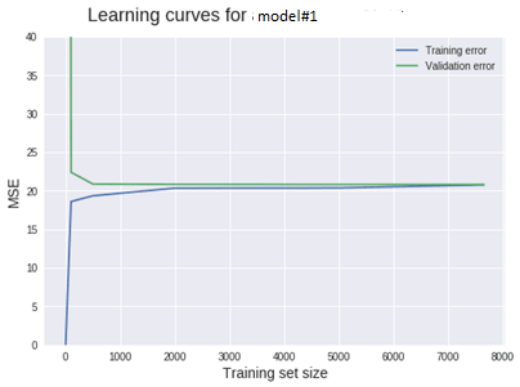
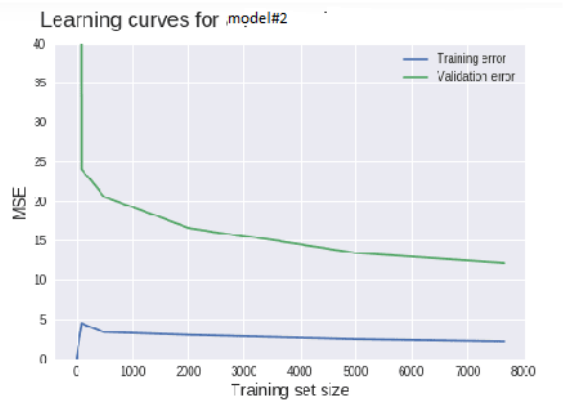
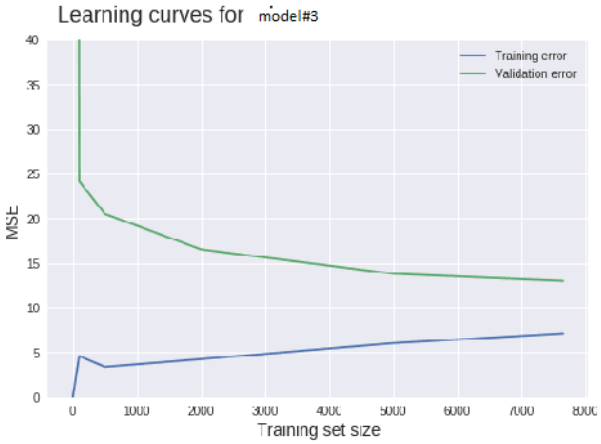
Assignment#2(Classification Algorithms)
15 Points
Due Date: 15 Feb 2023

Instructions:
Answer all four questions.

1. **Fitting problem:** Below are three learning curves plotted for three models (model#1, model#2, model#3). Which one of these are the best model? Does anyone of these suffer from fitting problem? If so, which one suffers from which fitting problem? What could be the causes and solutions to these fitting problems. ---- (3 point)







2. **Naïve Bayes algorithm:** Given is a dataset with three attributes: Color, Type and Origin. The target is Stolen and it can be either yes or no.

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Using Naïve Bayes algorithm determine if a **Red Domestic SUV** will be stolen or not without using Python Programming.

Note: Show all necessary steps. Answers without all the necessary steps will be graded as zero. ---- (4 point)

3. **Model Building:** Find the jupyter notebook and hypothyroid.csv file provided. ----(8 points)

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

1. A word document with answers to question 1 and 2.
2. A ran jupyter notebook file for question 3. 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