1 point

1. What percentage of the predictions on sample_validation_data did decision_tree_model get correct?

- [ ] 25%
- [x] 50%
- [ ] 75%
- [ ] 100%

---

1 point

2. Which loan has the highest probability of being classified as a safe loan?

- [ ] First
- [ ] Second
- [ ] Third
- [x] Fourth

---

1 point

3. Notice that the probability preditions are the exact same for the 2nd and 3rd loans. Why would this happen?

- [x] During tree traversal both examples fall into the same leaf node.
- [ ] This can only happen with sheer coincidence.

---

1 point

4. Based on the visualized tree, what prediction would you make for this data point?

- [ ] +1
- [x] -1

---

1 point

5. What is the accuracy of decision_tree_model on the validation set, rounded to the nearest .01 (e.g. 0.76)?

0.64

---

1 point

6. How does the performance of big_model on the validation set compare to decision_tree_model on the validation set? Is this a sign of overfitting?

- [ ] big_model has higher accuracy on the validation set than decision_tree_model. This is overfitting.
- [ ] big_model has higher accuracy on the validation set than decision_tree_model. This is not overfitting.
- [x] big_model has lower accuracy on the validation set than decision_tree_model. This is overfitting.
- [ ] big_model has lower accuracy on the validation set than decision_tree_model. This is not overfitting.

---

1 point

7. Let us assume that each mistake costs money:

- Assume a cost of $10,000 per false negative.
- Assume a cost of $20,000 per false positive.

What is the total cost of mistakes made by decision_tree_model on validation_data? Please enter your answer as a plain integer, without the dollar sign or the comma separator, e.g. 3002000.

50280000