



SCHOOL OF COMPUTER TECHNOLOGY

AASD 4001 Mathematical Concepts for Machine Learning

Term Project:

1. Provide a description of the dataset used including explanation of various features.
2. Pre-process and clean the dataset as appropriate.
3. Use following approaches for classification of the dataset:
 - A. Logistic Regression
 - B. Decision Tree
 - C. Random Forest
 - D. SGD
 - E. SVM
4. Use GridSearchCV to tune the parameter of each of the above models. Can you obtain better results in this step for any of the models? Discuss your observations.
5. Randomly (or based on certain hypothesis) remove some features and re-evaluate the models. Document your observations with respect to models performances.