Machine Learning Homework 5

Hairui Yin

1.

(a) Consider using a SVM to find the margin perceptron between A and B, mapping brand 'A' to 1, brand 'B' to -1, we can construct the SVM given code

```
import pandas as pd
import numpy as np
from sklearn.svm import SVC

df = pd.read_csv('Pizza.csv')

dfAB = df[(df['brand']=='A') | (df['brand']=='B')]
X = dfAB.iloc[:, 2:].to_numpy()
Y = dfAB.iloc[:, 0]
Y = Y.map({'A':1, 'B':-1})

svm = SVC(kernel='linear', C=1.0, random_state=42)
```

The matrix w^T and intercept b are as follows:

```
w = svm.coef_.T
b = svm.intercept_
```

Using this method, we can get margin perceptron seperating AB(map A to 1, B to -1), seperating AC(map A to 1, C to -1) and seperating BC(map B to 1, C to -1), the result is as follows:

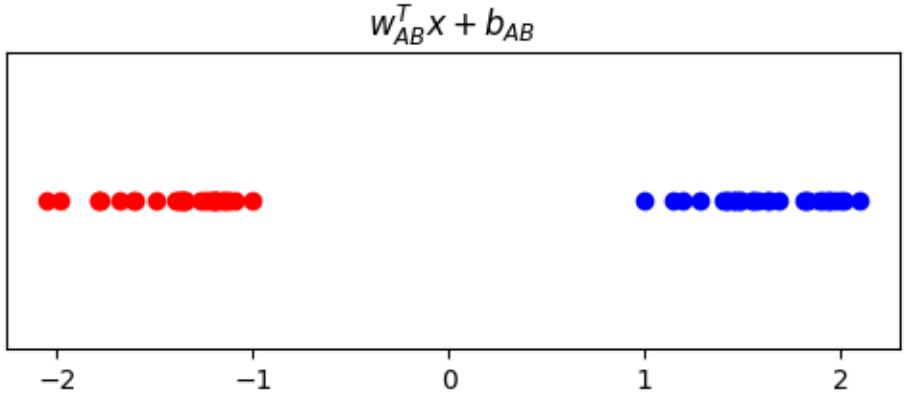
```
the matrix wT of hyperplace between A and B is
[[-0.08365963  0.0263273   0.06290964  0.00899566  0.00371821 -0.01457297
   0.00611705]]
the intercept b of hyperplace between A and B is
[0.81032823]
----------------------------------------------------------------------
the matrix wT of hyperplace between A and C is
[[-0.0611547  -0.01665783  0.0655635   0.00745803  0.00472285  0.00479101
   0.00544715]]
the intercept b of hyperplace between A and C is
[0.64815436]
----------------------------------------------------------------------
the matrix wT of hyperplace between B and C is
[[ 0.0720765  -0.15818992  0.04085258  0.00489641  0.00675163  0.04036443
  -0.00102341]]
the intercept b of hyperplace between B and C is
[-1.72875753]
```

To verify the result, we input data of brand A and B into the hyperplane of AB, then calculate

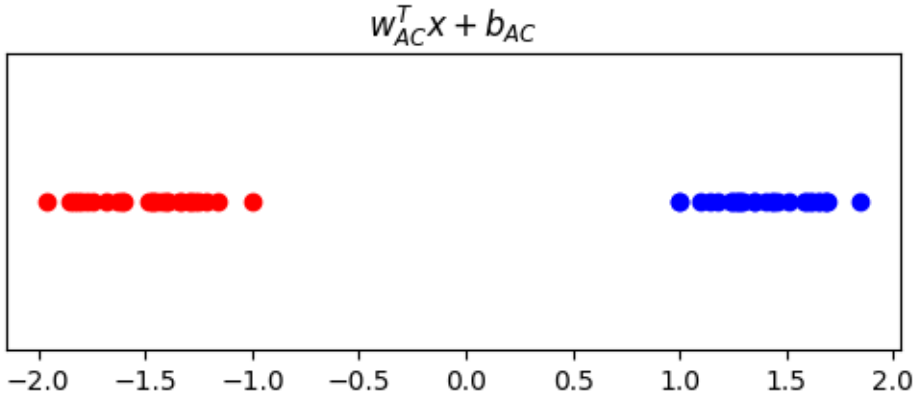
$$w^T x + b$$

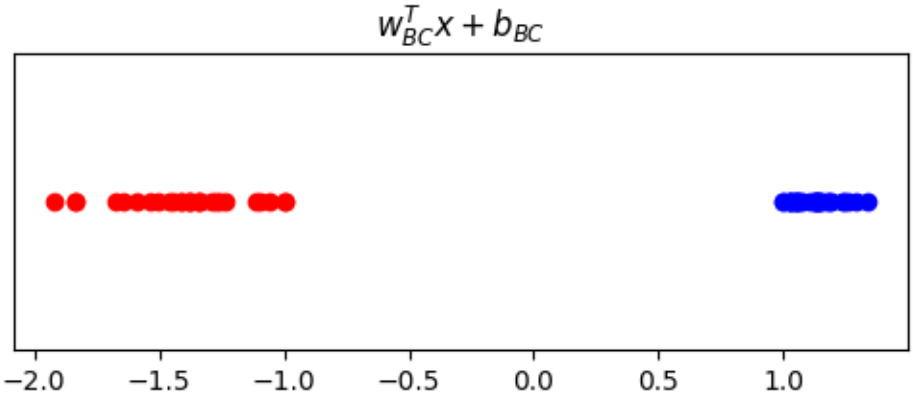
```
1        X @ wAB.T + bAB
```

The picture is shown below:



Similarly, the pictures using hyperplane AC, BC to classify the data are


$w_{AC}^T x + b_{AC}$
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0


$w_{BC}^T x + b_{BC}$
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0

(b) The margins can be calculated through

$$\text{margin} = \frac{2}{||w||}$$

```
the marginal of hyperplace between A and B is
18.261816747296752
-----------------------------------------------
the marginal of hyperplace between A and C is
21.761630456581393
-----------------------------------------------
the marginal of hyperplace between B and C is
10.912859531408142
```

(c) In this problem, we need three new hyperplane

$$\begin{cases} \text{A vs. BC} \\ \text{B vs. AC} \\ \text{C vs. AB} \end{cases}$$

Then, we assign x to label given by

$$y \in \arg \max_{j=A,B,C} w_j^T x + b_j$$

Firstly, finding matrix w_j and intercept b_j is similar to q(a). The results are

```
wA and bA are respectively
[[-0.08337843  0.02314028  0.06481282  0.00903022  0.00384769 -0.01360488
   0.00620176]]
[0.77596636]
-----------------------------------------------
wB and bB are respectively
[[ 1.14713811e-01 -1.66949615e-01  3.38913915e-02 -8.77824409e-03
   9.74604384e-05  2.71226561e-02 -2.67806239e-03]]
[-3.42281938]
-----------------------------------------------
wC and bC are respectively
[[-0.01674098  0.19812401 -0.09580719 -0.00745718 -0.01431877 -0.07811867
  -0.00392921]]
[-0.50011631]
```

After than, input $s1, s2, s3$ to $\arg\max_{j=A,B,C} w_j^T x + b_j$. These three instances are respectively labeled as

$$\begin{cases} \text{s1 : Brand C} \\ \text{s2 : Brand A} \\ \text{s3 : Brand B} \end{cases}$$