Adhiraj Budukh abudukh@wpi.edu

HW1 Part2

Library used: scikit-learn.

Criteria used: entropy.

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

Thus, given train data after encoding:

For part2.2 we are only changing value from ‘0’ to ‘2’.

Name	Debt	Income	Married	Owns_Property	Gender	Risk
Tim	0	0	0	0	0	0
Joe	2	2	1	1	0	0
Sue	0	2	1	0	1	0
John	1	0	0	0	0	2
Mary	2	0	1	0	1	2
Fred	0	0	1	0	0	2
Pete	0	1	0	1	0	0
Jacob	2	1	1	1	0	0
Sofia	1	0	0	0	1	2

Thus, Test data after encoding:

Name	Debt	Income	Married	Owns_Property	Gender
Tom	0	0	0	1	0
Ana	0	1	1	1	1

Output1: Decision tree when Sofia's risk is low/0:

```
Decision Tree Rules:
|--- Income <= 0.50
|   |--- Married <= 0.50
|   |   |--- Gender <= 0.50
|   |   |   |--- Debt <= 0.50
|   |   |   |   |--- class: 0
|   |   |   |--- Debt >  0.50
|   |   |   |   |--- class: 2
|   |   |--- Gender >  0.50
|   |   |   |--- class: 0
|   |--- Married >  0.50
|   |   |--- class: 2
|--- Income >  0.50
|   |--- class: 0
```

Prediction when Sofia's risk is low/0:

```
Tom's predicted credit risk: 0
Ana's predicted credit risk: 0
```

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

Output2: Decision tree when Sofia's risk is high/2:

As we can see Sofia's Credit Risk is changed to 'High' from 'Low', feature attribute 'Gender' is no longer exist in the Decision Tree.

```
Decision Tree Rules:
|--- Income <= 0.50
|   |--- Debt <= 0.50
|   |   |--- Married <= 0.50
|   |   |   |--- class: 0
|   |   |--- Married >  0.50
|   |   |   |--- class: 2
|   |--- Debt >  0.50
|   |   |--- class: 2
|--- Income >  0.50
|   |--- class: 0
```

Prediction when Sofia's risk is high/2:

```
Tom's predicted credit risk: 0
Ana's predicted credit risk: 0
```

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

Also, in both decision trees, 'Owns_Property' plays no role in decision tree formation.