Assignment 3

CS4104 Applied Machine Learning

Deadline: 11th December 2021

Question 1

You are provided with the dataset in Table 1:

1. Load the dataset into the Pandas DataFrame.
2. Split the dataset into training and validation data randomly with 8 instances as training and 4 as validation.
3. Apply Linear SVM classifier with the regularization parameter values of [0.1 to 1.0] and display the results as SVM1 to SVM10.
4. Apply the weighted majority voting on the results of SVM1 to SVM10 with the weights of [0.1 to 1.0].
5. Apply the hard majority voting on the same results as in step 3 (SVM1 to SVM10).

Table 1:

Student Id	Marks >50	Attendance >50	Subjects >20	Grade(S/U)
P181234	1	1	1	S
P180001	1	1	0	S
P180002	0	1	1	U
P180003	1	0	0	S
P180004	0	1	1	U
P180005	1	0	1	S
P180006	1	1	1	S
P180007	0	0	0	U
P180007	0	0	1	U
P180008	1	1	1	U
P180009	0	1	1	S
P180010	1	0	0	S
P180011	0	1	1	S

Question 2

You are provided with the dataset in Table 1, apply the decision tree classification algorithm using 70% as training and 30% as validation dataset. Compute the evaluation measures of listed below:

1. Precision
2. Recall
3. Accuracy
4. F2-Score
5. MCC
6. Jaccard Index

Question 3

Using the dataset in the Table 1, convert it into three clusters using K-means clustering (Code). Show the clusters computed by the K-means. Apply the DBSCAN clustering on the same dataset on paper (not coding). Show the clusters computed by DBSCAN.