

DOKTORANTŪROS STUDIJŲ DALYKO SANDAS

Dalyko pavadinimas	Mokslo kryptis (šaka) kodas	Fakultetas	Institutas, katedra
Mašininis mokymasis	T 007	Matematikos ir informatikos fakultetas	Duomenų mokslo ir skaitmeninių technologijų institutas

Studijų būdas	Kreditų skaičius ECTS	Studijų būdas	Kreditų skaičius
paskaitos	1 (rudens semestras)	konsultacijos	1
individualus	4	seminarai	1

Dalyko anotacija
Reikalavimai: Tikimasi, jog doktorantas pasirenkantis šį dalyką turės tiesinės algebros, matematinės analizės bazines žinias, bei mokės Python programavimo kalbą. Pagrindinis kurso tikslas – padėti studentams įgyti reikiamas žinias apie mašininio mokymosi metodus ir su jais susijusias technologijas, siekiant juos pritaikyti moksliniuose tyrimuose. Dalyko temos: - Įvadas į mašininį mokymąsi. Pagrindinės sąvokos. Mašininio mokymo sistemos (mokymas su mokytoju ir be, mokymas realiu laiku ir paketais, modelių ir individų panašumu paremtas mokymas). Pagrindiniai iššūkiai mašiniame mokyme. Modelių testavimas ir validavimas. - Tipinio mašininio mokymosi projekto struktūra ir atlikimas. Problemos formulavimas. Duomenų gavimas. Duomenų vizualizavimas ir pažinimas. Duomenų parengimas mašiniam mokymuisi. Modelio parinkimas ir mokymas. Modelio derinimas. - Duomenų klasifikavimas. Binarinio klasifikavimo uždavinys. Klasifikatoriaus rezultatų vertinimo metrikos. Paklaidų analizė. Daugelio klasių ir daugelio žymenų klasifikavimas. Artimiausių kaimynų ir Bajeso klasifikatoriai. - Regresijos metodai. Tiesinės, polinominės ir logistinės regresijos modeliai. Gradientinis nusileidimas. Parametrų regularizacija. Mokymosi kreivės. - Kiti klasifikavimo ir regresijos metodai. Atraminų vektorių mašinos (tiesinės, netiesinės). Sprendimo medžiai ir informacijos matavimo matai. Klasifikavimo metodų ansambliai (balsavimas, pakavimas (angl. bagging), auginimas (angl. boosting) ir kt.). Atsitiktinių miškų metodas (angl. Random forests) - Dimensijos mažinimo metodai. Projekcijos metodai – principinių komponentų analizė ir įvairios jos versijos. Daugdaros išmokymo metodai – LLE. Kitos dimensijos mažinimo technikos. - Neuroniniai tinklai ir gilus mokymas. Įvadas į dirbtinius neuroninius tinklus. Neuroninių tinklų ir gilių neuroninių tinklų mokymas. - Konvoliuciniai neuroniniai tinklai. - Rekurentiniai neuroniniai tinklai. - Autoenkoderiai. Autoenkoderiai – efektyvus duomenų reprezentavimo priemonė. Gilūs autoenkoderiai. Reti ir tikimybiniai autoenkoderiai. Autoenkoderių taikymas triukšmo šalinimui duomenyse, signalų suspaudime. - Natūralios kalbos apdorojimas. Žodžių vektoriai, įterpiniai, teksto klasterizavimas ir klasifikavimas. Mašininis vertimas. Transformerių neuroniniai tinklai. - Skatinamasis mokymasis. Taisyklių paieška. Įvadas į OpenAI Gym. Strategijų mokymas. Markovo sprendimų priėmimo procesas. Mokymasis iš pokičių laike ir Q-mokymasis. - Grafų neuroniniai tinklai. Praktinės užduotys: - Klasikinių klasifikavimo ir regresijos metodų taikymas. Atsiskaitomas projektinis darbas iš 1-5 temų.

2. Specializuotų algoritmų kūrimas. Atsiskaitomas projektinis darbas iš doktoranto vykdomos mokslinės temos parinktos užduoties, kurios sprendimui būtų naudojami neuroniniai tinklai, autoenkoderiai arba skatinamojo mokymosi metodai.

Modulį sudarys 14 paskaitų, 14 seminarų, du atsiskaitomieji darbai, du namų darbai.

Pagrindinė literatūra

Aurélien Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, 2019

Ian Goodfellow and Yoshua Bengio and Aaron Courville. Deep 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 Andrew 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

Konsultuojančiųjų dėstytojų vardas, pavardė	Mokslo laipsnis	Svarbiausieji darbai mokslo kryptyje (šakoje) paskelbti per pastaruosius 5 metus
Virginijus Marcinkevičius	Mokslų daktaras	http://www.elaba.mb.vu.lt/dmsti/?aut=Virginijus+Marcinkevičius
Linas Petkevičius	Mokslų daktaras	http://www.elaba.mb.vu.lt/mif/?aut=Linas+Petkevičius
Vytautas Valaitis	Mokslų daktaras	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.