TÉCNICO LISBOA

Learning-based Multimedia Processing

2022/2023

IST number: _____

Name: _____

Quiz #4 **Duration: 20 minutes**

Provide clear, legible, and succinct answers.
Always justify your assumptions.

Questions

1. In the context of machine learning problems, what is understood by cross-validation? When is it typically applied?

2. In the context of linear regression, the mean square error (MSE) is typically used to compute the cost function. Starting with an initial guess of the linear regression model parameters, how can those parameters be adjusted to reduce the cost function?

3. When using the gradient descent technique, is it guaranteed that the learning process will converge to the global minimum of the cost function?

4. When using the gradient descent technique, what are the main differences between using the batch, stochastic, or mini-batch approaches?

(batch)

(stochastic)

(min-batch)

5. A trained binary classifier model doesn't provide a good fit to the training data nor to the validation data.
What could be the problem of this model? How could it be improved?

1 –

Cross-validation partitions data into complementary subsets, training on one (larger) subset and testing on a disjoint subset; all combinations are evaluated and the results are combined (e.g. averaged) to estimate the model's predictive performance.

2 –

Use the gradient descent technique – compute the cost function as a function of the parameter values and update the parameters to follow direction of the cost function derivative;
Alternative: use the analytical solution of deriving the cost function mathematical expression and equalling it to zero – Normal equation

3 –

If the Cost function is not convex, the learning process may converge to a local minimum, or "get stuck" in a plateau.
If the cost function is convex (as the case of the MSE) it will converge to the global minimum.

4 -

Batch gradient descent: each step uses all training examples
Stochastic gradient descent: each step uses one training example
Mini-batch gradient descent: each step uses a subset of the training examples

5 –

The model is underfitting the data
Use an higher order / more complex model – e.g., add additional parameters