(Gedächtnis-)protokoll
Punktanzahl insges. 100, Punkte stehen in den Klammern

1. Verschiedenes (35)

What is the disserence between regression and classification? (2)

You are a reviewer for the International Mega-Conference on Machine Learning of Outrageous Stuff, and you read a paper that selected a small number of features out of a large number of features for a given classification problem (Wie Übung 2, Thema 5: Kernels). Would you accept or reject the paper? Why? − 3 reviews (6)

Multiple choice: statements about LDA (4)

Multiple choice: which of these Algorithms solve classification problems? (4)

Multiple choice: which methods do $\max\limits_{\omega} \frac{\omega^T S \omega}{\omega^T \omega}$ (4)

Multiple choice: Given $\lambda_0 = 1$ and for all other Lambdas $\lambda_{Rest} = 0$. What can we conclude? (4)

Multiple choice: statements about perceptrons (4)

Explain wheather perceptron and LDA are depending on initialization or not? (3)

What strategies have we learned in the lecture how to handle data that is not linearly separable? (4)

2. (10)

Sketch a tapical LDA and an OLS data set (4)

Imagine, you would like to do for Twitter something, that filters out only the 10000 relevant words (Text war länger, mehr Informationen). Which algorithm would you choose for that? (6)

3. (11)

Which of these given covariance matrices belong to which of the given data-sets (drawn as pictures) (3)

Draw the Eigenvectors in the pictures above (2)

Compute Eigenvalues and Eigenvectors of $\begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$. (6 Punkte)

4. (8)

What is the main idea behind k-means-clustering and PCA? (4)

What are their similarities and what are their differences? (2)

Give an example for each of them (2)

5. (13)

Explain overfitting and underfitting. Why does cross validation help? (6)

Given this pseudocode for Cross Validation Algorithm. Find all bugs and correct them. (4)

Draw overfitting, underfitting and good fit in these three pictures with given datapoints. (3)

6. (10)

OLS: We have a given matrix X and vector y: compute ω (4)

Apply the model to all 4 datapoints (2)

evaluate your results, is it a good model? (4)

7. (13)

What is the difference between linear and ridge regression? (4)

apply forward propagation to all 4 datapoints (same datapoints as in task 6) (5)

compare your solution to the OLS solution in task 6. Which one is the better model? Why? (4)