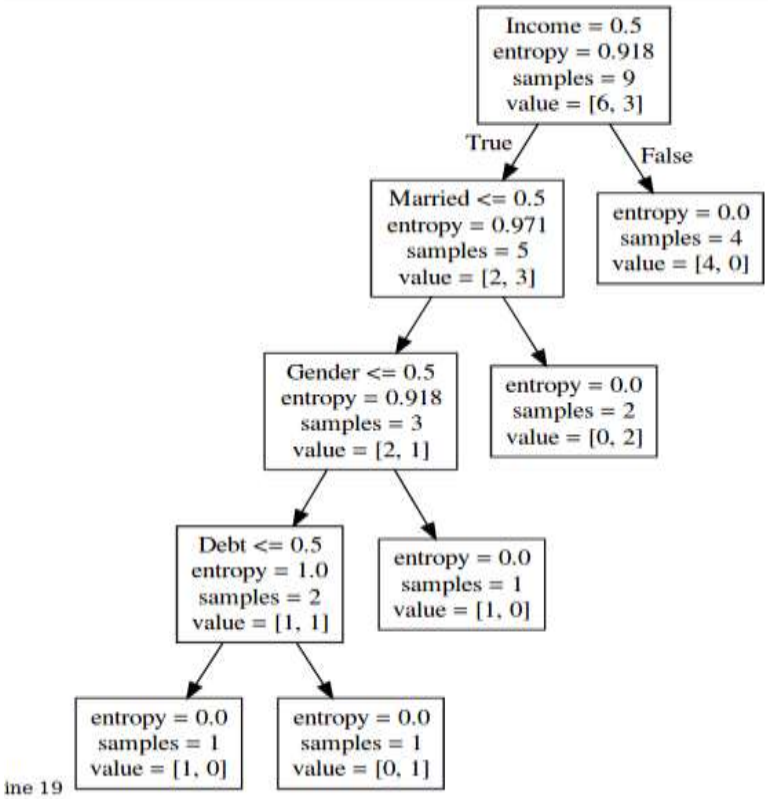
CS539: Machine Learning

HW1-Part2: Credit Risk Prediction

Task 2-1:



ine 19

Training Data:

Risk	Debt	Income	Married	Owns_Property	Gender
0	0	0	0	0	0
0	2	2	1	1	0
0	0	2	1	0	1
2	1	0	0	0	0
2	2	0	1	0	1
2	0	0	1	0	0
0	0	1	0	1	0
0	2	1	1	1	0
0	1	0	0	0	1

Legends:
Risk: 0=low, 2=high
Debt: 0=low, 1=medium, 2=high
Income: 0=low, 1=medium, 2=high
Married: 0=no, 1=yes
Owns_Property: 0=no, 1=yes
Gender: 0=male, 1=female

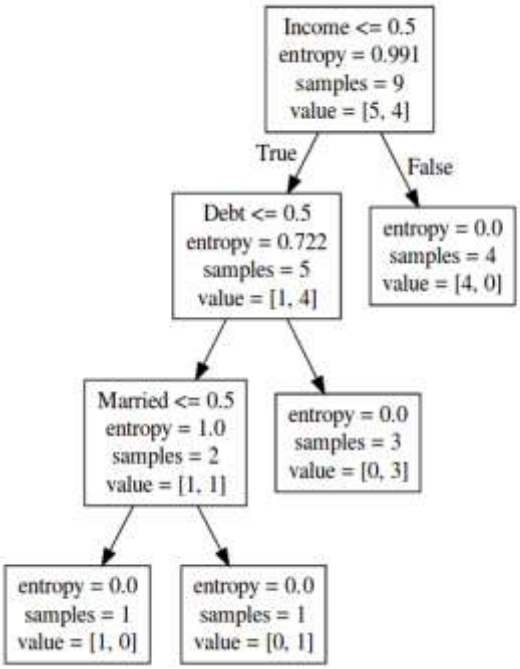
Since the Scikit learn algorithm required all the feature variables to be coded into its equivalent integer codes, all the string variables have been encoded to equivalent integer value, see legends on the right to understand.
All the data given in file credit.txt is used to train the model and following data is used as the test data to predict the output.

Debt	Income	Married	Owns_Property	Gender
0	0	0	1	0
0	1	1	1	1

```
Predicted Result: [0 0]
```

this means:
Credit Risk (Tom) = Low
Credit Risk (Ana) = Low

Task 2-2:



When Sofia's Credit Risk is changed to 'High' from 'Low', feature attribute 'Gender' no longer is a node of Decision Tree.
Also, in both of the decision tree, 'Owns_Property' plays no role in decision tree formation.