

School of Computer Science Engineering and Technology

Course- BTech
Course Code- 301
Year- 2022
Date- 18-04-2022

Type- Core
Course Name-AIML
Semester- Even
Batch- 4th Sem (SPL)

Lab Assignment No. 12

Objective: To Implement KNN Classifier

Problem Statement: Build a KNN Classifier using Sklearn library for identifying the types of glass (Multi-class Classification).

About Dataset: This dataset summarizes a types of glass defined in terms of their oxide content (5 M).

Data Set Characteristics:	Multivariate	Number of Instances:	214	Area:	Physical
Attribute Characteristics:	Real	Number of Attributes:	10	Date Donated	1987-09-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	426980

This is a copy of the test set of the UCI ML glass datasets

<https://archive.ics.uci.edu/ml/datasets/glass+identification>

From USA Forensic Science Service; 6 types of glass; defined in terms of their oxide content (i.e. Na, Fe, K, etc). The study of classification of types of glass was motivated by criminological investigation. At the scene of the crime, the glass left can be used as evidence...if it is correctly identified!

Attribute Information:

1. Id number: 1 to 214
2. RI: refractive index
3. Na: Sodium (unit measurement: weight percent in corresponding oxide, as are attributes 4-10)
4. Mg: Magnesium
5. Al: Aluminum
6. Si: Silicon
7. K: Potassium
8. Ca: Calcium
9. Ba: Barium
10. Fe: Iron
11. Type of glass: (class attribute)
 - 1 building_windows_float_processed
 - 2 building_windows_non_float_processed
 - 3 vehicle_windows_float_processed
 - 4 vehicle_windows_non_float_processed (none in this database)
 - 5 containers
 - 6 tableware
 - 7 headlamps
 - .

School of Computer Science Engineering and Technology

Steps

1. **Dataset:** Download the dataset from the link (5)
<https://archive.ics.uci.edu/ml/datasets/glass+identification>
2. Read the dataset (3)
3. Check the shape of data (2)
4. Extract the Independent and Dependent Variable (5)
5. Split the dataset into 80% for training and rest 20% for testing (sklearn.model_selection.train_test_split function) (10)
6. Perform the feature scaling (Hint: StandardScaler and MinMaxScaler) (10)
7. Build a KNN model using Sklearn with default parameters. (10)
8. Predict the target values in the testing set. (5)
9. Apply classification metrics and visualize the results as graphs. (10)
10. Playing with KNN: Change the following parameters of the KNN and analyze their performance for training and testing using the evaluation measures. (20)
 - a) n_neighbors
 - b) weights
 - c) algorithm
 - d) leaf_size
 - e) p
 - f) metric
 - g) metric_params
 - h) n_jobs
11. Compare the performance of the KNN model with other classification models (20).

Suggested Platform: Python: Azure Notebook/Google Colab Notebook, packages such as numpy, pandas, sklearn, matplotlib.