

# Student's exam questions

You can choose among the following topics that cover the whole supervised learning part of the lecture. Ask questions testing the understanding or transferring capabilities of your fellow students.

- Bayesian Decision Theory
- Empirical Risk Minimization
- Linear Least Squares Regression
- Lasso
- Ridge
- Cross Validation
- Logistic Regression
- Evaluation of classification results
- Kernels
- Support Vector Machines
- Random Forests
- Boosting
- Bootstrap
- (kernel) PCA

## Exercise 1 (Kernels, understanding)

Name one advantage and disadvantage of the Kernel Trick

SOLUTION: The main advantage of the Kernel Trick is that one can solve a non-linear problem with a linear classifier. One huge disadvantage is that the output depends on the kernel choice: If one assumes any underlying distribution of the data, the kernel can be chosen accordingly. However, if there are no underlying assumptions or inductive biases the kernel trick does not work well with arbitrary kernels.

## Exercise 2 (Linear Least Squares Regression, understanding)

Explain the terms bias and variance in the Regression Setting and relate them to ridge vs lasso.

SOLUTION: Bias describes underlying assumptions that are used to simplify and generalize the model. The goal is to decrease the complexity of the model and its sensitivity to outliers.

Variance describes the errors of a model to individual data points. Low variance means that the individual data points are fitted closely, which increases the risk of overfitting. Meanwhile higher variance is often related to a strong generalization of the data points which corresponds to underfitting in the extreme case.

The squared regularization term of ridge regression penalizes the weights of the fitted function strongly. Ridge regression rather focuses on reducing the variance.

Lasso regression sets multiple coefficients to zero resulting in the decreasing complexity of the model. Therefore Lasso rather focuses on decreasing the bias

## Exercise 3 (Model Complexity, transfer)

For each of the following methods, describe briefly how the bias can be decreased (no need to explain the methods)

- Logistic Regression
- K-nearest Neighbors
- Polynomial (Function) Fitting

- Lasso Regression

SOLUTION:

- Logistic Regression: increase number of regression coefficients
- K-nearest Neighbors: decrease k
- Polynomial (Function) Fitting: increasing polynomial degree
- Lasso Regression: decreasing the regularization parameter  $\lambda$