CA2: Machine Learning

This assignment is worth 40% of the total marks that are available for this module.

This is an individual assignment. (estimated workload (outside class time): 8 hours.)

Can we predict if a patient won't show (no show) for an appointment?

Requirements

You are required to take the dataset "Medical Appointment No Shows.csv" and **using either KNN or Naïve Bayes,** make a prediction if a patient will be a "no show or a show" at their medical appointment.

The output of your program must:

1. Visualize the data and prediction
2. Print out the message: "The patient will: " (show/No show):-**Dependant on your prediction**.
3. Compare Model accuracy (KNN VS Naïve Bayes)
4. Print out the accuracy or error of the ML model you are using.
5. Explain in the mark-up (As Text) why you picked the specific algorithm
6. Explain in the mark-up (As Text) your code

Marking Scheme (Sliding Scale)

Data	Feature Engineering	15
	Use of Appropriate Data Sample	5
	Clear Model Code	5
	Clear Training/ testing code	5

Modelling	Clear Explanation of Algorithm choice. This should be achieved through use of the notebook mark-up	20
Code Explanation	Clear, concise explanation of code and the process used to answer the question asked. This should be achieved through use of the notebook mark-up	15
Accuracy of Model	Compute the accuracy or error of the Model	10
	Compare Model accuracy (KNN VS Naïve Bayes)	10
Questions	Visualize the data and prediction/classification	10
	Print out the message: ""The patient will: " (show/No show):-**Dependant on your prediction.**	5
	TOTAL MARKS:	100

Please note: No other output will be graded

Deadline

This assignment is due at 23:55 on Sunday 18th April 2020. **Late submission will marked as per CCT policy. Failure to upload your assignment will result in a 0% grade for this assessment.**

Required Files

You are required to upload a Jupyter Notebook file onto Moodle, name it as yourName_Class_CA2.ipynb.