**INTRODUCTION MACHINE LEARNING**

**EXERCISE 3**

A red sign with white text

Description automatically generated

|  |
| --- |
| **TEACHER:** |
| Johannes Kiesel |

|  |
| --- |
| **GROUP:** |
| Group 16 |
|  |
| **SUBMITTED BY:** |
| Aaron Perez Herrera |
| Cesar Fernando Gamba Tiusaba |
| Chun Ting Lin |
| Olubunmi Emmanuel Ogunleye |

Exercice 1: Linear Models:

1. Name these concepts:

The point-wise loss

Global loss

 : objective function or error function

: Parameter vector or (p+1)-dimensional hypothesis.

* direction vector excluding wo or the p-dimensional direction vector.

1. How would the figure below change if wo is halved?

The separating line in the figure corresponds to the equation of the decision boundary:

When w0 halved, the intercept of the line on the x​ plane changes. Specifically:

W0 affects the position of the decision boundary but **not its orientation**, as the weights w1​​ determine the slope.

* The absolute value of X​ increases, meaning the intercept moves further away from the origin. The boundary shifts outward along the direction normal to w1​.
* Since w1​ remains unchanged, the slope of the decision boundary does not change. Thus, the boundary retains its orientation, and only its position is affected.
* The blue line shifts is the original equation, when W0 is halved, the orange line is the shifted equation, some of the points (red ones) keep the label, but some other points (black points) change of the label

1. What is the difference (if any) between decision boundaries for linear and logistic regression:

The key difference is the predicted output.

* Linear regression: predicts a continuous output and its decision boundary is a straight line.
* Logistic regression: predicts a probability (between 0 and 1) and uses a sigmoid function to map a linear combination of features to this probability. The decision boundary in logistic regression can take a nonlinear form.

1. The lecture notes slides state that a key difference between ridge  and lasso  regression is that, with lasso regression, parameters can be reduced to zero. Explain why.

This is because Lasso regression uses L1 regularization, which adds a penalty term proportional to the absolute value of the coefficients. As a result, some coefficients can be reduced to zero, effectively removing the corresponding features from the model

1. Why can the gradient descent method not be applied for ?

can not be expressed as a differentiable function, also because the loss function considered when applying 0/1 loss is typically non-convex.

Exercise 2: Pointwise Loss Functions

In the lecture notes, slide ML:III-63 on loss computation for logistic regression in detail, the rightmost plot "Loss over hyperplace distance" shows the pointwise logistic and 0/1 loss for a logistic regression model for ,that means, for examples with c = 1. In this exercise you will investigate the case of examples with c = 0.

1. Show that

For c = 0

1. Draw the plot "Loss over hyperplane distance" for examples with c = 0, showing both logistic loss and 0/1 loss.