DSC 40A - Discussion 05 - Perceptrons and Probability

February 18, 2020

Problem 1.

Let $\vec{x}^{(1)}, \ldots, \vec{x}^{(n)}$ be a set of n data points with associated labels $y_1, \ldots, y_n \in \{-1, 1\}$. Recall from lecture that the perceptron update rule is

$$\vec{w}^{(t)} = \vec{w}^{(t-1)} - \frac{\alpha}{n} \sum_{i \in M} \begin{cases} \operatorname{Aug}(\vec{x}^{(i)}), & \operatorname{Aug}(\vec{x}^{(i)}) \cdot \vec{w}^{(t-1)} \geq 0 \\ -\operatorname{Aug}(\vec{x}^{(i)}), & \operatorname{Aug}(\vec{x}^{(i)}) \cdot \vec{w}^{(t-1)} < 0 \end{cases}$$

where M is the set of data points which are currently misclassified by $\vec{w}^{(t-1)}$.

Suppose we gather the following data:

$$\begin{array}{ll} x_1 = (1,0)^T & y_1 = 1 \\ x_2 = (0,1)^T & y_2 = 1 \\ x_3 = (-1,1)^T & y_3 = 1 \\ x_4 = (1,2)^T & y_4 = -1 \\ x_5 = (2,2)^T & y_5 = -1 \\ x_6 = (2,1)^T & y_6 = -1 \end{array}$$

Run the perceptron algorithm by hand with w initialized to $(1, 1, 1)^T$, and with a step size of $\alpha = 3$. Stop the algorithm when all points are classified correctly. What is the weight vector w at each step? How many training examples are misclassified at each step? For the purpose of this problem, a training example which lies on the decision boundary is considered to be misclassified.

Solution:
Step 0
Current weight vector $\vec{w}^{(0)} = (1, 1, 1)^T$
Misclassified points: x_4, x_5, x_6
Number of misclassified points: 3
Update rule to find $\vec{w}^{(1)}$:

$$\begin{aligned} \vec{w}^{(1)} &= (1,1,1)^T - (3/6)(\operatorname{Aug}(\vec{x}^{(4)}) + \operatorname{Aug}(\vec{x}^{(5)}) + \operatorname{Aug}(\vec{x}^{(6)})) \\ &= (1,1,1)^T - (1/2)((1,1,2)^T + (1,2,2)^T + (1,2,1)^T) \\ &= (1,1,1)^T - (1/2)(3,5,5)^T \\ &= (1,1,1)^T - (1.5, 2.5, 2.5) \\ \vec{w}^{(1)} &= (-0.5, -1.5, -1.5)^T \end{aligned}$$

Step 1
Current weight vector $\vec{w}^{(1)} = (-0.5, -1.5, -1.5)^T$
Misclassified points: x_1, x_2, x_3
Number of misclassified points: 3
Update rule to find $\vec{w}^{(2)}$: *The rest is left for you to complete until $\vec{w}$ converges.*

Problem 2.

I have 10 shirts, 6 pairs of pants, and 3 jackets. Every day I dress at random, picking one of each category. What is the probability that today I am wearing at least one garment I was wearing yesterday?

Solution:

The sample space is the set of possible outfits I can wear today. It is easier to calculate the complement, and find the probability that I am wearing all different clothes from yesterday.

Let A be the event that I am wearing all different clothes from yesterday.

We can see that $P(A) = (9/10) * (5/6) * (2/3)$. The reason for this is that today I can wear any of the other garments that I didn't wear yesterday.

Thus the probability of wearing at least one of the same garments as yesterday is $1 - P(A) = 1 - (9/10) * (5/6) * (2/3)$.

Problem 3.

Suppose that bitstrings (sequences of 0s and 1s) of length 4 are generated randomly with equal probability. Assume that positions are numbered $1, 2, 3, 4$ reading from left to right.

For each question below, find the probability. You do not need to simplify your answers at all. For example, you can leave answers like $2/4$ or $3/18 + 1/3$.

a) What is the probability that a bitstring starts with a one?

> **Solution:** $1/2$.
>
> There is one choice for the first element of the string, and two choices for the remaining 3 elements. This gives a total of $2^3 = 8$ possibilities. There are $2^4 = 16$ possible bitstrings. Thus the probability is $8/16 = 1/2$

b) What is the probability that a bitstring has a 1 in position 2 and a 0 in position 3?

> **Solution:** $1/4$.
>
> The second and third elements of the string are fixed, but there are 2 choices for the other two elements. This gives $2^2 = 4$ possibilities, and the probability is $4/16 = 1/4$.

c) What is the probability that a bitstring has the same first and last bits (the same bits in positions 1 and 4)?

> **Solution:** $1/2$.
>
> There are two ways in which a string can have the same first bit as last bit: both can be 1 or both can be zero.
>
> If both are one, there are 2^2 ways to pick the remaining two elements. Likewise is both are zero. Therefore there are $2 \cdot 2^2 = 8$ ways for the bitstring to have the same first element as last, and the probability of the event is $8/16 = 1/2$.

d) What is the probability that a bitstring has all four bits the same?

> **Solution:** There are two ways this can happen: 1111 or 0000. Therefore the probability is $2/16 = 1/8$.

e) What is the probability that a bitstring has more 0s than 1s?

Solution: 5/16

If the bitstring has more zeros than ones, then there are either three zeros or four zeros.

There is only one way to have four zeros: 0000.

There are $\binom{4}{3} = 4$ ways to have three zeros: 1000, 0100, 0010, 0001.

In total, the event has $1 + 4 = 5$ elements, and so the probability is 5/16.

Note: you should be able to do this problem for strings of longer length. In that situation you can't feasibly enumerate all strings with exactly three zeros, and so you have to use the formula for the number of combinations.

Problem 4.

There are 5 teams in a sports league. Over the course of a season, each team plays every other team exactly once. If the outcome of each game is a fair random coin flip, and there are no ties, what is the probability that some team wins all of its games?

> **Solution:** The sample space is the set of all possibilites for who wins each game played in the league.
>
> As there 5 teams in the league, each team plays 4 games, so for a team to be undefeated, it must win all 4 games. Every game to be won has a probability of 1/2. The probability that a particular team wins all 4 games is $(1/2)^4$.
>
> Let E_i be the event that team i is undefeated, for $i = 1$ to 5. The event that some team is undefeated is $E_1 \cup E_2 \cup \cdots \cup E_5$, so we want $P(E_1 \cup E_2 \cup \cdots \cup E_5)$. But the events E_i are disjoint subsets of the sample space since we can't have two undefeated teams, which means we can add their probabilities together and instead compute $P(E_1) + P(E_2) + \cdots + P(E_5)$, by the addition rule.
>
> The probability that team i wins all of its games is $P(E_i) = (1/2)^4$. As there are 5 teams, the total probability that some team is undefeated is $5 * (1/2)^4$.

Problem 5.

A standard deck of cards contains 13 cards (2 – 10, J, Q, K, A) in 4 suits (spades, clubs, diamonds, hearts) for a total of 52 cards. Hearts and diamonds are red; clubs and spades are black.

Suppose one card is drawn. What is the probability of the card being:

a) the ace of spades?

> **Solution:** 1/52

b) a diamond?

> **Solution:** 13 / 52

c) a heart or a diamond?

> **Solution:** $26/52 = 1/2$

d) an ace?

> **Solution:** $4/52 = 1/13$

e) a face card (jack, queen, or king)?

> **Solution:** $(3 \cdot 4)/52 = 12/52 = 3/13$

f) a red face card?

> **Solution:** There are 6 red face cards; three diamonds and three hearts: $6/52 = 3/26$

g) even-numbered?

Solution: There are 20 even-numbered cards; 5 for each suit. $20/52 = 5/13$

h) a club and ≤ 10?

Solution: There are 10 such cards: $10/52 = 5/26$

i) red and even-numbered and ≤ 6?

Solution: There are 6 such cards; three diamonds and three hearts. $6/52 = 3/26$.

Now a card is drawn, placed back into the deck; then a second card is drawn. What is the probability that:

j) the ace of spades is drawn, then the king of hearts?

Solution: There are 52^2 equally-probable outcomes. This is just one of them: $1/52^2$.

k) aces are drawn both times?

Solution: There are 16 outcomes consisting only of aces. $16/52^2$.

l) a red ace is drawn and a black ace is drawn?

Solution: There are 4 outcomes of the form (red ace, black ace): $4/52^2$.

m) the first card is red and the second card is black?

Solution: There are $26 \cdot 26$ such outcomes: $26^2/52^2 = 1/4$

n) at least one of the draws is a heart?

Solution: The event can be broken into three disjoint sets of outcomes: 1) outcomes of the form (heart, not a heart), 2) outcomes of the form (not a heart, heart), 3) outcomes of the form (heart, heart).

There are $13 \cdot (52 - 13)$ outcomes of the first kind, $(52 - 13) \cdot 13$ outcomes of the second kind, and $13 \cdot 13$ outcomes of the third kind. Therefore, there are $2 \cdot 13 \cdot 39 + 13 \cdot 13 = 13 \cdot (78 + 13) = 13 \cdot 91$ outcomes in the event. Hence the probability is $(13 \cdot 91)/52^2 = 1183/52^2 = 0.4375$

Another way of solving is that P(at least one heart) = 1 - P(no hearts) = $1 - (3/4) * (3/4) = 7/16 = 0.4375$