REQUIREMENT ANALYSIS DOCUMENT

Vision:

Data labeling is a process of classfining instances(text,images,recordes) and assigning each group of instances to a suitable label - category - using a classification model chosen by expert users. This process gave us a result dataset which usually helps in training machine learning models which use these datasets to understand the pattern and take a decision based on the predicted pattern.

Scope:

The data labeling project will provide :

- Multi-users will be able to use the program and choose a labeling mechanism .
- For the first iteration the program will support the random labeling mechanism only.
- The program will start by reading the dataset from a json file and start classifying them one by one .
- For some instances multi-labeling will be provided with a limited number of labels for each instance .
- After the classification process is done the program will create a json file containing all the instances with suitable labels and user information .

System constraints

GUI part is not included at this iteration . The required displayments are made in the console.

Project Members:

- 150117046 Ahmet TURGUT
- 150117048 Fatih BAŞ
- 150117010 Süleyman Ahmet SÖNMEZ
- 150117035 Alperen KÖKER
- 150117037 Koray Emre ŞENEL
- 150118504 Mehmet Etka UZUN
- 150119529 Esraa Ismail Taha Turky.

Glossary of Terms:

- Labeling mechanisms : models process the date and assign instances to suitable labels
- Random labeling   : model assign random label form the list of labels to an instance
- Multi-labeling : proved more than one label to one instance

System Sequence Diagram: