Machine Learning
Università della Svizzera italiana

Assignment 1

Surname Name

April 19, 2021

The assignment is split into two parts: you are asked to solve a regression problem, and answer some questions. You can use all the books, material, and help you need. Bear in mind that the questions you are asked are similar to those you may find in the final exam, and are related to very important and fundamental machine learning concepts. As such, sooner or later you will need to learn them to pass the course. We will give you some feedback afterwards.

!! Note that this file is just meant as a template for the report, in which we reported **part of** the assignment text for convenience. You must always refer to the text in the README.md file as the assignment requirements.

1 Regression problem

This section should contain a detailed description of how you solved the assignment, including all required statistical analyses of the models' performance and a comparison between the linear regression and the model of your choice. Limit the assignment to 2500 words (formulas, tables, figures, etc., do not count as words) and do not include any code in the report.

1.1 Task 1

Use the family of models $f(\mathbf{x}, \boldsymbol{\theta}) = \theta_0 + \theta_1 \cdot x_1 + \theta_2 \cdot x_2 + \theta_3 \cdot x_1 \cdot x_2 + \theta_4 \cdot \sin(x_1)$ to fit the data. Write in the report the formula of the model substituting parameters $\theta_0, \ldots, \theta_4$ with the estimates you've found:

$$f(\mathbf{x}, \boldsymbol{\theta}) = \_ + \_ \cdot x_1 + \_ \cdot x_2 + \_ \cdot x_1 \cdot x_2 + \_ \cdot \sin(x_1)$$

Evaluate the test performance of your model using the mean squared error as performance measure.

1.2 Task 2

Consider any family of non-linear models of your choice to address the above regression problem. Evaluate the test performance of your model using the mean squared error as performance measure. Compare your model with the linear regression of Task 1. Which one is **statistically** better?

1.3 Task 3 (Bonus)

In the **Github repository of the course**, you will find a trained Scikit-learn model that we built using the same dataset you are given. This baseline model is able to achieve a MSE of **0.0194**, when evaluated on the test set. You will get extra points if the test performance of your model is better (i.e., the MSE is lower) than ours. Of course, you also have to tell us why you think that your model is better.

2 Questions

2.1 Q1. Training versus Validation

1.Q. Explain the curves' behavior in each of the three highlighted sections of the figures, namely (a), (b), and (c);

1.A.

2.Q. Is any of the three section associated with the concepts of overfitting and underfitting? If yes, explain it.

2.A.

3.Q. Is there any evidence of high approximation risk? Why? If yes, in which of the below subfigures?

3.A.

4.Q. Do you think that by further increasing the model complexity you will be able to bring the training error to zero?

4.A.

5.Q. Do you think that by further increasing the model complexity you will be able to bring the structural risk to zero?

5.A.

2.2 Q2. Linear Regression

Comment and compare how the (a.) training error, (b.) test error and (c.) coefficients would change in the following cases:

1.Q. x_3 is a normally distributed independent random variable $x_3 \sim \mathcal{N}(1, 2)$

1.A.

2.Q. $x_3 = 2.5 \cdot x_1 + x_2$

2.A.

3.Q. $x_3 = x_1 \cdot x_2$

3.A.

2.3 Q3. Classification

1.Q. Your boss asked you to solve the problem using a perceptron and now he's upset because you are getting poor results. How would you justify the poor performance of your perceptron classifier to your boss?

1.A.

2.Q. Would you expect to have better luck with a neural network with activation function $h(x) = -x \cdot e^{-2}$ for the hidden units?

2.A.

3.Q. What are the main differences and similarities between the perceptron and the logistic regression neuron?

3.A.