MARMARA UNIVERSITY
FACULTY OF ENGINEERING
DEPARTMENT OF COMPUTER ENGINEERING

CSE3063
OBJECT ORIENTED SOFTWARE DESIGN
PROJECT 1 – ITERATION 2
REQUIREMENT ANALYSIS

PROJECT NAME:
DATA LABELING SYSTEM

Vision

The objective of the this project is to create a Java based object oriented implementation of Data Labeling System. This project will run on terminal where necessary parameters are taken from the user.

The project has three main iterations. According to the demands and feedback of the customer, necessary changes and developments will be made in the project.
The requirements and feedback will be listed in each step. After each iteration changes between two iterations will measure depending on some metrics.

Scope

Data labeling is the process of assigning one of the several predetermined labels such as class labels, categories, tags to a group of instances through a user interface by human experts. A group of instances are known as a dataset.

There can be several labeling mechanisms. A labeling mechanism takes a user, a single instance and a set of class labels as an input and one of these labels are assigned to the instance. Then this mechanism returns the assigned label or labels associated with the given user. If words or phrases are labeled it will return a set of pairs, each pair including word or phrase and its label(s).

System Constraints

- This project runs on any Java based platform.
- It run as a simulation on the console with any device that has Java Runtime
- Environment installed.
- This system runs from command line and also prints its actions to the command line
- and a log file one by one.
- It runs without GUI part.
- There will be no databases in this project. Any database system can not be used.
- All requirements are not known in the beginning. Rest of the requirements will be
- revealed during next iteration.
- Datasets are taken from a json file.

Rules

- This project is designed as a multi-user system.
- There are several labeling mechanism.
- A user can label many instances.
- A user can assign more than one class labels to an instance.
- An instance can be labeled by one or more users.
- A single instance can be assigned more than one class labels.
- There must be at least 3 users in the config.json file

Stakeholders

- Murat Can GANİZ (Customer)
- Lokman ALTIN (Customer)
- Veysi ÖZ
- Muhammet Yasin TUFAN
- Mikail TORUN
- Ahmet Hamza DEMİR
- Enes GARİP
- Abbas Göktuğ YILMAZ
- Belgin TAŞTAN

Glossary of Terms

Data Labeling : Refers to the process of segmenting and assigning labels to data

Random Labeling Mechanism: Randomly chose one of the labels from the set of labels and assigns it to the instance

Data Set: A group of instances

Data Set: File: Includes the set of labels, maximum number of labels to tag for an instance, a set of instances

User: A simulated person who labels the instances

Label: Identifies data features

Instance: Refers to the data that are taken from user

Printer: Prints instances, labels(ID) and datasets received

Starter: The class that works as a thread, waiting for the user to press the enter key while the program is running, and terminates the system when the enter is pressed

Config: Parse where config.json file is read

Technologies and Control Mechanisms

- Github

Project Deadlines

Iteration I – 5 Aralık 2020
Iteration II – 24 Aralık 2020
Iteration III – 2 Ocak 2021

Domain model