Name: Monu Tayal

Roll Number : 2018201042

Decision Tree

Decision tree classifier to predict which valuable employee will leave the company.

Q1.

```
tayal@tayal-Lenovo-G50-80:~/Downloads/a1_2018201042$ python3 q-1-1.py
TP = 2 TN = 6843 FP = 1 FN = 2144
Accuracy : 76.14 %
Precision : 66.67 %
Recall : 0.09 %
f1-score : 0.19 %
```

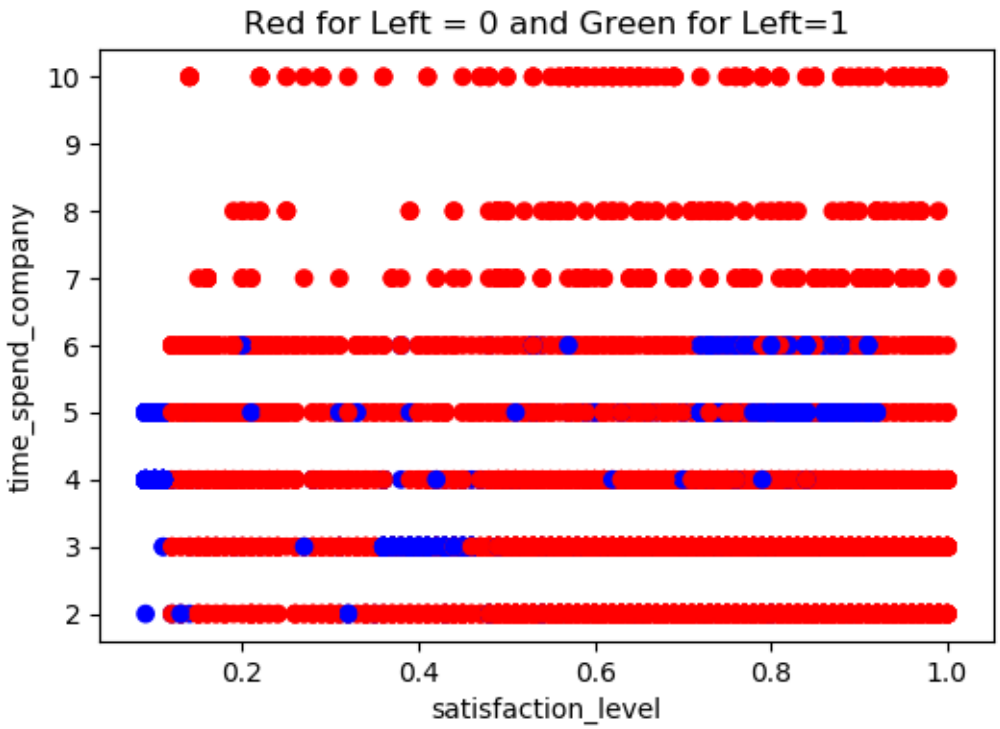
Q2.

```
tayal@tayal-Lenovo-G50-80:~/Downloads/a1_2018201042$ python3 q-1-2.py
TP = 1994 TN = 6589 FP = 265 FN = 142
Accuracy : 95.47 %
Precision : 88.27 %
Recall : 93.35 %
f1-score : 90.74 %
```

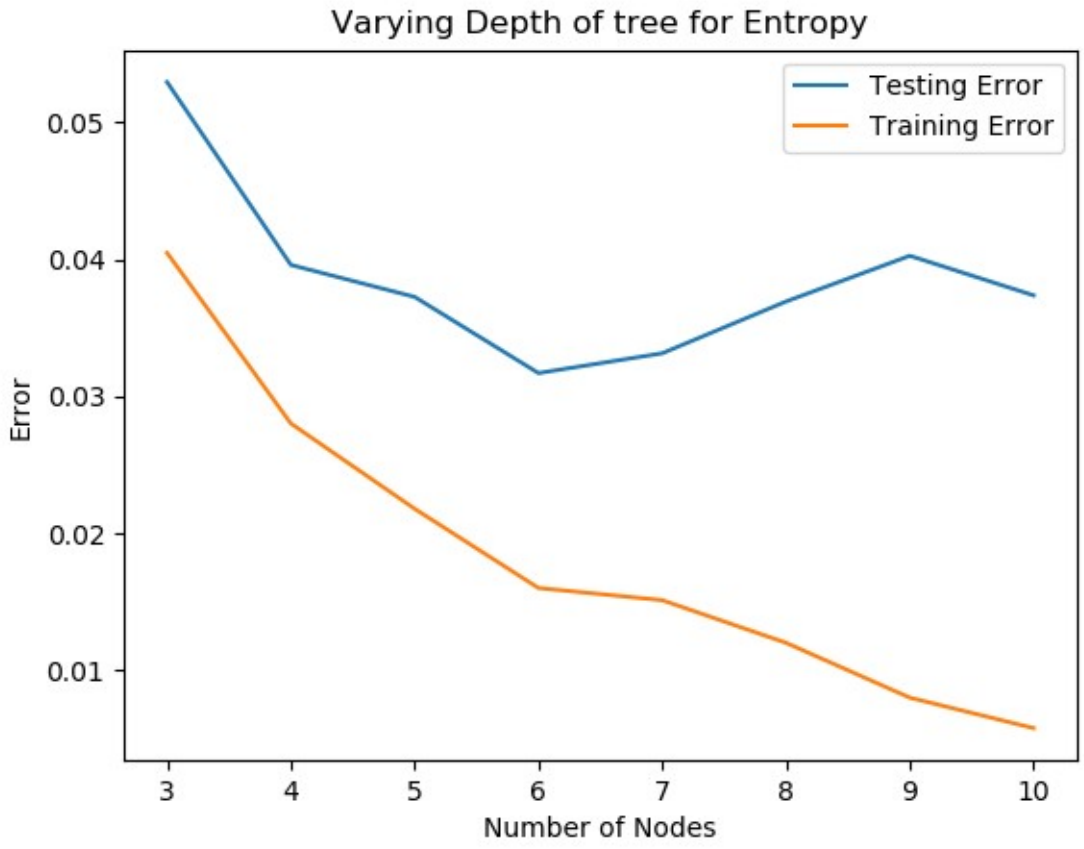
Q3.

```
tayal@tayal-Lenovo-G50-80:~/Downloads/a1_2018201042$ python3 q-1-3.py
TP = 1958 TN = 6656 FP = 206 FN = 170

For  MisClassification Rate
TP = 1958 TN = 6656 FP = 206 FN = 170
Accuracy : 95.82 %
Precision : 90.48 %
Recall : 92.01 %
f1-score : 91.24 %

TP = 1965 TN = 6664 FP = 198 FN = 163

For  Gini
TP = 1965 TN = 6664 FP = 198 FN = 163
Accuracy : 95.98 %
Precision : 90.85 %
Recall : 92.34 %
f1-score : 91.59 %

TP = 1960 TN = 6701 FP = 161 FN = 168

For  Entropy
TP = 1960 TN = 6701 FP = 161 FN = 168
Accuracy : 96.34 %
Precision : 92.41 %
Recall : 92.11 %
f1-score : 92.26 %
```

Q4.



Q5.



Q6.
Decission tree for handling missing values
Mean and Mode

*If data is from numerical column then we consider the average of entire column.
*If data is from categorical column then we consider the value those frequecy is maximum in that column.

Traverse through all branches

*While predicting for a certain row and at a particular node we encounter missing value then will will traverse through every branch of that node and consider the result which will have more occurance.