the system constraints, User characteristics and assumptions about the product. The third section provides the requirements specification in detail and a description of the different interfaces. Different specification techniques are used in order to specify the requirements more clearly for different audiences.

The rest of the sections that organized this document are Project perspective and descriptions, Different interfaces that the system consists, Requirements of the system, Summary of major functionality, Users and characteristics of the system and the background of the general factors affect the system.

2 Overall Description

Criticality Analyzer

This component will provide the criticality level of the situation going on by analyzing social media post. As a input of this component it will talk the social media post that filtered and cleared from the upper component of the system. This component will be develop using NLP techniques and will use natural language processing toolkit of python language.

There will be a new classification algorithm which can predict criticality level of the situation going on. For develop that algorithm I will be maintain a data set library to train the model. Each of the word in the library given a weigh and that data set library will maintain by using mongoDB. By using the data set model will be trained. Then by using that trained model it can be predict the critical level of the new social media post which filtered from upper component.

The output of the component will be show in the UI tool and it can be reached outsiders by the API.

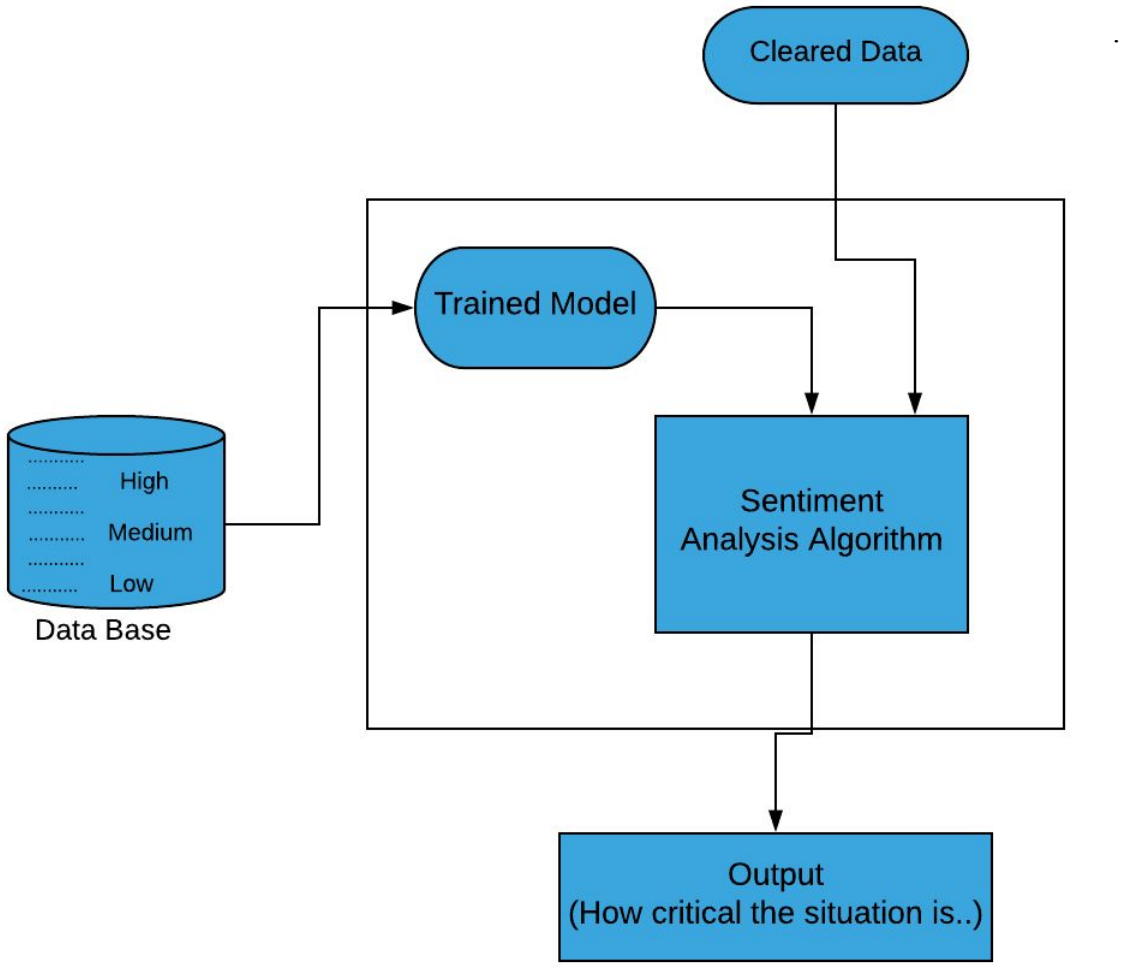

Figure 1: Critivality Analyzing Component

2.1 Product Perspective

There are several application around the world which are perform disaster management in social media. The have also used NLP techniques, Sentiment analysis algorithms and semantic analysis algorithms. But in those systems, they have use sentiment analysis and semantic analysis algorithm to show the feeling of the peoples. Those system haven't built mechanism to predict the criticality level of the situation going on.

In this component I will use sentiment analysis and semantic algorithms for predict the how critical the situation is. It will really helpful for supporting teams to get the idea about criticality of the situation and to get the correct decisions. Because if the supporting teams know the critical level of the situation and how are the people need help first they can be ready for the situation before they go to the area where the disaster is happening. That will really helpful for saving people's lives which is a very important thing in disaster management. [1], [2],[4]

Features	Twitris	Senseplace 2	EMERSE	AIDR	Proposed System
Automated Classification	✓	✓	✓	✓	✓
Prioritizing	✗	✗	✗	✗	✓
Criticality Analysis	✗	✗	✗	✗	✓
Accuracy Validation	✗	✗	✗	✗	✓
Text Summarization	✗	✗	✗	✗	✓

2.1.1 System interfaces

To develop the criticality analysis algorithm system will refer several NLP base APIs.

- WordLift API
- Indata lab API

2.1.2 User Interfaces

The main outcome of the system is a publicly available API which is user can get the data processed by the system. This API is publicly available for outsiders. They can do there work interdependently by interacting with the API.

2.1.3 Hardware Interfaces

- Modern PC with enough performance.

2.1.4 Software Interfaces

- Windows 7 or higher version

2.1.5 Communication Interfaces

- Internet facility to call the API and get data.

2.1.6 Memory Constraints

To call the API and get data, minimum memory requirement is 512 MB of RAM in the PC.

2.1.7 Operations

To access the API user need to call the rest API by passing relevant parameters. Basically it will be traditional get request.

2.2 Product functions

Use case diagram for criticality analysis component

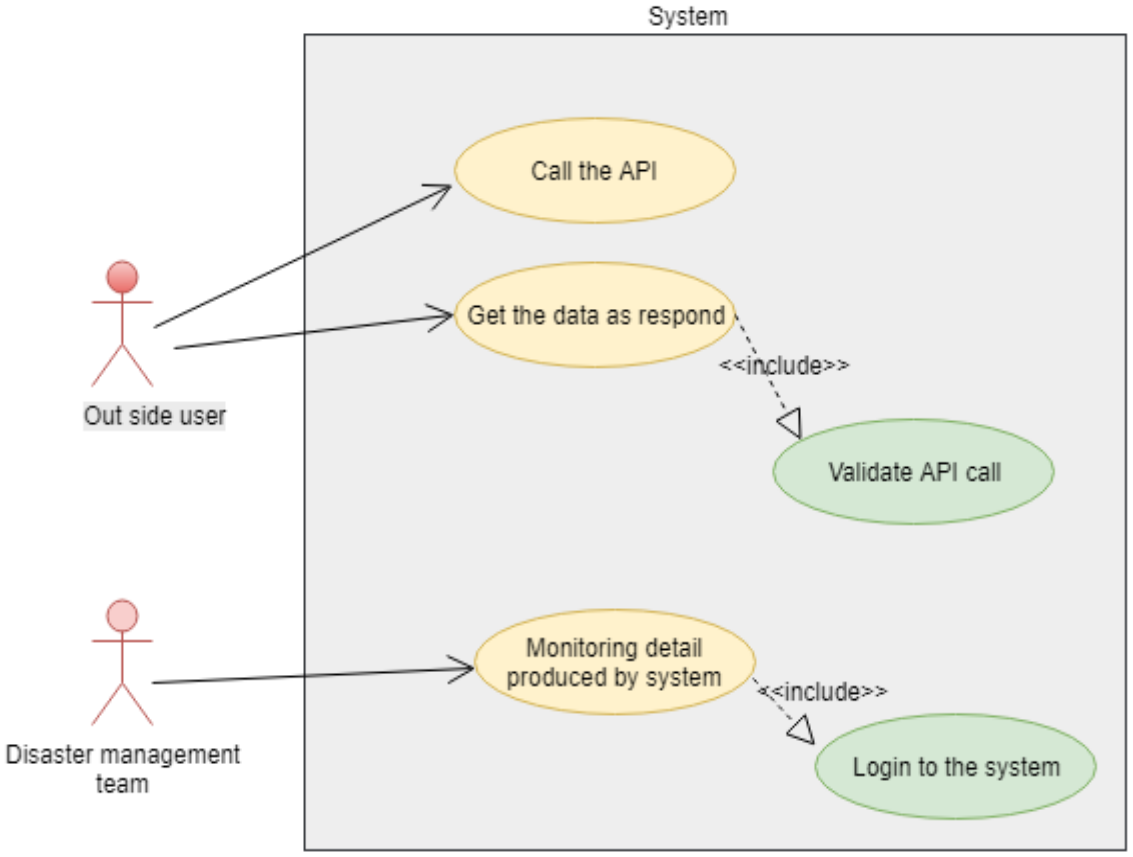

Figure 2: Use case diagram for Criticality analysis component

Use case diagram illustrates the requirements of the system from user's perspective. This system is giving a public API to the outside users and there is a UI tool for inside used. Out side user interact with the system only trough that API. Disaster management teams use UI tool to monitoring the disaster situation.

Use Case Scenarios

Use case name	Call the API
Goal	Getting possessed data
Pre-conditions	Internet connection must be stable User passed correct parameters
Actor	User
Main success scenario	1. Get the processed data abut the disaster situation going on
Extensions	

Use case name	Monitoring details produced by the system
Goal	Monitoring details of disaster situations going on
Pre-conditions	Log in to the system Internet connection is stable
Actor	Disaster management teams
Main success scenario	Find the details of disaster situation
Extensions	

2.3 User characteristics

This system will used by two type of users primarily. First type is the disaster management supporting teams who using the system to monitoring disaster situations. Second type of users will be the outsiders who need possessed disaster management data for there systems.

First type users use the UI tool to interact with the system and second type of users use publicly available API to interact with the system.

2.4 Constraints

- Information exchange rate depends on the specification of the PC and connectivity options available to the user.
- Internet connection must be stable.

2.5 Assumptions and dependencies

- Out side users have the knowledge how JSON object works.
- Out side user familiar with how to make API call.
- Connectivity between the system components is secure.

2.6 Apportioning of requirements

In section 1.5 it described the overview of the proposed system and the section 2 provided the detailed overall description about the system and the requirements. Section 3 contains detailed requirements that should be followed while designing those requirements. The methodology of implementing the system may slightly different from the content described in this document. During the system designing, requirements specified will not be changed and the system released will totally contains its purpose and objectives.

3 Specific requirements

3.1 External interface requirements

3.1.1 User interfaces

In this system the main outcome is publicly available API. But for testing purpose and for usage of disaster management supporting teams there will be a UI tool as well. This is a mock-up of the UI tool.

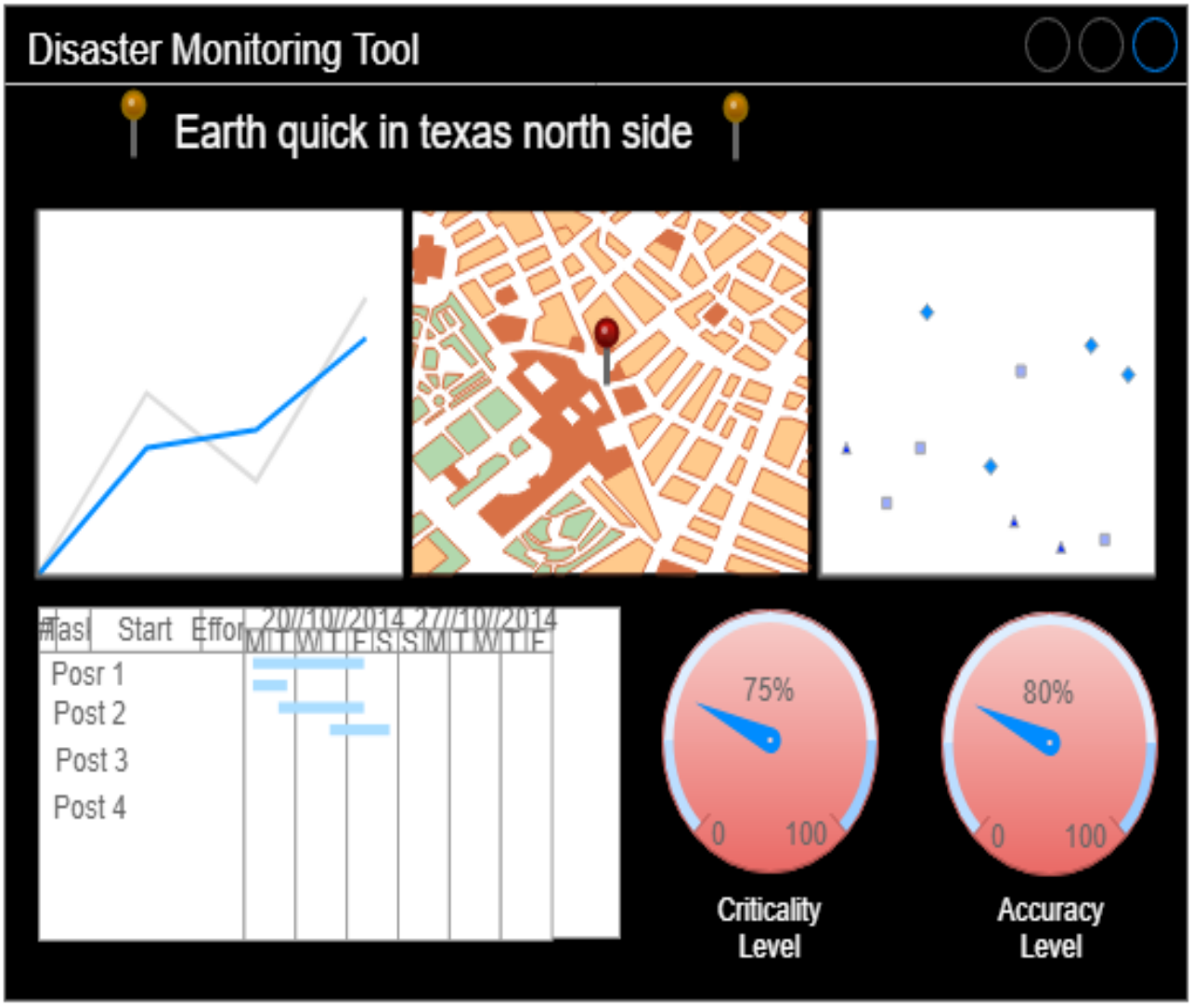

Figure 3:User Interface of UI tool of the system

Other interface in client side is API interface. It is the only layer between outside user and the system. It is showing in this diagram.

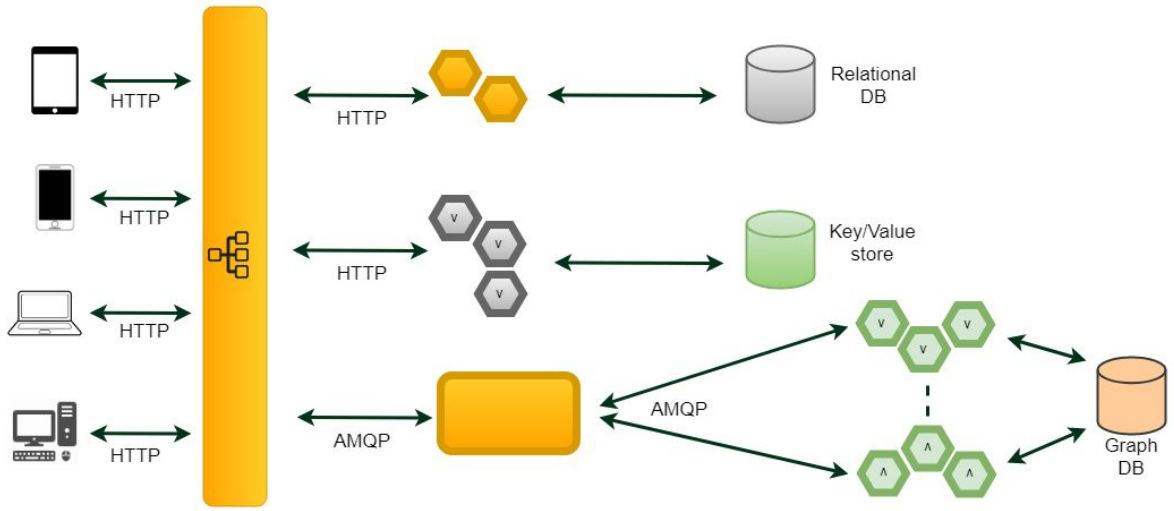

Figure 4:API interface of the system

3.1.2 Hardware interfaces

PC which has the operating system window 7 or above and minimum RAM of 1GB (this is only to access the the API or install the UI tool).

3.1.3 Software interfaces

jupyter notebook with python installed.

NetBeans IDE with Java to develop the UI tool

MongoDB

3.1.4 Communication interfaces

Internet router or modem with stable internet connection

3.2 Functions

Criticality level prediction algorithm

This function is to predict how critical the situation is. The main purpose of the function is to help supporting teams to identify and prioritize people who need help first.

The main data source of the function will be a data set library which is maintain in mongoDB server. Each of the words in that data set given a weight. Using that data set library model will be train.

Then that model can predict the critical level of the situation by analyzing new social media posts which are cleared and filtered from upper component of the system.

UI tool development

This function is to develop the front end part of the criticality analysis component. The criticality level of the situation shown as graphically. As well as user can see the numerical value of the criticality level.

API development

This function is to develop API part of the criticality analysis component. It will give the result of the component for out size users. It will be a JSON object.

3.3 Performance requirements

- Computer with at least 1800MHz processing power and 1 GB RAM to install UI tool or access the API.
- Internet access with at least 512kbps connection speed.

3.4 Design Constraints

Programming Language

In machine learning Python is the most stable language. As well as Python has lots of inbuilt methods and libraries which are specified to ML tasks.

Development Tools

Server side will be develop using the jupyer note book and front end side will be develop using NetBeans IDE.

Database

MongoDB will use as the database because key value pairs is help full to maintain word set with the weight.

3.5 Software system attributes

3.5.1 Reliability

Reliability is the probability that an application will accurately perform its specified task under stated environmental conditions. Simply, that is how much a user can depend on the system. So in the proposed system there is a mechanism to check the accuracy of the system predictions. It will handle in separate component.

3.5.2 Availability

The system will be hosted in AWS so, in there we hope to manage backup instance as well. If any failure happen second instance can be use. So it will increase the availability of the system.

3.5.3 Security

Since the system going to be hosted in AWS, there are many security features to protect the user.

3.5.4 Maintainability

Maintainability is defined as the probability of performing a successful repair action within a given time. In other words, maintainability measures the ease and speed with which a system can be restored to operational status after a failure occurs. In this system we use component base architecture. Each main function we develop as a separate component. So if any change to be done, it will easy to maintain.