**Notes for Assignment\_2 – Gabriel Augusto Arsego**

Normalization

* Before going into the assignment’s questions, I normalized the data.

A screenshot of a computer program

AI-generated content may be incorrect.A screenshot of a computer program

AI-generated content may be incorrect.

Question 1

* For question 1, Customer\_1 was created as a “data.frame”. knn\_model1 was trained using the Training\_data that was partitioned earlier.
* Customer\_1 would be classified as 0, meaning that he wouldn’t accept the loan offer.

A screenshot of a computer program

AI-generated content may be incorrect.

Question 2

* A white background with blue text

  AI-generated content may be incorrect.Using the train() function I created knn\_model2 to find the best k and found k = 5.

Question 3

A screenshot of a computer

AI-generated content may be incorrect.A computer screen shot of text

AI-generated content may be incorrect.

Question 4

A close-up of a computer code

AI-generated content may be incorrect.

Same client as question 1, when applying the model with k = 5, we achieve the same result. A client with these characteristics wouldn’t accept the loan offer.

Question 5

* knn\_model3 created after portioning the data in 50, 30 and 20%. First dividing into 80% and 20%. Then for the Train and Validation a proportion of 62.5 and 37.5% respectively.

A computer screen shot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

* The differences weren’t great when it comes to accuracy, 95% for the Validation data, and 96.5% for the Test data. Other statistics like Kappa was significantly higher for the Test data, meaning that it had a better result when it comes to how more accurate the prediction was than a random guess.