Answer Sheet for Assignment: Decision Tree

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Fill in the answer sheet below.*

|  |  |  |
| --- | --- | --- |
|  | **Question** | **Answer** |
| **Part 1. Decision Tree in R** **(Complexity factor = 0.005)** |
| 1 | How often will this tree make a correct prediction (include decimals)? Provide your answer for both the training set and the validation set. |  |
| 2 | How likely is a customer to pay back their loan if they have one child and make $35,000 per year?*(NOTE: When asked “how likely…” cite the percentage!)* |  |
| 3 | How likely is a customer to pay back their loan if they are married, make $45,000 per year, have no children, and no mortgage? |  |
| 4 | How likely is a customer to pay back their loan if they make $83,000 per year and have no children? |  |
| 5 | Describe the profile of the least likely customer to successfully repay their loan. |  |
| 6 | Describe the profile of the most likely customer to successfully repay their loan. |  |
|  **(Complexity factor = 0.05)** |
| 7 | How often will this new tree make a correct prediction (include decimals)? Provide your answer for both the training set and the validation set. |  |
| 8 | Is this model better or worse than the first model at predicting who will repay their loan? Explain how changing the complexity factor affected the tree using **no more than two sentences.** |  |
| 9 | How likely is a customer to pay back their loan if they have one child and make $35,000 per year? |  |
| 10 | Does marriage increase or decrease the likelihood that a customer will pay back their loan? |  |
| **Part 2 Compute and Evaluate Decision Trees** |
| 11 | What is the correct classification rate for Tree #1? |   |
| 12 | What is the correct classification rate for Tree #2? |  |
| 13 | Which decision tree (Tree #1 versus Tree #2) has higher classification accuracy? |  |