**Q1.1: Why does it make sense to discretize columns for prediction problem?**

It is convenient for us to discretize columns especially for this type of task that we are facing since having many values will make it more challenging in terms of measuring frequency. It would require a lot of unique records for each value. With this CA, it would make the analysis more challenging especially since the size is quite big. With the help of discretizing, the volume in front of us is reasonable and an overall effective analysis can be done.

**Q1.2: What might be the issues (if any) if we DID NOT discretize the columns?**

As briefly mentioned in the previous question, a lot more additional work would have to take place. As an analyst dealing with data, time-efficiency is very important. When we can reduce the values without affecting the overall result, it should be done. With the “Tree Decision Analysis” the number of leaves will be extremely excessive. Reducing the number of leaves will make our analysis simples, but also cleaner. We want to avoid to overfit our model.

**Q8.1: How long was your total run time to train the model?**

It took me roughly 2 hours to run the model as I was struggling to make the right code. Once I have the code down, I can imagine it would be reduced significantly in the future, even with bigger datasets.

**Q8.2: Did you find the BEST TREE?**

I ran various combinations, and I was able to find a tree with good accuracy, precision, and f1 score. I used information from “medium.com”. The link is attached in the notebook.

**Q8.3: Draw the Graph of the BEST TREE Using GraphViz:**

**Diagram

Description automatically generated**

**Q8.4: What makes it the best tree?**

It is the best tree because of the evaluation I conducted after I made it. The evaluation consists of accuracy, precision and f1 score. The AUC value was the highest with this tree as well. Overall, looked at all these measures, I concluded that this was the best tree.

**Q10.1 and Q10.2:**

I was not able to use the right codes to find the answers I was looking for.