

Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham – Coimbatore

III Year B.Tech. CSE V Sem
19CSE305 – Machine Learning
Lab Evaluation – II

Date & Time: 30th Dec 2022 & 02:30 to 03:45 pm

Instructions:

- The total mark for the evaluation is 25, of which 5 marks are for data preprocessing + 15 marks are for building of models + 5 marks are for the comparison of models.
 - The dataset to be used for this question set is “seeds_dataset.txt”. A txt file by name “readme” is also shared to let you know the details about the dataset.
 - The questions should be neatly worked out in Google Colab/Jupyter and it needs to be made sure that the python notebook (all models should be created in one single ipynb) is named as Roll_Number_Eval1.ipynb (e.g., CB.EN.U8CSE80563_Eval1.ipynb).
 - The original notebook along with its pdf exported version should be submitted as solution to the assignment that will be created on the day of evaluation.
-

1. Data Preprocessing (5 marks)

- a. Thoroughly examine the dataset in order to determine whether there is a need for data preprocessing or not.
- b. Missing values, if any, may be handled properly by using the statistical properties such as mean/median/mode.
- c. Outliers and duplicate observations, if any, need to be resolved.
- d. Feature scaling, if the feature values are not in similar unit/range.

2. Model Building (15 marks)

- a. Four machine learning classification models have to be created in the following manner:
 - i. Model 1 – Logistic Regression
 - ii. Model 2 – Decision Tree (ID3)
 - iii. Model 3 – Random Forest
 - iv. Model 4 – K Nearest Neighbor
- b. The models can be created/trained/tested using the built-in utilities available in Scikit package.
- c. Use one single ipynb for all these models.

3. Model Comparison (5 marks)

- a. The classification reports of the models created in the previous step have to be undergoing a comparative study.
 - b. Your inference about the result of the comparison need to be included in the ipynb as well as presented to the examiner who conducts viva for you.
-