Project Data Labeling

Requirement Analysis Document

This Project is dedicated to build a data labeling software .Data Labeling can be defined as a process of describing and associating some texts (instances) with single or couple of words(labels) so that they can be classified and analyzed to use.

Functional Requirements

System can read a digital file entered and can parse the information included in it.

System is supposed to allow users to associate instances with labels which are entered to system in a digital file.

System can make a random selection from labels included in dataset and associate selected label with an instance.

System can produce an output file including labeled texts and their labels

Non-Functional Requirements

System must be user friendly, easy to use and secure.

Glossary

Instance:

An instance represents a piece of text which can consist of eather a single word or couple of sentences.Instances take place in a dataset which is entered to the system as a file. Instances are labeled by Users.

Label:

A label is one or couple of words that will describe an instance or it's feature briefly in a point of view.One or more labels can be associated with an instance.

Dataset:

A dataset is a group of information including instances and labels can be used for these instances.A dataset also includes information of maximum number of labels that can be associated with a single instance. Datasets are entered to system in digital files.

Label Assignment:

Label Assignment represents an associating activity of one or more labels with an instance.

Maximum Number Of Labels:

Maximum number of labels is the answer to the question such that how many labels can be associated with a single instance at maximum.

User:

A user is a person or a mechanism that will associate labels and instances in the dataset which is entered to the system.

Domain Model