**SRS And UML documentation**

**SRS:**

(Software Requirement Specification)

Requirement specifications for our project --- Evaluating Turkiye Students dataset by using kmeans and Hierarchial clustering algorithms are:

**Software requirements:**

* Operating System: Windows 7, Windows 8, (or higher versions)
* Language: Python 3.5 and other libraries likes numpy, pandas, matplotlib, seaborn and scikit-learn.
* Browsers: Mozilla Firefox, Chrome (or any browser)

**Hardware requirements:**

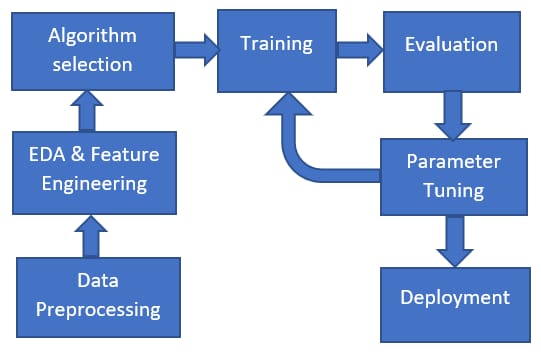
* Processor: Pentium 3, Pentium 4 and higher
* RAM: 2GB/4GB RAM and higher
* Hard disk: 40GB and higher

**UML Diagrams**

(Unified Modeling Language-UML)

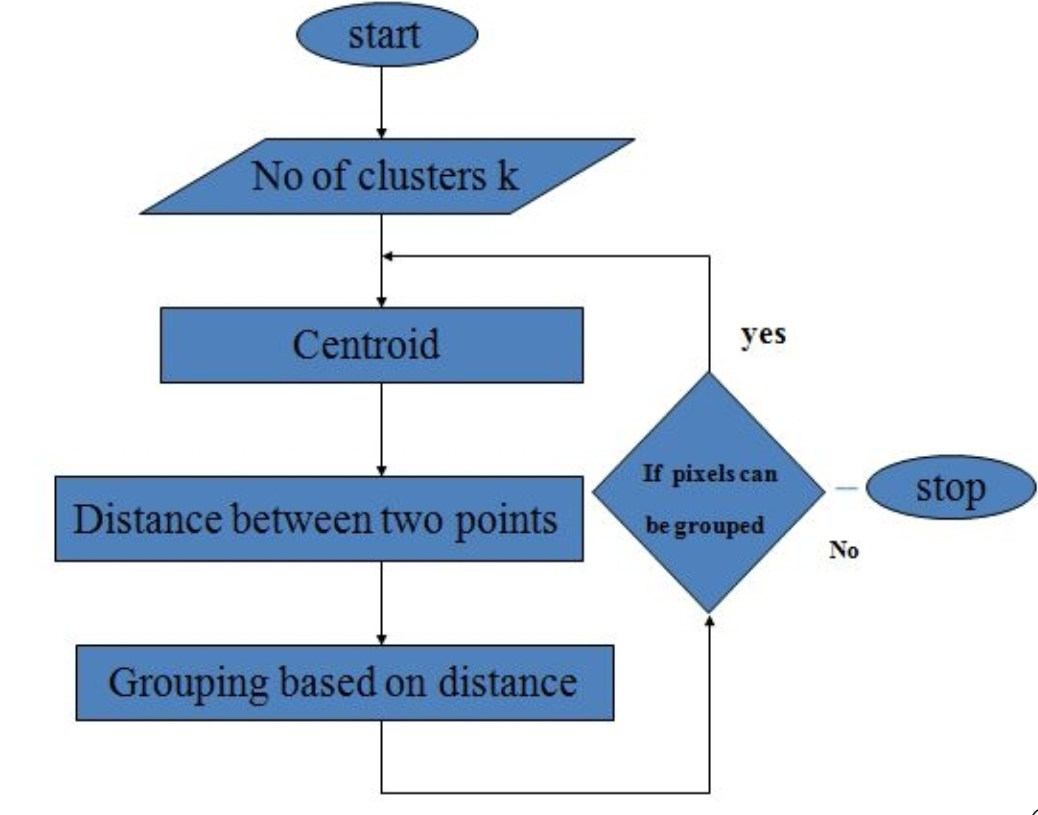
**Data preprocessing:**

Data preprocessing steps helps us to enhance the quality of data by transforming raw data into understandable and readable format .In machine learning it refers to the technique of preparing the raw data to make it suitable for building and training machine learning models.

****

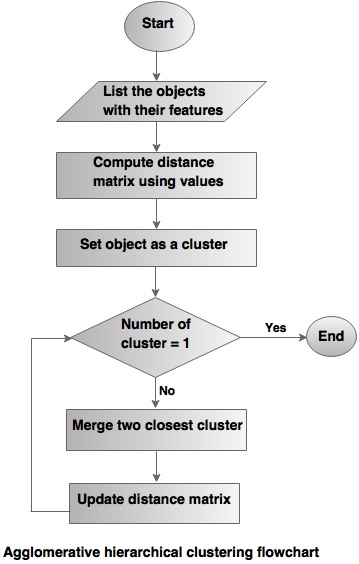
**Kmeans :**

Kmeans algorithm is one of the unsupervised machine learning algorithm used to partition the dataset into kpre-defined distinct non-overlapping subgroups (clusters) where each datapoint belongs to only one group. The less variation we have with in the clusters , the homogeneous the datapoints are with in the same cluster .



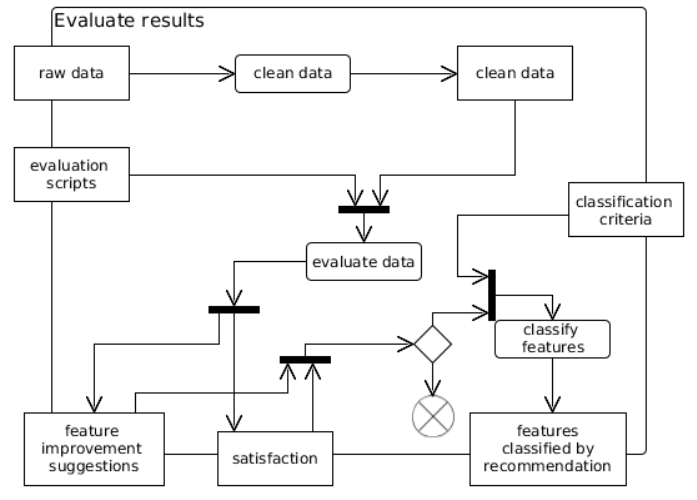
**Hierarchical Agglomerative Clustering :**

The agglomerative clustering is the most common type of hierarchical clustering used to group objects in clusters based on their similarity. The algorithm starts by treating each object as a singleton cluster.

****

**Process of trained model :**

Initially the data get cleaned after that it evaluate the data by finding the clusters then for later it applies classification algorithm to classify the data

****

clustering