CSE3063 Object Oriented Software Design / Java Project 3rd

Iteration Requirement Analysis

Group Members

150119752 - Ayça İdil Kaya

150119620 - Hafsa Ulusal

150119073-Zeynep Naz Akyokuş

150119728-Ecem Çeşmeciler

150119863- Taha Kızmaz

150119634-Mustafa Gürel

150119654-Mehmet Nusret Odabaşı

Table of Contents

1. About the Project

This project focuses on textual data labeling which is the process of identifying textual data. Project's main goal is to display a platform that clients can label and see the given labels.

Although the dataset used in the testing phase involves movies and genres, the main aim of this project is to be strictly in object oriented manner so that any given dataset (with the same input labels) will work on the project.

2. Requirements

2.1-Functional Requirements

- ***User authentication***
  - Priority(1-lowest/5-highest): 5
  - Criticality: In order to use the system client must register first

- ***Labeling***
  - Priority: 5
  - Criticality: This the main idea behind the project

2.2-Non-functional Requirements

- This project will be a multi-user system.
- There will be several labeling mechanisms.
- Labeling will be done randomly in the first phase, which will assign genres randomly to the text.
- Every text can be labeled by various users.
- Every text can have more than one label.
- A user will be able label a particular instance more than once in order to measure their consistency
- For a user, one of the labeled instances by the user will be shown to this user again with a certain probability in a dataset as an independent parameter. (i.e. ConsistencyCheckProbability)
- Clients will be able to add users through the config file.
- Client will be able to add datasets through config file
- Client will be able to assign any number of existing users in config json to a particular dataset for labelling.

- Client will be able to set an existing dataset to start labeling simulation from the config json file.
- If an instance is labeled more than once by the same user or other users, the most frequent class label will be assigned as the final label.
- Clients will be able to stop the simulation at any time and access reports.
- Users' logins and labeling will be handled through the command line.
- Since there will be many human users, user credentials will be stored in the config file along with the ConsistencyCheckProbability.
- The user will be able to terminate the system any time they want.
- If the username and password are entered blank, the system will act as it did in the 2nd Iteration, meaning bot users will handle random labeling.
- Additional bot users will be added.
- System will ask for a username and password. If the username and password entered matches one of the user credentials in our config file then this human user will be able to label instances in the current dataset one by one.
- If the username and password does not match any user credentials in our config file then the wrong username and password message will be displayed and the system will ask the username and password again.

3. Domain Model