1. Gini index is 0.64 and corresponding entropy is 0.4416
2. decision trees are prone to over fitting,

Random forests collects these decision trees to arrive at a solution.

Decision trees cannot guarantee optimal trees.

Random forests are robust to outliers while decision trees are not.

There is a lower risk of overfitting in decision trees.

1. Scaling is necessary as some features have g=higher range while some have lower range. In other case the difference is units is alos there.

So to bring down all the features at one level scaling is necessary.

* Min-max Scaler
* Power transformer

1. We can speed up gradient descent by scaling.

This is because θ will descend quickly on small ranges and slowly on large ranges, and so will oscillate inefficiently down to the optimum when the variables are very uneven

1. If the dataset is highly imbalanced accuracy cannot be used as a metric or accuracy cannot prove to be a good metric as the model will have a better idea(training will be better) of the greater occurring class and will tend to predict that class correctly.

On other hand even if the model predicts the lower occurring class incorrectly the accuracy will tend to be higher as the greater occurring class I predicted well.

1. F score is harmonic mean of precision and recall.

It is given by

F1=2[(precision\*recall)/(precision+recall)]

F1=[tp/(tp+1/2(fp+fn))]

1. fit() - It is used for calculating the initial filling of parameters on the training data (like mean of the column values) and saves them as an internal objects state

transform() - Use the above calculated values and return modified training data

fit\_transform() - It joins above two steps. Internally, it just calls first fit() and then transform() on the same data.