**Project Report Vivek Bothra--1001354640**

In this assignment I have implemented the KNN algorithm for finding the k nearest neighbors of the testing example from the training dataset.

I have implemented the Euclidean distance and the polynomial kernel for finding the nearest neighbors.

The datasets I have used are Ecoli, glass and yeast.

During preprocessing the dataset I have performed the normalization and K-Fold cross validation .

The Initialize functions are used to initialize all the datasets reading raw input data and adding it to various data structures to perform KNN algorithm upon them .

Kfold cross validation is performed adding the training datasets and testing data set to the superset list.

I have made classes for individual attributes containing name , value and class name and id for the attribute and separate class for individual instance called exampleSet containing the list of attributes.

I have added the columns to a separate list so as to perform normalization on the data easily.

The process starts by running the test in which various functions are called to initialize the dataset, perform normalization doing k fold cross validation and finally starting the classification to use various distance metrices to find k nearest neighbors of the testing instance in the training dataset.

Here I have calculated the k nearest neighbors for the testing index instance in the training set and shown their respective classes.

The results are shown as below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Distance metric | dataset | Value of k | Training set | Testing data index and class | Training set index and class |
| euclidean | ecoli | 1 | Trset-1 | 69 -- cp | 82--cp |
| euclidean | ecoli | 2 | Trset1 | 42 --cp | 326-pp, 136--cp |
| euclidean | ecoli | 5 | Trset1 | 149 --im | 203—im, 151—im,196—im,249—imu,66--cp |
| polynomial | ecoli | 1 | Trset1 | 24--cp | 131--cp |
| polynomial | ecoli | 2 | Trset1 | 243--imu | 15—cp, 138--cp |
| polynomial | ecoli | 5 | Trset1 | 206--im | 124—cp,24—cp,129—cp,49—cp,76--cp |
| euclidean | glass | 1 | Trset1 | 159--vehiclewindowfloatprocessed | 46-buildingwindowfloatprocessed |
| euclidean | glass | 2 | Trset1 | 197--headlamps | 214—headlamps,212--headlamps |
| euclidean | glass | 5 | Trset1 | 131--buiwinfloatprocessed | 171—containers,200—headlamps,161—vehiclewindowfloatprocessed,157—vehiclewindowfloatprocessed,46—buildingwindowfloatprocessed. |
| polynomial | glass | 1 | Trset1 | 29--buildingwindowfloatprocessed | 208--headlamps |
| polynomial | glass | 2 | Trset1 | 196--headlamps | 162—vwfp,44--bwfp |
| polynomial | glass | 5 | Trset1 | 93--bwnfp | 109—bwnfp,204—headlamps,170—containers,108—bwnfp,183--tableware |