QUESTION 1

Question 1 : Linear Regression · · · · · · · · · · · · · · · **19.5** / 22 pts

1.1	Data properties	**3** / 3 pts
1.2	Linear Model	**3** / 3 pts
1.3	Display	**2.5** / 3 pts
1.4	Custom implementation	**3** / 3 pts
1.5	MSE	**3** / 3 pts
1.6	MSE 2	**2** / 3 pts
1.7	Analysis	**3** / 4 pts

QUESTION 2

Question 2 : Nonlinear Regression · · · · · · · · · · · **17** / 18 pts

2.1	Polynomial regression	**5** / 5 pts
2.2	Bar plot	**3** / 3 pts
2.3	Analysis	**3** / 4 pts
2.4	RBF	**6** / 6 pts

QUESTION 3

Question 3 : Decision Trees · · · · · · · · · · · · · · · · · · **23** / 26 pts

3.1	Dataset analysis	**2** / 4 pts
3.2	Analysis	**4** / 4 pts
3.3	Decision Trees	**2** / 2 pts
3.4	DT Depth	**2** / 3 pts
3.5	Hyperparameter tuning	**6** / 6 pts
3.6	Attribute importance	**5** / 5 pts
3.7	Analysis	**2** / 2 pts

QUESTION 4

Question 4 : Evaluating Binary Classifiers · · · · · · · **12** / 14 pts

4.1	Classification accuracy	**2** / 4 pts
4.2	AUC	**4** / 4 pts
4.3	ROC plots	**6** / 6 pts

MISTAKES

1.3:
– 0.5 pts You did not add a legend

1.6:
– 0.5 pts You could comment on why MSEs are the same/different
– 0.5 pts You included too many decimal places in your answer. Four or less is more appropriate

1.7:
– 1 pts You failed to mention that the plot was convex/that there is a single minimum rather than multiple minima.

2.3:
– 1 pts You did not mention that the weight vector entry corresponding to x_{4i} is very small.

3.1:
– 0.5 pts You did not report the correct number of positive labels in the train set (2335)
– 0.5 pts You did not report the correct number of negative labels in the train set (2465)
– 0.5 pts You did not report the correct number of positive labels in the test set (592)
– 0.5 pts You did not report the correct number of negative labels in the test set (608)

3.4:
– 1 pts You only gave one example of what happens when you use a maximum depth that is too large e.g. overfitting, large trees requiring larger storage, and can be slow to evaluate

4.1:
– 1 pts Your answer is too generic or you failed to mention a limitation of using a FIXED threshold i.e. the threshold might not be optimal for each of the different set of predictions
– 1 pts You failed to give a better way of choosing the threshold for each model e.g. using a held out validation set