MASTER 2 MICAS

Machine Learning Communications, and Security

Cours : MIACS911 Introduction To Statistic Learning.

Établissement : INSTITUT POLYTECHNIQUE DE PARIS

Enseignant : ASLAN TCHAMKERTEN

Sujet : Rapport Perceptron 06/10/2020...

Etudiant : Abdel Aziz Bounhar
Etudiant : Panongbene Jean Mouhamed Sawadogo.

I. SYNTHETIC DATA

Question 1 : Evaluate the computational complexity of the perceptron in terms of arithmetic operations per iteration.

Answer : For a Perceptron having for input variable of size m (size (X) = m) therefore a vector of weight W of size also m and an activation function HardLimiter :

$$\varphi(x) = \begin{cases} 1 & \text{if x} > 0 \\ 0 & \text{if x} \leq 0 \end{cases}$$

we have :
For the forward : For one iteration, we have

$$y = \varphi(W^T X) = \varphi(\sum_{i=0}^{m} w_i x_i)$$

so we have m multiplications and m additions followed by a comparaison so the complexity is $O(\mathbf{m^2})$
For the forward : For one iteration, we have

$$W = W + \eta(d - Y)X$$

so we have 1 subtraction, m+1 multiplication and m addition so the complexity is $O(\mathbf{m^2})$.
In the end, the complexity of the perceptron is $O(\mathbf{m^2})$

Question 2 : Look at the notebook for the code Consider four different values 0.05, 0.25, 0.50, 0.75 of the noise variance σ^2. For each of these values, run the perceptron over 50 randomly generated sets, compute the average error e(σ^2) and its standard deviation

$$s(\sigma^2) = \sqrt{\frac{1}{50}\sum_{i=1}^{50}(e_i - e)^2}$$

, where e_i denotes the fraction of misclassified points. Represent graphically e(σ^2)and s(σ^2) for the four values of σ^2 (use error bars). Comment.

Answer : Look at the notebook for the code We find that the mean value of the error and the standard deviation increases gradually with the value of the variance. Thus, we can say that the larger the standard ecc deviation, the larger the error interval.

Question 3 : Generate one data set with $\sigma^2 = 0.15$. A new random data set is now obtained by flipping each label d(n) with probability p to obtain d^1(n). Considering the generated 200 data set $\{x(n), d^1(n)\}_{n=1}^{200}$. repeat the previous experiments for p $\in \{0\%, 5\%, 10\%, 20\%\}$ and evaluate e(p) and σ^2(p). Comment.

Answer : Unlike question one, we find that the larger the value of p, the smaller the interval defined by the standard deviation and also the smaller the error. So we can say that the larger p, the smaller the interval defined by the standard deviation.

II. REAL DATA

2.1 Application to binary classification

Question 1 :

Answer : Look at the notebook

Question 2 :

Answer : Look at the notebook for the code
For a binary classification between the Iris-setosa and Iris-versicolor classes, the perceptron presents very good performances (the errors =0) . In fact, we observe in the graphic representation of the data that the two classes are linearly separated which justifies these good performances

2.2 (Optional) Application to multi-label classification

Answer : Look at the notebook