Technical University of Kaiserslautern

Memory Protocol of

Machine Learning 1

Given by Prof. Marius Kloft

Written Exam

Date: 12.02.2018

1 General

[3 + 3 + 3 + 3 + 3 + 3 = 18 Points]

[3 Points]

Why is optimizing nice with convex functions?

☐ They are easy to compute

☐ ???

☐ ???

[3 Points]

What is k-Means?

☐ Unsupervised learning

☐ Classification

☐ Regression

[3 Points]

The gradient of a function f goes in the _______ direction.

☐ Steepest ascent

☐ Steepest descent

☐ global Minimum

[3 Points]

What is Supervised learning?

[3 Points]

Explain Overfitting.

[3 Points]

Explain Random Forest.

2 SVM

[5 + 5 + 5 = 15 Points]

a) Consider the following form of Regression:

$$\min_{w \in \mathbb{R}^d, \xi \in \mathbb{R}^n} \frac{1}{2} \parallel w \parallel^2 + \frac{C}{2n} \sum_{i=1}^{n} \xi^2$$

Calculate the dual Lagrangian Problem

[5 Points]

b) Calculate the derivations.

[5 Points]

c) What is the Lagrangian dual problem.

3 Kernels

[3 + 3 + 3 + 4 + 4 = 17 Points]

[3 Points]

a) Explain the Kernel trick.

[3 Points]

b) Given the polynomial Kernel with Input Data x_1, x_2. What is roughly the dimension of the feature map if $d = 2$?

- ☐ $log(d)$
- ☐ d^2
- ☐ d^3
- ☐ e^d

[3 Points]

c) Give an example but exact definition for a kernel of your choice.

[4 Points]

d) Let $k(x, y)$ be a kernel. Further we define a function $\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$ if $x \neq y$. Show that $k'(x, y) = k(x, y) + \delta(x, y)$ is also a kernel.

[4 Points]

e) Until now, we only considered Kernels they map into the Euclidean space. Now we want to consider a kernel with the mapping function $\phi : \{0, 1\} \rightarrow \{0, 1\}$ with the following Definition: $k(0, 0) = k(1, 1) = 1$ and $k(0, 0) = k(1, 1) = \frac{1}{2}$. Show, k is also a kernel.

4 Regression

[? + ? + ? + 5 = ? Points]

[? Points]

a) Consider the following form of Regression:

$$\operatorname*{argmin}_{w \in \mathbb{R}^d} \sum_{i=1}^{n} (y_i - \cancel{w^T x_i})^2 \;=\; \operatorname*{argmin}_{b \in \mathbb{R}} \sum_{i=1}^{n} (y_i - b)^2$$

Calculate the Optimal value of b.

[? Points]

b) Consider the following two Optimization Problems:

$$\operatorname*{argmin}_{b \in \mathbb{R}} \sum_{i=1}^{n} (y_i - w^T x_i)^2$$

and

$$\operatorname*{argmin}_{b \in \mathbb{R}} \sum_{i=1}^{n} |y_i - w^T x_i|$$

Give an advantage and disadvantage for each of those Regression forms.

[? Points]

c) ???

[5 Points]

d) Calculate a closed form solution for w^*

$$w^* \coloneqq \operatorname*{argmin}_{w \in \mathbb{R}} \frac{1}{2} \parallel w \parallel^2 + \sum_{i=1}^{n} (y_i - w^T x_i)^2$$

5 k-Means Algorithm

[5 + 3 = 8 Points]

a) Given Pseudocode of the k-Means implementation. Find the error in the code.

Algorithm 1 Create_Cycle_Basis

```
procedure K_MEANS(Input Data x_1, x_2 ..., x_n; Clusters Center c_1, c_2, ..., c_m)
    index_set := {1, 2, ... m}
    for i = 1 to m do
        j ← Random_Entry(index_set)
        c_i ← x_j
        index_set.remove(j)
    end for
    repeat
        for i = 1 to m do
            clusterAssignment[i] == ← argmin_j ||x_i − c_j||^2
        end for
        for i = 1 to n do
            c_j ← average{x_j | clusterAssignment[j] == i}
        end for
        for i = 1 to m do
            c'_j ← c_j
        end for
    until c'_j == c_j for all j ∈ {1, 2, ..., m}
    return c_1, c_2, ..., c_m
end procedure
```

b) Describe in detail how you would fix the bug.

6 Principle component analysis

[5 + 5 + 5 = 15 Points]

[5 Points]

a) Write in pseudocode how to reduce given input $x_1, ..., x_n$ with principle component vectors $v_1, ..., v_m$ into lower dimension m.

[5 Points]

b) ???

[5 Points]

c) Write **code** how to plot given input onto the Euclidean plane (scatter plot) in a language of your preference. Make sure to also import any library you may use.

7 Neural Networks

[3 + 4 + 4 = 11 Points]

[3? Points]

a) A Neural Networks can calculate a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ that calculates for each given input $x \in \mathbb{R}^d$ the function value $f(x)$:

True ☐

False ☐

[4? Points]

b) What does the learning rate do? How would you go about choosing a specific learning rate?

[4? Points]

c) Describe in 2 or 3 sentences what backpropagation does.