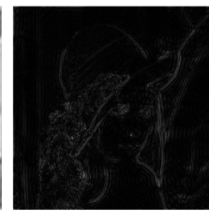




Exame Recurso

1. The notion of “feature” is among the most important in machine learning. What is a feature and why is it that important for the whole domain of knowledge?
2. Consider the problem of distinguishing “cats” from “dogs” using color images. Imagine and describe features used in such problem that:
 - a. Are permanent, but not discriminant.
 - b. Are discriminant, but not permanent.
 - c. A very good feature, providing discriminability and permanence simultaneously.
3. Gradient Descent. Consider the following function: $f(x) = 3\frac{x^2}{2} - 2x$.
 - a. During optimization, is it possible to get trapped in a local minimum? Why?
 - b. Implement a Python function that, upon a starting point, a learning rate and a number of iterations, provide the optimized value of “x”.


```
def otimize(start_pt, l_r, tot_iterations)
```
4. Unsupervised Learning. One of the main issues of unsupervised learning is to perceive the topology of the underlying features in the data set.
 - a. Illustrate one set of 2D instances (≥ 20 elements), where the space topology would be important to measure the similarity between elements.
 - b. For such “topology-sensitive” cases, the K-means algorithms is rarely the best choice. Why is this?
5. Comment the following statement: “*the number of parameters used in a model is directly related to the probability of overfitting, during optimization.*”
6. Consider the widely known “Lena” image as input (below, at left) and the images at right as the output of the convolution between “Lena” and two filters (one low-pass (A) and one high-pass (B)). Which images correspond to the outputs of “A” and “B”? Justify your answer.



7. During the empirical development and evaluation of one model (e.g., a neural network), what are the hyper-parameters? At what point they should be defined/chosen?
8. Comment the following statement: “*if using deep learning models in structured data (e.g., tabular), we lose most of the rationale behind this kind of models. However, they still might enable to obtain good performance, when compared to traditional models.*”
9. Illustrate one feature space in a 2D space, composed of 10 instances (5O + 5X), where a neural network of a single layer would have very poor discrimination performance.