**2. What percentage of the predictions on sample\_validation\_data did decision\_tree\_model get correct?**

25%

**50%**

75%

100%

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

First

Second

Third

**Fourth**

**4. Notice that the probability preditions are the exact same for the 2nd and 3rd loans i.e 0.472267584643798. Why would this happen?**

**During tree traversal both examples fall into the same leaf node.**

This can only happen with sheer coincidence?

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

+1

**-1**

**6. What is the accuracy of decision\_tree\_model on the validation set, rounded to the nearest .01 (e.g. 0.76)?**

0.64

**7. 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.

**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.

**8. 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