

# CS4101 Applied Machine Learning

## Assignment 02

Due Date: 24th October 2023

Total Marks: 40

---

1. Use Table 1 to create Decision Tree using C4.5 and CART. Use Depth = 2. Show all steps clearly.  
**[20 Points]**
2. Use any datasets from Kaggle competition not earlier than 2020, compare and analyze the performance of decision trees on that dataset (with different hyper-parameter tuning) using F-measure and Accuracy. Plot performance curves and discuss. Use 10 Fold Cross Validation and random train/test split (70%, 30%).  
**[20 Points]**

Refund	Status	Tax Income	Cheat
Yes	Single	115	No
Yes	Single	60	No
Yes	Single	40	Yes
Yes	Married	130	Yes
Yes	Married	80	No
Yes	Divorced	130	Yes
Yes	Divorced	80	No
No	Single	75	No
No	Single	65	Yes
No	Married	125	Yes
No	Married	80	No
No	Divorced	90	Yes
No	Divorced	75	Yes

Table 1: Training Data.