

CSE 3063 Fall 2020 Java Project 1
Group Number 14
Iteration 1 Requirement Analysis Document

1.Goal

This program is basically a labeling program which analyzes the comments done by users and labels them. This enables users to create labeled and classified datasets. These datasets can be used for Machine Learning models in the future.

2.Problem Statement

- 2.1** The program should support a multi-user system.
- 2.2** There will be input files of instances in JSON format.
- 2.3** The program should read the dataset from these JSON files.
- 2.4** After reading the dataset from JSON files, the program should randomly label the data with the corresponding user ID's.
- 2.5** Since this is a multi labeling system, the user may assign more than one class label ID to a single instance.
- 2.6** After the labeling is done, the program should create a JSON output file.

3.Scope

The program is going to label the given datasets and these datasets may be used for the training of Artificial Intelligence models such as Machine Learning models.

4. Glossary of Terms

4.1 Dataset: A group of instances to be labeled.

4.2 Class Labels: Predetermined categories to assign the instances.

4.3 User: People who make the labeling of the datasets.

4.4 JSON: JavaScriptObjectNotation files that are used to showcase structured data.

5. Use Cases

5.1 Aid researchers while labeling big datasets.

5.2 Help users categorize their own data offline.

5.3 Multiple users can label the same data instance with different labels.