M144	**Statistical learning**	L	P	S	ECTS
		2	0	3	7

Course objectives. To acquaint students with concepts, methods and algorithms of statistical learning. Introduce abstract learning models and basic concepts of statistical learning theory. Cover the main methods of supervised and unsupervised learning such as decision trees, support vector machines and neural networks with special emphasis on data applications.

Prerequisites. Statistics. Multivariate analysis.

Course content.

1. Introduction. Types of learning problems and illustrations on practical examples.
2. Theory of statistical learning. Formal learning models. PAC learning and the loss function. Consistency and convergence of the learning process. Vapnik-Chervonenkis dimension.
3. Introduction to methods. Model validation and selection.
4. Overview of linear methods. Selection. Ridge and lasso regression. Linear models for classification. Additive models.
5. Trees. Boosting and additive trees. Random forests.
6. Support vector machines for classification and regression. Generalized linear discriminant analysis.
7. Deep learning. Neural networks. Parameter estimation methods.
8. Unsupervised learning methods.

LEARNING OUTCOMES

No.	LEARNING OUTCOMES
1.	Demonstrate the characteristics and properties of the models covered in the course content.
2.	Recognize problem types suitable for statistical learning techniques and appropriate methods.
3.	Explain the processes of statistical learning with mathematical arguments and understand theoretical concepts.
4.	Identify appropriate methods, estimate model parameters and apply validation procedures.
5.	Critically analyze models and understand the limitations of applied methods.
6.	Study and apply new literature on statistical learning methods.
7.	Present modeling results and application possibilities to experts and a wider audience.

RELATING THE LEARNING OUTCOMES, ORGANIZATION OF THE EDUCATIONAL PROCESS AND ASSESSMENT OF THE LEARNING OUTCOMES

TEACHING ACTIVITIES	ECTS	LEARNING OUTCOME **	STUDENT ACTIVITY*	EVALUATION METHOD	POINTS	
					min	max
Attending lectures and exercises	0.5	1-7	Lecture attendance, discussion, teamwork, independent work on given tasks and short examination	Attendance lists, tracking activities	0	5
Homework	2.5	4-7	Solving theoretical and practical problems	Evaluation	15	25
Written exam (Mid-terms)	2	1-5	Preparing for written exam	Evaluation	25	50
Final exam	2	1-7	Revision	Oral exam	10	20

TOTAL	7				50	100

Teaching methods and student assessment. Lectures and exercises are obligatory. The final exam is oral, taken after the completed lectures and exercises and achieved minimum number of credits at the midterm exams and homework.

Can the course be taught in English: Yes

Basic literature:

1. T. Hastie, R. Tibshirani, J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction, New York: Springer, 2nd edition, 2016.
2. G. James, D. Witten, T. Hastie, R. Tibshirani. An introduction to statistical learning, New York: Springer, 2nd edition, 2021.

Recommended literature:

1. V. Vapnik, The nature of statistical learning theory, Springer science & business media, 1999.
2. S. Ben-David, S. Shalev-Shwartz, Understanding machine learning: From theory to algorithms, Cambridge University Press, 2014.
3. K. P. Murphy, Probabilistic machine learning: an introduction, MIT press, 2022.
4. M. Mohri, A. Rostamizadeh, A. Talwalkar, Foundations of machine learning, MIT press, 2018.