Software Engineering
Software Requirements Specification (SRS) Document

Contents

1. Introduction

1.1 Introduction

The purpose of this document is to define and describe the requirements of the project and to spell out the system's functionality and its constraints.

1.2 Scope of this Document

The customer and the user for the system are the students, researchers who use the datasets.

1.3 Overview

The product is a website which provides end user analytics like number of times a dataset is being downloaded, number of papers being published for each dataset and standard analytics like mean, median, mode on the datasets in the data foundation.

2. General Description

2.1 Product Functions

The product should make it easy for user to perform analysis on any dataset. The analysis should also be accurate.

2.2 System Information

The system displays a list of available datasets and fetch top 50 rows of the user's chosen dataset and display. The user is provided with analysis of the chosen dataset. The system not only provides analysis of individual dataset but also analysis across different datasets.

2.3 User Characteristics

The users include the students, researchers who wants to use the datasets in the data foundation. To use the website, the users do not require any skills, since the interface would be user friendly.

2.4 User Problem Statement

Studying and understanding the data is very important to get useful and usable information and to make the most informed decisions possible. This data analysis might be difficult when users lack required knowledge or skills.

2.5 User Objectives

The user wants a interface which shows dataset Analysis that makes studying data a lot simpler and more accurate. The datasets are stored in database.

3. Functional Requirements

1. **Standard Analytics.**

When we're trying to describe and summarize a sample of data, we probably start by finding the mean (or average), the median, and the mode of the data. These are central tendency measures

and are often our first look at a dataset. The mean, median, and mode are three metrics that are commonly used to describe the center of a dataset.

- **Mean:** The average value in a dataset.
- **Median:** The middle value in a dataset.
- **Mode:** The most frequently occurring value(s) in a dataset.

Individuals and companies use these metrics all the time in different fields to gain a better understanding of datasets. Few examples are: (a) Insurance analysts and actuaries in the healthcare industry; (b) Used often by real estate agents.

2. **End user analytics.**

Though standard analytics are very useful in understanding many of the datasets, They doesn't make much sense when most of the columns in the dataset are not numerical . Even when values are numerical, standard analytics are of no use in some cases. Example: Covid dataset(which might contain information like "sample id" which is numerical but mean, median doesn't provide any useful information.). So we perform end user analysis on all the datasets in the data foundation. We analyze the usage of all the datasets in various ways.

- **Downloads:** We will be able to record the number of times each dataset is downloaded.
- **Citations**: Keep track of how many papers have been published on each dataset in the data foundation.
- **Visualization:** Perform monthly analysis and show the visualization and report the analysis.
- **Comparison:** Perform comparison on the analysis of different datasets.

5. Software Requirements

Aside from a computer and an internet connection, most of the tools need to build the website are software programs. You need a text or HTML editor, a graphics editor, web browsers.

Required Setup

1. **Git Bash:** The user of all operating systems can use it. All the Django related commands and Unix commands are done through it. For downloading the Git bash.
2. **Text-Editor:** Any Text-Editor like Sublime Text or Visual Studio Code can be used.
3. **Python 3:** The latest version of Python can be downloaded in Download Python.

6. Hardware Requirements

The website is developed using Django and React.

1. Minimum requirements:
   - **Memory:** 4 GB
   - **Graphics Card:** AMD Radeon R5 M230
   - **CPU:** Intel Core i3-2340UE
   - **File Size:** 3 GB
   - **OS:** Windows 10
2. Recommended requirements
   - **Memory:** 8 GB
   - **Graphics Card:** AMD Radeon R7 A10-7850K
   - **CPU:** Intel Core i5-4400E
   - **File Size:** 3 GB
   - **OS:** Windows 10