Name: Sai Anish Garapati
UIN: 650208577

Assignment 3

1) Question Answering:

Q1) Discriminative model: In discriminative model, we estimate the parameters of $P(C_k/X)$ directly from the training set.
Generative model: In generative model, we estimate the likelihood probability $P(X/C_k)$ and prior probability $P(C_k)$ from training set and use Bayes theorem to calculate $P(C_k/X)$.

Q2) The Naive Bayes algorithm assumes that the features are independent from each other given a class.

Q 3)

1) For class 1 (C_1)

Bag of words:

a	b	c	d
5	5	5	3

Learned parameters:

$P(C_1) = \frac{5}{10} = 0.5$

$P(a|C_1) = \frac{5}{18} = 0.28$

$P(b|C_1) = \frac{5}{18} = 0.28$

$P(c|C_1) = \frac{5}{18} = 0.28$

$P(d|C_1) = \frac{3}{18} = 0.17$

For class 2 (C_2)

Bag of words:

a	b	c	d
4	1	3	6

Learned parameters

$P(C_2) = \frac{5}{10} = 0.5$

$P(a|C_2) = \frac{4}{14} = 0.29$

$P(b|C_2) = \frac{1}{14} = 0.07$

$P(c|C_2) = \frac{3}{14} = 0.21$

$P(d|C_2) = \frac{6}{14} = 0.43$

(2) Prediction for "abbd"

$P(C_1 / abbd) \propto P(C_1) \cdot P(abbd / C_1) = P(C_1) \cdot P(a/C_1) \cdot (P(b/C_1))^2 \cdot P(d/C_1)$

$= (0.5) \cdot (0.28) \cdot (0.28)^2 \cdot (0.17)$

$= 0.0019$

$P(C_2/abbd) \propto P(C_2) \cdot P(abbd/C_2) = P(C_2) \cdot P(a/C_2) \cdot (P(b/C_2))^2 \cdot P(d/C_2)$

$= (0.5) \cdot (0.29) \cdot (0.07)^2 \cdot (0.43)$

$= 0.0003$

Prediction for "bbcc"

~~P(C1)~~ $P(C_1 / bbcc) \propto P(C_1) \cdot P(bbcc/C_1) = P(C_1) \cdot (P(b/C_1))^2 \cdot (P(c/C_1))^2$

$= (0.5) \cdot (0.28)^2 \cdot (0.28)^2$

$= 0.003$

$P(C_2 / bbcc) \propto P(C_2) \cdot P(bbcc/C_2) = P(C_2) \cdot (P(b/C_2))^2 \cdot (P(c/C_2))^2$

$= (0.5) \cdot (0.07)^2 \cdot (0.21)^2$

$= 0.0001$

(3) For "abbd", $P(C_1) > P(C_2)$. It belongs to Class 1

For "bbcc", $P(C_1) > P(C_2)$. It belongs to class 1

Q4)

Q 4) Let $W_{c_1} = [1, 1, 2]$, $W_{c_2} = [1, -1, 1]$, $W_{c_3} = [-1, 1, -1]$

x_1	x_2	a_{c_1}	a_{c_2}	a_{c_3}
1	1	4	1	-1
0	0	2	1	-1
-10	10	2	-19	19
100	-50	52	151	-151
-20	35	17	-54	54

where $a_{c_i} = W_{c_i}[2] + W_{c_i}[0]\, x_1 + W_{c_i}[1]\, x_2$

(1) For $x_1 = 1,\ x_2 = 1$:-

$$P(C_1 \mid \{x_1 = 1,\ x_2 = 1\}) = \frac{e^4}{e^4 + e^1 + e^{-1}} = 0.946$$

$$P(C_2 \mid \{x_1 = 1,\ x_2 = 1\}) = \frac{e^1}{e^4 + e^1 + e^{-1}} = 0.047$$

$$P(C_3 \mid \{x_1 = 1,\ x_2 = 1\}) = \frac{e^{-1}}{e^4 + e^1 + e^{-1}} = 0.006$$

For $x_1 = 0,\ x_2 = 0$:-

$$P(C_1 \mid \{x_1 = 0,\ x_2 = 0\}) = \frac{e^2}{e^2 + e^1 + e^{-1}} = 0.705$$

$$P(C_2 \mid \{x_1 = 0,\ x_2 = 0\}) = \frac{e^1}{e^2 + e^1 + e^{-1}} = 0.259$$

$$P(C_3 \mid \{x_1 = 0,\ x_2 = 0\}) = \frac{e^{-1}}{e^2 + e^1 + e^{-1}} = 0.035$$

For $x_1 = -10,\ x_2 = 10$:-

$$P(C_1 \mid \{x_1=-10,\ x_2=10\}) = \frac{e^{2}}{e^{2}+e^{-19}+e^{19}} = 4.13 \times 10^{-8}$$

$$P(C_2 \mid \{x_1=-10,\ x_2=10\}) = \frac{e^{-19}}{e^{2}+e^{-19}+e^{19}} = 3.13 \times 10^{-17}$$

$$P(C_3 \mid \{x_1=-10,\ x_2=10\}) = \frac{e^{19}}{e^{2}+e^{-19}+e^{19}} = 0.99$$

For $x_1 = 100,\ x_2 = -50$:-

$$P(C_1 \mid \{x_1=100,\ x_2=-50\}) = \frac{e^{52}}{e^{52}+e^{151}+e^{-151}} = 1.01 \times 10^{-43}$$

$$P(C_2 \mid \{x_1=100,\ x_2=-50\}) = \frac{e^{151}}{e^{52}+e^{151}+e^{-151}} = 1.0$$

$$P(C_3 \mid \{x_1=100,\ x_2=-50\}) = \frac{e^{-151}}{e^{52}+e^{151}+e^{-151}} = 6.96 \times 10^{-132}$$

For $x_1 = -20,\ x_2 = 35$:-

$$P(C_1 \mid \{x_1=-20,\ x_2=35\}) = \frac{e^{17}}{e^{17}+e^{-54}+e^{54}} = 8.53 \times 10^{-17}$$

$$P(C_2 \mid \{x_1=-20,\ x_2=35\}) = \frac{e^{-54}}{e^{17}+e^{-54}+e^{54}} = 1.24 \times 10^{-47}$$

$$P(C_3 \mid \{x_1=-20,\ x_2=35\}) = \frac{e^{54}}{e^{17}+e^{-54}+e^{54}} = 0.99$$

(3) For $x_1=1,\ x_2=1$; $P(C_1) > P(C_2) > P(C_3)$. Belongs to Class 1

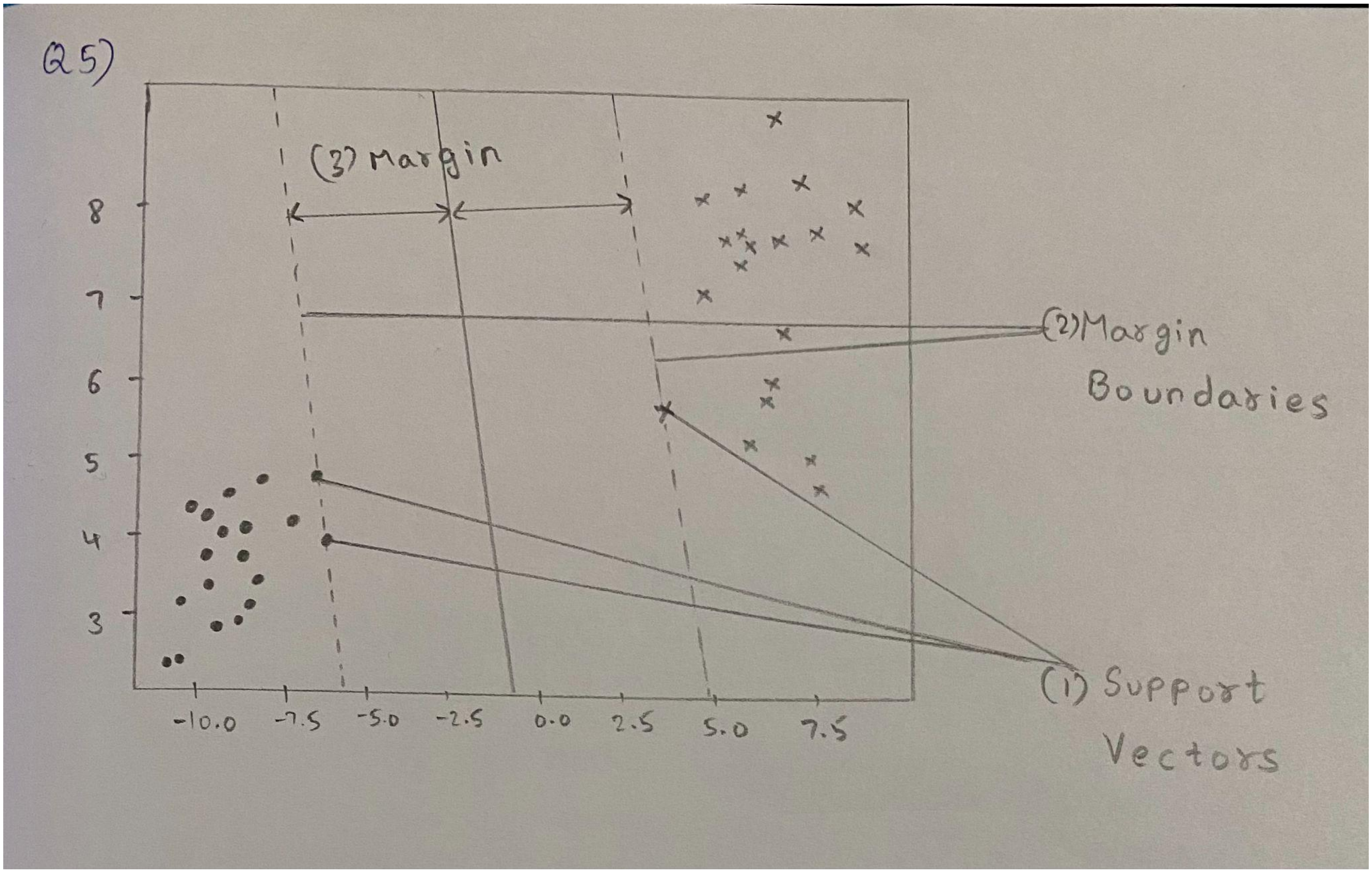
For $x_1=0,\ x_2=0$; $P(C_1) > P(C_2) > P(C_3)$. Belongs to Class 1

For $x_1=-10,\ x_2=10$; $P(C_3) > P(C_1) > P(C_2)$. Belongs to Class 3

For $x_1=100,\ x_2=-50$; $P(C_2) > P(C_1) > P(C_3)$. Belongs to Class 2

For $x_1=-20,\ x_2=35$; $P(C_3) > P(C_1) > P(C_2)$. Belongs to Class 3

Q5)



2) Programming

P1)

2.1) Predicted Class for [1.0 1.0]: Class 1

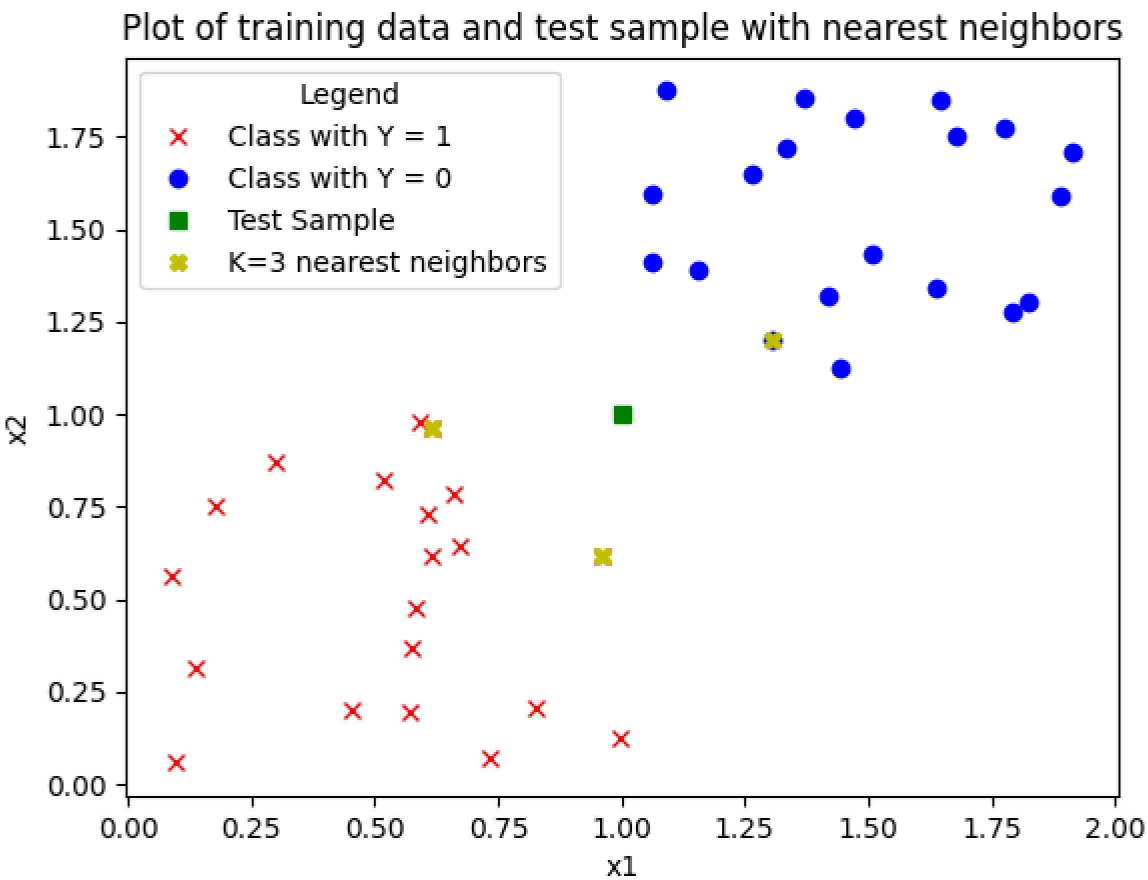
2.2) K=3 nearest neighbors and corresponding classes:

[1.30754347 1.20132266] - Class 0

[0.61652471 0.96384302] - Class 1

[0.96239599 0.61655744] - Class 1

3)



P2)

2) K: 1, Fold: 1, accuracy: 0.75
K: 1, Fold: 2, accuracy: 0.875
K: 1, Fold: 3, accuracy: 0.625
K: 1, Fold: 4, accuracy: 0.75
K: 1, Fold: 5, accuracy: 0.875
K: 1, Avg_accuracy: 0.775

K: 3, Fold: 1, accuracy: 0.875
K: 3, Fold: 2, accuracy: 1.0
K: 3, Fold: 3, accuracy: 0.75
K: 3, Fold: 4, accuracy: 0.75
K: 3, Fold: 5, accuracy: 0.875
K: 3, Avg_accuracy: 0.85

K: 5, Fold: 1, accuracy: 0.875
K: 5, Fold: 2, accuracy: 0.875
K: 5, Fold: 3, accuracy: 0.875
K: 5, Fold: 4, accuracy: 1.0
K: 5, Fold: 5, accuracy: 0.875
K: 5, Avg_accuracy: 0.9

K: 7, Fold: 1, accuracy: 0.875
K: 7, Fold: 2, accuracy: 0.875

K: 7, Fold: 3, accuracy: 0.75
K: 7, Fold: 4, accuracy: 1.0
K: 7, Fold: 5, accuracy: 0.875
K: 7, Avg_accuracy: 0.875

K: 9, Fold: 1, accuracy: 0.875
K: 9, Fold: 2, accuracy: 0.875
K: 9, Fold: 3, accuracy: 0.75
K: 9, Fold: 4, accuracy: 0.875
K: 9, Fold: 5, accuracy: 0.625
K: 9, Avg_accuracy: 0.8

Optimal K value: 5

P3)
2) Predictions using Logistic regression model:
Predicted class for [0.5, 0.5]: 1
Predicted class for [1, 1]: 0
Predicted class for [1.5, 1.5]: 0

3) Predictions using SVM model:
Predicted class for [0.5, 0.5]: 1
Predicted class for [1, 1]: 0
Predicted class for [1.5, 1.5]: 0

4)