Exam themes:

- bias-variance trade-off and overfitting/underfiting in relation to training and testing error
- difference between parametric and non-parametric methods
- different validation strategies used to estimate testing error
- common model evaluation metrics; R^2, adjusted R^2, Accuracy, Sensitivity, Specificity; RSS, MSE
- How are decision trees constructed?
- Comparison of decision trees and different tree ensemble methods (bagging, boosting, random forest)
- Rationale of different method stacking
- Main problem with high dimensional data
- Main problem when we have p>n (or p>>n) situation and possible ways to deal with it
- difference between lasso and ridge methods
- different ways to perform feature selection
- General approach for hyperparameter selection in various models
- ROC curve interpretation
- Main idea of meanshift, DBscan, k-means and hierarchical clustering algorithms (intuition and understanding, formulas are not needed)
- KNN algorithm
- Four main hierarchical clustering linkage methods, how is the tree influenced by them?
- How to evaluate if clustering is good?
- PCA plot interpretation
- What is labeled and unlabeled data?
- What is a decision boundary?
- Which classifiers create a linear and which create a nonlinear decision boundary?
- How a formula of a linear discriminant (LDA) looks like?
- LDA formula is a formula of a hyperplane (or a line if in two dimensions)
- If the data points in the dataset follow multivariate Gaussian distribution what parameters of the distribution define coefficients of the linear discriminant?
- What is a geometrical meaning of a weight vector of linear discriminant?
- Which are two distinctive features of Support Vector Machine that makes it different from a simple linear discriminant?
- Bayes formula that forms basis to build Naive Bayes classifier
- What simplest operation does an artificial neuron performs on its inputs?