CDS: Machine Learning 2023 // Tutorial Week 2

Jochem Pannekoek, Daria Mihaila & Jasper Pieterse

September 26, 2023

Percepton

1.

We start with the following expression of the perceptron capacity:

$$C(P,N) = 2\sum_{i=0}^{N}\binom{P-1}{i}$$

a)

For $P \leq N$, using the fact that $\binom{n}{k} = 0$ for $n < k$, we can stop our summation at N = P - 1. This gives us the following expression:

$$C(P,N) = 2\sum_{i=0}^{P-1}\binom{P-1}{i}$$

Using the binomial theorem with $(1+\alpha)^n = \sum_{i=0}^{n}\binom{n}{i}\alpha^i$, we note that in our case $\alpha = 1$, thus we have:

$$C(P,N) = 2\sum_{i=0}^{P-1}\binom{P-1}{i} = 2(1+1)^{P-1} = 2^P$$

Therefore all problems (2^P) are linearly separable.

b)

For $P = 2N$, we have $P - 1 = 2N - 1$, thus the sum becomes:

$$C(P,N) = 2\sum_{i=0}^{2N-1}\binom{2N-1}{i} = 2\frac{1}{2}2^(2N-1) = 2^{2N-1} = 2^{P-1}$$

Therefore, exactly half of the problems are linearly separable.

2.

See the Jupyter notebook attached to the Brightspace submission for this exercise.

3.

See the Jupyter notebook attached to the Brightspace submission and Fig. 1. When $P > N$, the bound will be used, which is quite generous as you can see from Fig. 1.

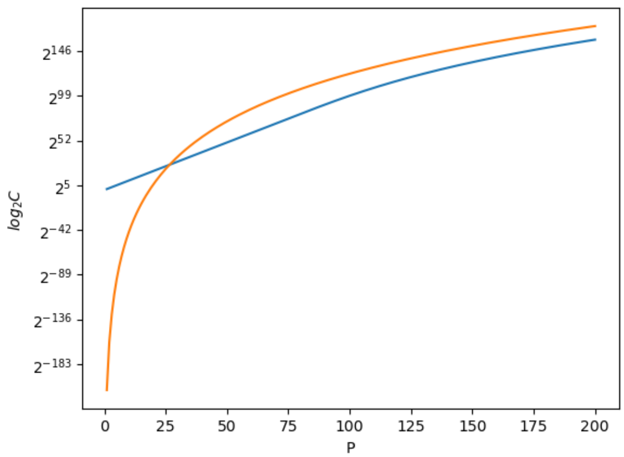

Figure 1: This shows the C(N,P) in blue and its bound in orange.

4.

a)

From the initial expression fot the generalisation bound, we have:

$$\delta = 4m(2P)\exp(\frac{\epsilon^2 P}{8})$$

After taking the logs on both sides and rearanging, we get the following expression for epsilon:

$$\epsilon = \sqrt{\frac{8}{P}\ln(\frac{\delta}{4m(2P)})}$$

For $\delta = 0.01$, $N = 10$ and ensuring that $\epsilon \approx 0.1$ and repeating for $N = 20, 30, 40, 50$ we have used the code in the attached Jupyter file ro observe the linear scaling of P with increasing N. Repeating for $N = 20, 30, 40, 50$, we get the $P \approx$ *something* for each case.

b)

See the Jupyter notebook attached to the Brightspace submission for this exercise.