

Rubric for Problem Set 3

1.1 (5 points)

Author clearly explains how they chose the variables in their candidate learner, as well as interactions (if any).

- 5 points. Full marks
- 4 points. Good. Could be clearer
- 2 points. Partially explained
- 0 points. No marks

1.2 (5 points)

Author clearly explains how they chose one or more candidate learners. From the writeup, I understand why the author chose the approach they did. (Full points possible even if only 1 approach was used)

- 5 points. Full marks
- 4 points. Good. Could be clearer
- 2 points. Partially explained
- 0 points. No marks

1.3A Code correctness (16 points)

Split the data into train and test correctly on approved specifications

- 4 points. Full marks
- 2 points. Small mistake
- 0 points. Incorrect or unclear

Trained the model on the training set correctly

- 4 points. Full marks
- 2 points. Small mistake
- 0 points. Incorrect or unclear

Tested the model on the training set and made predictions correctly

- 4 points. Full marks
- 2 points. Small mistake
- 0 points. Incorrect or unclear

Reported MSE correctly

- 4 points. Full marks
- 2 points. Small mistake
- 0 points. Incorrect or unclear

1.3B Code aesthetics (4 points). Code does not have to be formatted exactly as in the solutions, but it does need to be easily readable.

Every line of code fits on the PDF page

- 1 points. Yes!
- 0.5 points. Almost all
- 0 points. Many errors

Code is properly indented

- 1 points. Yes!
- 0.5 points. Almost all
- 0 points. Many errors

Spaces are used in a consistent way throughout the code

- 1 points. Yes!
- 0.5 points. Almost all
- 0 points. Many errors

Every pipe |> or + is followed by a new line

- 1 points. Yes!
- 0.5 points. Almost all
- 0 points. Many errors

2. New Task

2.1 (5 points)

- 5 points if population stated clearly, so that one could hypothetically make a list of these units
- 2.5 points if answer has some unclear reasoning or mistakes
- 0 points if no reasonable answer given, or the whole answer was incorrect, confusing, or off-track

2.2 (5 points)

- 5 points if outcome is precisely stated. Should be one variable that could be measured for each unit in the population
- 2.5 points if answer has some unclear reasoning or mistakes
- 0 points if no reasonable answer given, or the whole answer was incorrect, confusing, or off-track

2.3 (5 points)

- 5 points if predictors are precisely stated. Should be one or more variables that could be measured for each unit in the population
- 2.5 points if answer has some unclear reasoning or mistakes
- 0 points if no reasonable answer given, or the whole answer was incorrect, confusing, or off-track

2.4 (5 points)

- 5 points if effectively explains the significance of the predictions were both accurate and inaccurate
- 2.5 points if answer has some unclear reasoning or mistakes
- 0 points if no reasonable answer given, or the whole answer was incorrect, confusing, or off-track

Grad Questions (20 points)

Trained 3 learners in the train set

- 3 points. Full marks
- 2 points. Small mistake/some learners trained incorrectly
- 0 points. Incorrect or unclear

Predicted the test set with all 3 learners

- 5 points. Full marks
- 2.5 points. Small mistake
- 0 points. Incorrect or unclear

Correctly calculated MSEs on all 3 learners

- 5 points. Full marks
- 2.5 points. Small mistake
- 0 points. Incorrect or unclear

Interpreted findings/output as well as relating the output to machine learning's impact on predictability beyond what can be achieved by OLS

- 5 points. Full marks
- 2.5 points. Small mistake
- 0 points. Incorrect or unclear

Every line of code fits on the PDF page

- 0.5 points. Yes!
- 0.25 points. Almost all
- 0 points. Many errors

Code is properly indented

- 0.5 points. Yes!
- 0.25 points. Almost all
- 0 points. Many errors

Spaces are used in a consistent way throughout the code

- 0.5 points. Yes!
- 0.25 points. Almost all
- 0 points. Many errors

Every pipe |> or + is followed by a new line

- 0.5 points. Yes!
- 0.25 points. Almost all
- 0 points. Many errors