MARMARA UNIVERSITY

FACULTY OF ENGINEERING

DEPARTMENT OF

COMPUTER ENGINEERING

Data Labelling System Requirement Analysis

CSE3063 Object Oriented Software Design

Team
Group Number: 2
150118013 - Gökberk Çelikmasat
150118065 - Şeref Kutay Akgün
150117005 - Kadir Acun
150119825 - Merve Rana Kızıl
150119824 - Zeynep Ferah Akkurt
150117905 - Erhan Yalnız

1.Introduction

Vision: The objective of the project is the implementation of a labeling system simulation with Java based object-oriented software design approach. The project is designed and developed by a team of 6 members.

Scope: Data Labelling System Simulation

Special Requirements

- Java
- Visual Studio
- Eclipse
- Git
- Intellij
- PlantUML
- Json-simple.jar
- Visual paradigm
- Log4j-api.jar
- Log4j-core.jar

Stakeholders:

- Murat Can Ganiz (Customer)
- Lokman Altın (Customer)
- Gökberk Çelikmasat
- Şeref Kutay Akgün
- Kadir Acun
- Merve Rana Kızıl
- Zeynep Ferah Akkurt
- Erhan Yalnız

2. Requirement Specifications

2.1 Functional Requirements

2.1.1 Must Have Requirements

- A configuration file must be created and copied into the source folder of the project.
- It is a multi-user system.
- An instance can be labeled by one or more users.
- An input file with a sample dataset that consists of meta data such as instances, labels and dataset information.
- External libraries like Json-simple.jar, Log4j-api.jar and Log4j-core.jar.
- Labeling operations will be done with command prompt.
- For the first iteration labeling will be done randomly only for random bots.
- System will print its actions to the command line and a log file one by one.

2.1.2 Could Have Requirements

- Improved artificial intelligence of the simulated labeling operations.
- There can be different types of labeling operations for different types of user types.

2.2 Non Functional Requirements

- Performance
- Portability
- Availability
- Security
- Safety
- Maintainability
- Reusability

2.3 Simulation Rules/Specifications

This is an insight into the way the group aims to achieve the implementation of the labeling system. This system provides information about different outcomes which about the labels may occur at the end of the simulation to customer with different types of users. This system may provide an information about how to enhance and upgrade the labeling mechanism. This is not a comprehensive list; it will be updated once the team has prepared the detailed design, class diagram and the sequence diagrams for the simulation.

2.3.1 Simulation Guidelines

- The system starts the simulation with reading configuration and input files.
- There will be simulated users which will take turns and label the instances one by one, one after another.
- In this system, users can be more than one different type.
- A user can label many instances.
- An instance can be labeled by one or more users.
- Datasets have an information about maximum number and according to this information, an instance can be labeled as much as this number.
- Number of labels per instance cannot be less than one.
- User randomly chooses one of the labels from the set of labels and chooses an instance which will be tagged.
- The system gives the output about all labeling operation, after the execution.

3. Glossary

- Instance: Like samples, examples, records, documents.
- Label: A descriptive or identifying word or phrase.
- Dataset: Group of instances and labels.
- LabelingAssignment: Holds required information about instances.
- LabelSystem: A system which holds all labeling operations.
- User: A type of bot that is behave like human who labels the instances.
- RandomLabeling: Type of bot that labels randomly.

4.DOMAIN MODEL