

Assignment#3(Classification Algorithms)

15 Points

Due Date: 24 Feb 2023

Instructions:

Answer all questions.

1. Model Building:

Build a **rbf kernel SVM** for the **Social_Network_Ads.csv** dataset and perform the following: **(6 point)**

- I. Prediction on a new test point [30, 87000]
- II. Print the confusion matrix.
- III. Visualizing the decision boundary and training set results
- IV. Visualizing the decision boundary and test set results

2. Model Building:

For instructions please refer Assignment#3_Q#2.ipynb **(6 point)**

3. You are given a dice. If you get a six, then you get \$12. But every time you need to play you need to give \$2. Will you play this game? Why or why not? **(1 point)**

4. Suppose you are building a decision tree model, which split a node on the attribute, that has highest information gain with **Gini**. Consider the training examples shown in Table below for a binary classification problem.

Which variable among “Student Background”, “Working Status”, and “online courses” will you consider for the root node of the decision tree based on Gini score. This question is to be answered without programing in Python. **(2 points)**

Resp srl no	Target variable	Predictor variable	Predictor variable	Predictor variable
	Exam Result	Other online courses	Student backgroun d	Working Status
1	Pass	Y	Maths	NW
2	Fail	N	Maths	W
3	Fail	y	Maths	W
4	Pass	Y	CS	NW
5	Fail	N	Other	W
6	Fail	Y	Other	W
7	Pass	Y	Maths	NW
8	Pass	Y	CS	NW
9	Pass	n	Maths	W
10	Pass	n	CS	W
11	Pass	y	CS	W
12	Pass	n	Maths	NW
13	Fail	y	Other	W
14	Fail	n	Other	NW
15	Fail	n	Maths	W

Submission Format: In the DC Connect, post the following:

1. A word document with answers to question 3 and 4.
2. A ran jupyter notebook file for question 1 and 2. Use the markdown cell in jupyter notebook to explain your findings. Use the Code cell in jupyter notebook to write the python code.

[Any submission other than the format of a notebook file(.ipynb) will be graded to zero.]

Assignment Rubric:

	Exceeds Expectations (9-10 points)	Meets Expectations (7-9 points)	Approaches Expectations (5-7 points)	Fails to meet Expectations (0-5 points)
Assignment Criteria	Assignment guidelines have been followed completely.	Assignment guidelines have been followed but 1 or 2 items missing.	Assignment guidelines have been followed more than two items.	Assignment guidelines have not been followed.
Organization of submitted documents	Assignment is exceptionally well organized.	Assignment is thoughtfully organized.	Some order/organization to submission but still some areas are unclear.	Assignment is disorganized and hard to follow.
Communication and Presentation of submitted documents	Superior communication	Reasonably understandable	Understandable with minor effort	Difficult to understand
Approach	The approach is well defined, clearly explained in detail and well formatted	The approach is well-defined, high-level explanation is given and well formatted	The approach has some flows, high level explanation is given and not well formatted	The approach has many flows, no explanation is given and not well formatted
Code	Code is run successfully, well documented, and complete	Code is run successfully but no documentation provided	Code has some flows and no documentation	Code is missing

Academic Integrity and Late submission:

Assignments are due by the due date announced in class and posted on DC Connect. At his or her own discretion, and depending on the nature of the assignment, each professor will provide a facility for the submission of late assignments up to a maximum of 72 hours after the assignment due date. All allowed late submissions will be assessed a penalty of 25% of the total possible grade for the

assignment. Assignments should be submitted on time, on a regular basis, to enable you to stay on track within the class.

Any violation of academic integrity will not be accepted and will be given a grade of zero (0) and reported. Find more information on academic integrity here

<https://durhamcollege.ca/mydc/learning-resources/academic-integrity>