

Assignment -02 – Decision Tree

Part 2

Task 2-1. Decision Tree:

The decision tree was built using the ID3 algorithm implemented in Part 1. Below is the text representation of the decision tree with the predicted credit risk for Tom and Ana:

```
● myenvtkaur@MacBookPro Assignment2 % python3 part2.py
Decision Tree:
Married? = ?
| Married? = 0
|   Predict: low
| Married? = 1
|   Predict: low
| Married? = 2
|   Owns_Property = ?
|     | Owns_Property = 0
|     |   Predict: high
|     | Owns_Property = 1
|     |   Debt = ?
|     |     | Debt = 1
|     |     |   Predict: low
|     |     | Debt = 2
|     |     |   Predict: low
|     |
|     Predicted Credit Risk:
|     Tom: low
|     Ana: low
○ myenvtkaur@MacBookPro Assignment2 %
```

The decision tree was used to predict the credit risk for the following individuals:

Name	Debt	Income	Married?	Owns Property	Gender	Predicted Risk
Tom	Low	Low	No	Yes	Male	Low
Ana	Low	Medium	Yes	Yes	Female	Low

Task 2.2. After Changing Sofia's Credit Risk and seeing any unused feature

Before: Decision Tree and prediction with Sofia's credit risk as low

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS
● myenvtkaur@MacBookPro Assignment2 % python3 part2.py
Decision Tree:
Married? = ?
| Married? = 0
|   Predict: low
| Married? = 1
|   Predict: low
| Married? = 2
|   Owns_Property = ?
|   | Owns_Property = 0
|   |   Predict: high
|   | Owns_Property = 1
|   |   Debt = ?
|   |   | Debt = 1
|   |   |   Predict: low
|   |   | Debt = 2
|   |   |   Predict: low
|   |   Predict: low
|   Predict: high
Predicted Credit Risk:
Tom: low
Ana: low

Unused Features in Decision Tree: ['Income', 'Gender']
```

After: Decision tree after changing Sofia's credit risk from low to high

```
● myenvtkaur@MacBookPro Assignment2 % python3 part2.py
Decision Tree:
Married? = ?
| Married? = 0
|   Predict: low
| Married? = 1
|   Predict: low
| Married? = 2
|   Debt = ?
|   | Debt = 1
|   |   Owns_Property = ?
|   |   | Owns_Property = 0
|   |   |   Predict: high
|   |   | Owns_Property = 1
|   |   |   Predict: low
|   |   Predict: low
|   | Debt = 2
|   |   Predict: high
|   Predict: high
Predicted Credit Risk:
Tom: low
Ana: low

Unused Features in Decision Tree: ['Income', 'Gender']
```

In the original tree, the split after Married? = 2 is on Owns_Property

The tree first splits on Married?.

- If Married? = 0 or Married? = 1, the prediction is low.
- If Married? = 2, the tree further splits on Owns_Property.

- If Owns_Property = 0, the prediction is high.
- If Owns_Property = 1, the tree further splits on Debt.
 - If Debt = 1 or Debt = 2, the prediction is low

In the modified tree, the split after Married? = 2 is on Debt

The tree still first splits on Married?.

- If Married? = 0 or Married? = 1, the prediction is low.
- If Married? = 2, the tree now splits on Debt instead of Owns_Property.
 - If Debt = 1, the tree further splits on Owns_Property.
 - If Owns_Property = 0, the prediction is high.
 - If Owns_Property = 1, the prediction is low.
 - If Debt = 2, the prediction is high.

The modified tree introduces a new branch for Debt = 2 with a prediction of High

Unused Feature – in both trees, the features Income and Gender do not play a role in the decision tree. The Income feature does not provide significant information gain for splitting the dataset. As a result, it is not used in the decision tree. This could be because the distribution of Income values does not strongly correlate with the target variable (Risk).

The Gender feature also does not provide significant information gain for splitting the dataset. This could be because the distribution of Gender values is not strongly correlated with the target variable (Risk).