DOCTORAL (PHD) STUDIES

COURSE UNIT DESCRIPTION

Course unit title	Scientific areas	Faculty	Institute, department
Machine learning	T 007	Mathematics and Informatics faculty	Institute of Data Science and Digital Technologies

Study method	Number of credits	Study method	Number of credits
Lectures	1 (autumn semester)	Consultations	1
Individual works	4	Seminars	1

Summary

Prerequisite: It is desirable for a doctoral student will have base knowlage to linear algebras', and mathematical analysis concepts, and will know Python programming language.

The aim of this subject is to gain knowledge about machine learning methods and technologies in order to apply them for scientific research.

Content:

1. Introduction to Machine Learning (ML). Basic concepts. Types of machine learning systems: supervised/unsupervised learning, batch and online learning, instance-based and model-based learning. Main challenges of machine learning. model testing and validating.
2. Structure and execution of typical ML project. Look at the Big picture. Get the data. Discover and visualize the data to gain insights. Prepare the data for ML algorithms. Select and train a model. Fine-tune model.
3. Classification. Training a binary classifier. Performance measures. Multiclass classification. Error analysis. Multilabel and multioutput classification. Bayesian classifier.
4. Regression. Linear regression. Polynomial and logistic regression. Gradient descent. Regularized linear regression models. Learning curves.
5. Other classification and regression methods. Support vector machines (SVM). Linear SVM classification, Nonlinear SVM classification. SVM regression. Decision trees. Gini Impurity and Entropy. Ensemble learning. Voting classifiers. Bagging, pasting, boosting, and stacking. Random Forests.
6. Dimensionality Reduction. The Curse of Dimensionality. Main approaches for dimensionality reduction. PCA, Incremental PCA, Kernel PCA, LLE, and other dimensionality reduction techniques.
7. Neural networks and deep learning. Introduction to artificial neural networks. Training Deep Neural Nets.
8. Convolutional neural networks.
9. Recurrent neural networks.
10. Autoencoders. Efficient Data Representation. Stacked, Sparse, Variational, Denoising, and other autoencoders.
11. Machine learning in Natural language processing. Word vectors, embeddings, text classification and clustering. Machine translation. Transformers.
12. Reinforcement Learning. Learning to optimize reward. Policy search. Introduction to OpenAI Gym. Policy gradients, Markov decision processes. Temporal difference learning and Q-learning.
13. Graph neural networks

Assignments:

1. Application of classical classification and regression algorithms. Projects based on 1-5 topics.
2. Development of specialized algorithms. Individual task taking into account the research topic of the doctoral student. In the tasks student should apply artificial neural networks, autoencoders or reinforcement learning.

The course will consist of 14 lectures, 14 seminars, two home works and two student projects.

First project will be from 1-5th lectures topics, and the second project will be based on doctoral student thesis topic.

Main literature
Aurélien Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, 2019
Ian Goodfellow and Yoshua Bengio and Aaron Courville. Deap Learning, 2016, MIT press, http://www.deeplearningbook.org
John D. Kelleher, Brian Mac Namee, Aoife D'Arcy.Fundamentals of Machine Learning for Predictive Data Analytics, The MIT Press , 2015
Hal Daumé III, A Course in Machine Learning, 2017, https://www.ciml.info/dl/v0_99/ciml-v0_99-all.pdf
Michael U. Gutmann, Pen and Paper Exercises in Machine Learning, 2022, https://arxiv.org/abs/2206.13446
Richard S. Sutton and Adrew G.Barto Reinforcement Learning: An Introduction. 2017, http://incompleteideas.net/book/bookdraft2017nov5.pdf
William L. Hamilton, Graph Representation Learning Book, McGill University, 2020, https://www.cs.mcgill.ca/~wlh/grl_book/
Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning:Data Mining, Inference, and Prediction. Second Edition, https://hastie.su.domains/ElemStatLearn/download.html
Valliappa Lakshmanan, Sara Robinson, Michael Munn, Machine Learning Design Patterns: Solutions to Common Challenges in Data Preparation, Model Building, and MLOps, 2020

Lecturer(s) (name, surname)	Science degree	Main publications
Virginijus Marcinkevičius	Dr.	http://www.elaba.mb.vu.lt/dmsti/?aut=Virginijus+Marcinekvičius
Linas Petkevičius	Dr.	http://www.elaba.mb.vu.lt/mif/?aut=Linas+Petkevičius
Vytautas Valaitis	Dr.	http://www.elaba.mb.vu.lt/mif/?aut=Vytautas+Valaitis

Doktorantūros komiteto teikimu patvirtinta fakulteto/instituto taryboje 20__ m.__ mėn.__d. , protokolo Nr.
Tarybos pirmininkas

Pastaba: jei doktorantūros teisė bus suteikta kartu su kita institucija, tvirtinama ne fakulteto taryboje , o jungtinėje komisijoje.