CSE514, Fall 2021, HW 5 Name: Anh Le Student ID: 488493

Note: This homework is worth a total of 15 points

Q1 (8pts): You have a dataset of one-dimensional samples. Within that dataset:

Sample 1: x = [2], y = -1

Sample 2: x = [10], y = +1

To optimize an SVM, you need to calculate the dot product between all possible pairs of samples in the dataset.

Q1a (1pt): What's the dot product between Sample 1 and Sample 2?

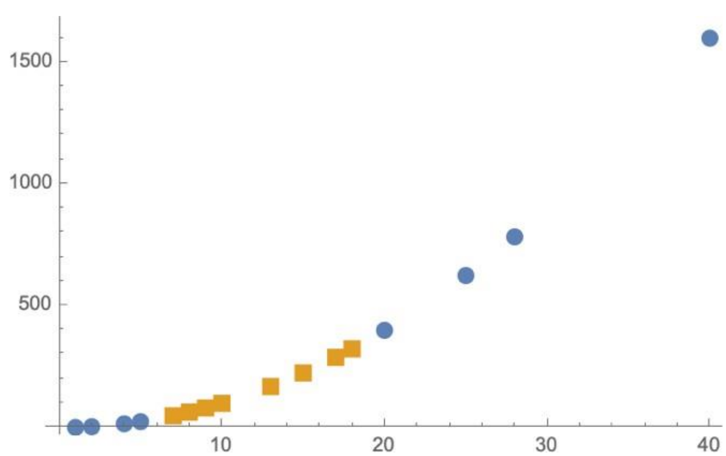
$2 \times 10 = 20$

You'd like to be able to separate the samples of the -1 class from the +1 class, but you can see that they aren't linearly separable:

So you decide to project the data into two dimensions. Updated samples:

Sample 1: x = [2, 4], y = -1

Sample 2: x = [10, 100], y = +1



Q1b (1pt): What's the dot product between Sample 1 and Sample 2 now?

$2 \times 10 + 4 \times 100 = 420$

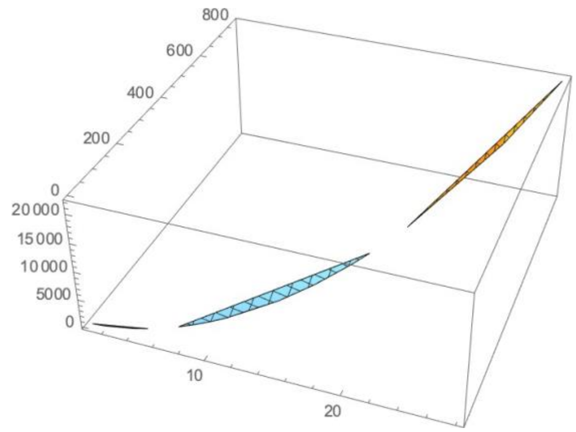
In two dimensions, these samples are separable, which is good, but the margins are tiny, which is not so good. Surely it would be better to find a boundary with larger margins? Let's try three dimensions:

Sample 1: x = [2, 4, 8], y = -1

Sample 2: x = [10, 100, 1000], y = +1

Q1c (1pt): What's the dot product between Sample 1 and Sample 2 now?

$$2 \times 10 + 4 \times 100 + 8 \times 1000 = 8420$$



Projecting onto three dimensions looks better than two, so maybe we should do SVM with even more?

Q1d (2pt): Project Samples 1 and 2 into five dimensions:

Sample 1: x = [2, 4, 8, 16, 32], y = -1
Sample 2: x = [10, 100, 1000, 10000, 100000], y = +1

Q1e (1pt): What's the dot product between Sample 1 and Sample 2 from your projection?

$$2 \times 10 + 4 \times 100 + 8 \times 1000 + 16 \times 10000 + 32 \times 100000 = 3368420$$

Q1f (1pt): Use the polynomial kernel with r = 1 to calculate the dot product between Sample 1 and Sample 2 in five dimensions.

$$\left(x_i \cdot x_j + 1\right)^5 = (2 \cdot 10 + 1)^5 = 4084101$$

Q1g (1pt): Use the RBF kernel with $\gamma = 1$ to calculate the dot product between Sample 1 and Sample 2 in infinite dimensions.

$$e^{-\gamma(a-b)^2} = e^{-1(a-b)^2} = e^{-1(2-10)^2} = e^{-64}$$

Q2 (7pts): *"Life is like a box of chocolates; you never know what you're going to get."* Unfortunately, I have a mild tree nut allergy, so if I'm going to pick out a chocolate from an unlabeled box, I'd prefer to pick one that's less likely to leave an itch in my throat!

From a labeled box, here are features of chocolates with or without tree nuts:

ID	Color	Shape	Appearance	Position	Flavor	Tree nuts?
1	Milk	Circle	Smooth	Edge	Coffee cream	No
2	Milk	Rectangle	Rippling	Edge	Toffee	No
3	Milk	Square	Dark drizzle	Edge	Caramel	No
4	Dark	Rectangle	Milk drizzle	Edge	Molasses	No
5	Milk	Oval	Speckles	Center	Maple nut	Yes
6	Milk	Oval	Smooth	Center	Chocolate cream	No
7	Milk	Rectangle	Dark drizzle	Center	Cherry nougat	Yes
8	Dark	Square	Dark drizzle	Edge	Butter caramel	No
9	Dark	Rectangle	Rippling	Edge	Toffee	No
10	Dark	Oval	Speckles	Center	Walnut nougat	Yes
11	Dark	Square	Speckles	Center	Coconut cream	No
12	Dark	Circle	Smooth	Center	Cherry cordial	No

From the unlabeled box, here are my choices:

ID	Color	Shape	Appearance	Position	Flavor	Tree nuts?
13	Milk	Rectangle	Speckles	Center	-	-
14	Dark	Circle	Smooth	Edge	-	-

Use a Multinomial Naïve Bayes classifier with pseudo-count = 1 to decide if:

Q2a (3pts): Chocolate 13 has tree nuts

$$P(Yes\ Tree\ Nuts\ |\ Chocolate\ 13) =\propto 0.25 \times 0.6 \times 0.2857 \times 0.375 \times 0.8 =\propto 0.0129$$

$$P(No\ Tree\ Nuts\ |\ Chocolate\ 13) =\propto 0.75 \times 0.4545 \times 0.3077 \times 0.1429 \times 0.3636 =\propto 5.45e^{-3}$$

→ Chocolate 13 **does** have tree nuts

Q2b (3pts): Chocolate 14 has tree nuts

$$P(Yes\ Tree\ Nuts\ |\ Chocolate\ 14) =\propto 0.25 \times 0.4 \times 0.1429 \times 0.125 \times 0.2 =\propto 3.5725e^{-4}$$

$$P(No\ Tree\ Nuts\ |\ Chocolate\ 14) =\propto 0.75 \times 0.5454 \times 0.2308 \times 0.2857 \times 0.6364 =\propto 0.172$$

→ Chocolate 14 **does not** have tree nuts

Q2c (1pts): If I want to try my luck with exactly one, which chocolate should I choose?

You should pick chocolate 14

Color:

	P(x|Yes Tree Nuts)	P(x|No Tree Nuts)
Milk	(2 + 1) / 5 = 0.6	(4 + 1) / 11 = 0.4545
Dark	(1 + 1) / 5 = 0.4	(5 + 1) / 11 = 0.5454
Total	(3 + 2) / 5 = 1.0	(9 + 2) / 11 = 1.0

Shape:

	P(x|Yes Tree Nuts)	P(x|No Tree Nuts)
Circle	(0 + 1) / 7 = 0.1429	(2 + 1) / 13 = 0.2308
Rectangle	(1 + 1) / 7 = 0.2857	(3 + 1) / 13 = 0.3077
Square	(0 + 1) / 7 = 0.1429	(3 + 1) / 13 = 0.3077
Oval	(2 + 1) / 7 = 0.4286	(1 + 1) / 13 = 0.1538
Total	(3 + 4) / 7 = 1.0	(9 + 4) / 13 = 1.0

Appearance:

	P(x|Yes Tree Nuts)	P(x|No Tree Nuts)
Smooth	(0 + 1) / 8 = 0.125	(3 + 1) / 14 = 0.2857
Rippling	(0 + 1) / 8 = 0.125	(2 + 1) / 14 = 0.2143
Dark drizzle	(1 + 1) / 8 = 0.25	(2 + 1) / 14 = 0.2143
Milk drizzle	(0 + 1) / 8 = 0.125	(1 + 1) / 14 = 0.1429
Speckles	(2 + 1) / 8 = 0.375	(1 + 1) / 14 = 0.1429
Total	(3 + 5) / 8 = 1.0	(9 + 5) / 14 = 1.0

Position:

	P(x|Yes Tree Nuts)	P(x|No Tree Nuts)
Edge	(0 + 1) / 5 = 0.2	(6 + 1) / 11 = 0.6364
Center	(3 + 1) / 5 = 0.8	(3 + 1) / 11 = 0.3636
Total	(3 + 2) / 5 = 1.0	(9 + 2) / 11 = 1.0

Flavor:

	P(x|Yes Tree Nuts)	P(x|No Tree Nuts)
Coffee Cream	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Toffee	(0 + 1) / 14 = 0.0714	(2 + 1) / 20 = 0.15
Caramel	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Molasses	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Maple nut	(1 + 1) / 14 = 0.1429	(0 + 1) / 20 = 0.05
Chocolate cream	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Cherry nougat	(1 + 1) / 14 = 0.1429	(0 + 1) / 20 = 0.05
Butter caramel	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Walnut nougat	(1 + 1) / 14 = 0.1429	(0 + 1) / 20 = 0.05
Coconut cream	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Cherry cordial	(0 + 1) / 14 = 0.0714	(1 + 1) / 20 = 0.1
Total	(3 + 11) / 14 = 1.0	(9 + 11) / 20 = 1.0

Tree Nuts:

P(Yes Tree Nuts)	3/12 = 0.25
P(No Tree Nuts)	9/12 = 0.75
Total	1.0