Data Science HW5 Report

學號：b08502141　　姓名：石旻翰　　系級：電機四

1 Please use gini index to derive and draw the resulting decision tree

The python code is provided in the link in Reference[1]
Mapping: {
"Gender": {"M": 1, "F": 0},
"Car Type": {"Family": 0, "Sports": 1, "Luxury": 2},
"Shirt Size": { "Small": 0, "Medium": 1, "Large": 2, "Extra Large" : 3},
"Class": {"C0": 0, "C1": 1}
}

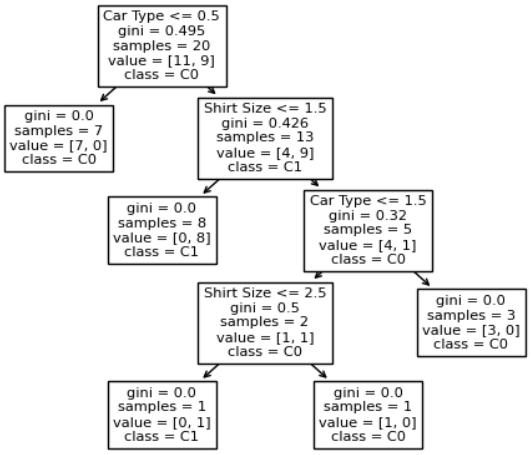

Figure 1: Decision Tree

2 Use naive Bayes classifier to classify input

Denote the input tuple (Gender=M, Car Type=Sports, Shirt Size=Medium)as X.
1. $P(X|C0) = P(\text{Gender=M} \mid \text{C0}) \cdot P(\text{Car Type=Sports} \mid \text{C0}) \cdot P(\text{Shirt Size=Medium} \mid \text{C0})$
$= \frac{7}{11} \cdot \frac{1}{11} \cdot \frac{2}{11} = \frac{14}{1331}$
2. $P(X|C0) = P(\text{Gender=M} \mid \text{C1}) \cdot P(\text{Car Type=Sports} \mid \text{C1}) \cdot P(\text{Shirt Size=Medium} \mid \text{C1})$
$= \frac{3}{9} \cdot \frac{5}{9} \cdot \frac{5}{9} = \frac{25}{81}$
And, $\frac{25}{81} > \frac{14}{1331}$. Thus, X should be classified as C1.

3 Drive the hyperplane by the SVM procedure.

The python code is provided in the link in Reference[2]

- Objective: Maximize $\frac{2}{|w^T|}$
- Constraints: $f(x) = \begin{cases} 1, \text{if } w^T \cdot X_i + b \geq 1 \\ -1, \text{if } w^T \cdot X_i + b \leq 1 \end{cases}$
- Two Support Vectors: [2, 1], [4, 3]
- The calculation steps:

```
X = np.array([[4, 3], [4, 8], [7, 2], [-1, -2], [-1, 3], [2, -1], [2, 1]])
Y = np.array([1, 1, 1, -1, -1, -1, -1])
clf.fit(X, Y)

w = clf.coef_[0]
a = -w[0] / w[1]
xx = np.linspace(-5, 5)
yy = a * xx - (clf.intercept_[0]) / w[1]

b = clf.support_vectors_[0]
yy_down = a * xx + (b[1] - a * b[0])
b = clf.support_vectors_[-1]
yy_up = a * xx + (b[1] - a * b[0])
```

Figure 2: Calculation steps

- (w, b) = ([0.5, 0.5], -2.5)
- Hyperplane

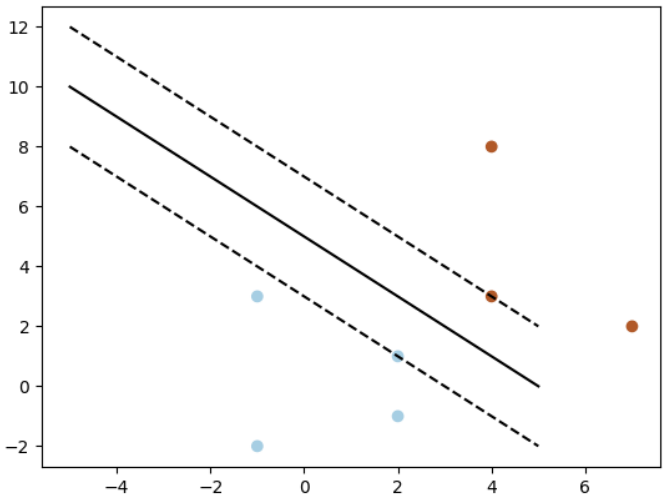

Figure 3: Hyperplane

References

[1] https://drive.google.com/file/d/1pKNbuopJVR3wIZ0s1DBVY0BQ3xNtAbhw/view?usp=share_link

[2] https://drive.google.com/file/d/10Vx8ZzsH-tyPbGNgI6x3Dw-eY3g3uR-F/view?usp=share_link