

DS502 Machine Learning for Data Science

Assignment 01

Due Date: 28th Feb 2020 (In Class)

Total Marks: 60

- What is the difference between labelled and unlabelled data? (Write in your own words) [**5 Points**]
- Calculate Accuracy, and F-Measure using 3-Fold CV for data provided in Table 1 using (i) Naive Bayes (ii) kNN Classifier (iii) C4.5 (upto 1 level only) (iv) CART (upto 1 level only) [**40 Points**]
- Write down a one page summary of the attached paper. 12 points (Times New Roman) of the following paper. What is new in this paper about nearest neighbour [**15 Points**]

Refund	Status	Tax Income	Cheat
Yes	Single	125	No
Yes	Single	80	Yes
No	Single	75	No
No	Single	65	Yes
Yes	Single	60	No
No	Married	120	Yes
No	Married	80	No
Yes	Married	90	Yes
No	Single	95	Yes

Table 1: Data.