



Frequência

1. One of the main roots/dangers of the machine learning domain is known as the “*curse of dimensionality*”. What is this? In what sense it has been constraining the development of novel automated solutions for everyday problems?
2. Gradient Descent. Consider the following function: $f(x) = 3\frac{x^2}{e^x} + 4$.
 - a. Implement a Python function that optimizes the function $f(x)$, according to the gradient descent algorithm, and a learning rate (passed by parameter, along with the number of iterations and the initial value). You should not use any auxiliary libraries, apart “*numpy*”.


```
def gradient_descent(initial_value, lr, tot_steps)
```
 - b. Consider the value $x=3$, at time $t=0$. What will be the value of x at time $t=1$, using the gradient descent algorithm and a learning rate $lr=1$? Provide all the calculations required to obtain the answer.
3. Comment the following statement: “*if matrix inversion could be easily done for very large dimensions, then the gradient descent algorithm would lose much of its usefulness*”.
4. Consider a bidimensional feature space (where each instance is represented by two variables “ x ” and “ y ”) that regards a problem to distinguish “cats” (**O**) from “dogs” (**X**). Draw a 2D plot with 10 instances (5 “cats” + 5 “dogs”) where using the PCA algorithm before classification will yield catastrophic results.
5. Consider a binary classification problem, with three different solutions (systems “A”, “B” and “C”). Draw 3 ROC curves (in a single plot) corresponding to the following scenario: “*A is better than B in terms of security, while B is better than C in terms of usability. Also, C is better than A only in a certain range of the performance space*”.
6. Consider a unidimensional feature space, with elements represented by a single value, of two classes (**O/X**). Represent 10 instances (5O + 5X) in two different spaces where:
 - a. A logistic regression classifier will yield similar classification performance to a neural network.
 - b. A neural network would obtain much better performance than a logistic regression classifier.
7. What is typically the purpose of the “transfer learning” phase, in the self-supervised learning (SSL) paradigm? Can you describe one practical application of SSL that does not use transfer learning, but still can be considered useful?
8. Comment the following statement: “*using any set of handcrafted features as input of a deep learning model, is only reasonable if a very short amount of learning data is available. Otherwise, the raw data is preferred as input*”.
9. Unsupervised Learning. Illustrate three bidimensional feature spaces, composed of 10 instances of classes **O/X** (5O + 5X), where the K-means and DBScan algorithms:
 - a. would yield similar results.
 - b. would yield very different results, with advantage to “K-Means”, for class separability purposes.
 - c. would yield very different results, with advantage to “DBScan”, for class separability purposes.