CDS503: Machine Learning
Academic Session: Semester 1, 2022/2023

School of Computer Sciences, USM, Penang

LAB EXERCISE (LAB 4)

Lab 4 Exercise

Re-look at the TAE data (tae.csv) from Lab 2 exercise.

Question 1

Explore the data and answer the following:
- What **type** of attributes of the data?
- Explore the data to check, whether is it **balance**?
- What attribute is **useful** and what is **not**? **why**?

Question 2

Experiment with **Naive Bayes** classifier introduced in this lab to *predict* the class of the tae data (tae.csv). Use *default* parameter configurations. Do data split and cross-validation as necessary. Compare the results using Naïve Bayes with the KNN.

Post your solution on Lab 04 Submission on **elearn@usm**. Make sure to include your name and lab# on the submission post.

Format: in .ipynb

The due date is **24 November 2022 5.00pm.**