Universität des Saarlandes
Prof. Dr. Dietrich Klakow
Lehrstuhl für Signalverarbeitung
NNTI Winter Term 2020/2021

Exercise Sheet 4

Machine Learning Basics (Part 2)

Deadline: 08.12.2020, 23:59

Instructions

This assignment covers the remaining concepts in Machine Learning basics including Logistic Regression, Support Vector Machines, and cost minimization.

Exercise 4.1 - Logistic Regression (1 + 1 + 1 points)

This week you were introduced to Logistic Regression (refer here and here), a machine learning approach used for classification. For the dataset $D_N = \{(x^{(i)}, t^{(i)})\}$ with target label $t^{(i)} \in \{0, +1\}$, logistic regression is defined using the following steps:

$$z = w^T x + b$$
$$y = \sigma(z)$$
$$L(y, z) = -t \log(y) - (1 - t) \log(1 - y)$$

a) Derive the equivalent cost minimization problem without the intermediate variables y and z. Your cost function should only depend on variables w and b, and the dataset D or number of entries N.

b) Instead of $t^{(i)}$, let us assume a new target label $\widetilde{t}^{(i)} \in \{-1, +1\}$. How does this change the minimization problem that your derived in the previous question?

c) Consider a logistic regression model where the class probabilities can be obtained using

$$z(x) = \sigma(w^T x + b)$$

where w and b are the parameters and we classify using the rule

$$y(x) = \mathbb{1}[z(x) > 0.5].$$

Prove that this model corresponds to a linear decision boundary in the input space.

Exercise 4.2 - Support Vector Machines (SVMs) (1 + 2 + 1 + 1 points)

Let us consider a classification problem with target class $y_i \in \{+1, -1\}$ which we want to solve using SVMs (see here and DL book page 139 to know more). For any input $\vec{x}$, consider the vector $\vec{w}$ (perpendicular to the classifier) such that $\vec{w}.\vec{x} + b \geq 0$ determines the class of the input.

a) Derive an expression for the width of the classification margin.

b) SVMs are known as *maximal-margin classifiers*. This constraint-based optimization problem can be solved using the *Lagrangian* (see here for more information) below:

$$L = X - \sum_i \alpha_i [y_i(\vec{w}.\vec{x_i} + b) - 1]$$

where α_i is the Lagrange multiplier and X is the term you want to optimize over i samples. Now try to maximize your margin from part (a) using the above equation. (Hint: Your answer should not contain w or b.)

c) What is the significance of the term α_i in SVMs? *(max. 2 sentences)*

d) Notice that we were dealing with linear data till now. Is it possible to change your maximized margin (from part (b)) so that it works for non-linear data as well? Justify your answer. There is no need for mathematical proofs or derivations, but you can support your answer with relevant equations. *(max. 3-4 sentences)*

Exercise 4.3 - Clash of the Kernels! (1 + 1 points)

Anika and Lena are childhood friends who both like to classify data, but using different methods. While Lena prefers SVMs, Anika loves logistic regression. Everyone loves Lena more because she claims she knows how to classify even non-linear data points using something called *kernels*. Recently, after reading this, Anika claims that she should also be able to do something like Lena!

a) Try to prove Anika's claim. Introduce the *kernel* function in the simple logistic regression model and try to relate to the value of $\vec{w}$ that you had obtained in the *Lagrangian* of Q.4.2(b) to support your answer. *(max. 3 steps)*

b) Briefly compare both their *kernel* methods on the basis of their complexities and support vectors. *(max. 2 sentences)*

Submission instructions

The following instructions are mandatory. If you are not following them, tutors can decide to not correct your exercise.

- You can and are encouraged to submit the assignment as a team of two students. Submitting as a team will be mandatory for the next assignment.
- Hand in zip file containing the PDF with your solutions and the completed iPython notebook, or you could also hand in just the iPython notebook containing everything.
- Therefore make sure to write the Microsoft Teams user name, student id and the name of each member of your team on your submission.
- Your assignment solution must be uploaded by only **one** of your team members to the 'Assignments' tab of the tutorial team (in **Microsoft Teams**).
- If you have any trouble with the submission, contact your tutor **before** the deadline.