

## DOKTORANTŪROS STUDIJŲ DALYKO SANDAS

Dalyko pavadinimas	Mokslo kryptis (šaka) kodas	Fakultetas	Katedra
Mašininis mokymasis	Informatikos inžinerija, 07 T	MIF	Duomenų mokslo ir skaitmeninių technologijų institutas
Studijų būdas	Kreditų skaičius ECTS	Studijų būdas	Kreditų skaičius
paskaitos	1 (rudens sem.)	konsultacijos	1
individualus	4	seminarai	1

### Dalyko anotacija

Pageidautina, kad doktorantas pasirenkantis šį dalyką išmanytų tiesinę algebrą, matematinę analizę ir Python programavimo kalbą.

Pagrindinis sando tikslas – įgyti 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, modeliu ir individų panašumu paremtas mokymas). Pagrindiniai iššūkiai mašiniame mokyme. Testavimas ir validavimas.
- Tipinio mašininio mokymosi projekto struktūra ir atlikimas. Problemos formulavimas. Duomenų gavimas. Duomenų vizualizavimas ir pažinimas. Duomenų parengimas mašininiam mokymuisi. Modelio parinkimas ir mokymas. Modelio derinimas.
- Duomenų klasifikavimas. Binarinio klasifikavimo uždavinys. Klasifikatoriaus rezultatų vertinimo metrikos. Paklaidų analizė. Artimiausių kaimynų ir Bajeso klasifikatoriai.
- Regresijos metodai. Tiesinės, polinominės ir logaritminės regresijos modeliai. Gradientinis nusileidimas. Reguliarizuotas tiesinės regresijos modelis. 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.). Atsitiktiniai miškai (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. Autoenkoderų taikymas triukšmo šalinimui duomenyse.
- Pastiprintas mokymasis. Taisyklių paieška. Įvadas į OpenAI Gym. Taisyklių gradientai. Markovo sprendimų priėmimo procesas. Mokymasis iš pavykių laike ir Q-mokymasis.

#### Praktinės užduotys:

- Klasikinių klasifikavimo ir regresijos metodų taikymas. Atsiskaitomas projektinis darbas iš 1–5 temų.
- Specializuotų algoritmų kūrimas. Pagal doktoranto vykdomą mokslinę temą parinkta užduotis, kurios sprendimui būtų naudojami neuroniniai tinklai, autoenkoderiai arba pastiprinto mokymosi metodai.

Modulį sudarys 9 paskaitos, 9 seminarai, du atsiskaitomieji darbai.
Pagrindinė literatūra
Aurélien Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, 2017
Ian Goodfellow and Yoshua Bengio and Aaron Courville. Deep Learning, 2016, MIT press, <a href="http://www.deeplearningbook.org">http://www.deeplearningbook.org</a>
John D. Kelleher, Brian Mac Namee, Aoife D'Arcy. Fundamentals of Machine Learning for Predictive Data Analytics, The MIT Press, 2015
CS229: Machine Learning, <a href="http://cs229.stanford.edu/syllabus.html">http://cs229.stanford.edu/syllabus.html</a>
Machine Learning group forum, <a href="https://www.reddit.com/r/MachineLearning/">https://www.reddit.com/r/MachineLearning/</a> , <a href="https://www.reddit.com/r/MachineLearning/wiki/index">https://www.reddit.com/r/MachineLearning/wiki/index</a>

Konsultuojančiųjų dėstytojų vardas, pavardė	Mokslo laipsnis	Svarbiausieji darbai mokslo kryptyje (šakoje) paskelbti per pastaruosius 5 metus
Virginijus Marcinkevičius	dr.	<p>Medvedev V.; Kurasova O.; Bernatavičienė J.; Treigys P., Marcinkevičius V. Dzemyda G., 2017. A new web-based solution for modelling data mining processes. Simulation Modelling Practice and Theory. ISSN: 1569-190X, Vol. 74, Elsevier.</p> <p>Jakštys, V.; Marcinkevičius, V.; Tichonov, J.; Treigys, P., 2016. Detection of the road pothole contour in raster images // Information technology and control. Kaunas: Technologija. ISSN: 1392-124X. 2016, Vol. 45, no. 3, p. 300-307.</p> <p>Bernatavičienė, J; Dzemyda, G.; Bazilevičius, G.; Medvedev, V.; Marcinkevičius, V.; Treigys, P. 2015. Method for visual detection of similarities in medical streaming data // International journal of computers communications &amp; control. Oradea: Universitatea Agora. ISSN 1841-9836. 2015, Vol. 10, no. 1, p. 8-21.</p>
Jolita Bernatavičienė	dr.	<p>Morkūnas, Mindaugas; Treigys, Povilas; Bernatavičienė, Jolita; Laurinavičius, Arvydas; Korvel, Gražina. Machine learning based classification of colorectal cancer tumour tissue in whole-slide images // Informatica. Vilnius : Vilniaus universiteto Matematikos ir informatikos institutas. ISSN 0868-4952. 2018, Vol. 29, no. 1, p. 75-90. DOI: 10.15388/Informatica.2018.158.</p> <p>Venskus, Julius; Treigys, Povilas; Bernatavičienė, Jolita; Medvedev, Viktor; Voznak, Miroslav; Kurmis, Mindaugas; Bulbenkienė, Violeta. Integration of a self-organizing map and a virtual pheromone for real-time abnormal movement detection in marine traffic // Informatica. Vilnius : Vilniaus universiteto Matematikos ir informatikos institutas. ISSN 0868-4952. 2017, Vol. 28, No. 2, p. 359-374. Prieiga per internetą: &lt;<a href="https://www.mii.lt/informatica/pdf/INFO1145.pdf">https://www.mii.lt/informatica/pdf/INFO1145.pdf</a>&gt;.</p> <p>Jucevičius, Justinas; Treigys, Povilas; Bernatavičienė, Jolita; Briedienė, Rūta; Naruševičiūtė, Ieva; Dzemyda, Gintautas; Medvedev, Viktor. Automated 2D segmentation of prostate in T2-weighted MRI scans // International journal of computers communications &amp; control. Oradea : Universitatea Agora. ISSN 1841-9836. eISSN 1841-9844. 2017, Vol. 12, No. 1, p. 53-60. Prieiga per internetą: &lt;<a href="http://univagora.ro/jour/index.php/ijccc/article/view/2783/1059">http://univagora.ro/jour/index.php/ijccc/article/view/2783/1059</a>&gt;.</p>

		<p>Medvedev V.; Kurasova O.; Bernatavičienė J.; Treigys P., Marcinkevičius V. Dzemyda G., 2017. A new web-based solution for modelling data mining processes. Simulation Modelling Practice and Theory. ISSN: 1569-190X, Vol. 74, Elsevier.</p>
		<p>Bernatavičienė, J; Dzemyda, G.; Bazilevičius, G.; Medvedev, V.; Marcinkevičius, V.; Treigys, P. 2015. Method for visual detection of similarities in medical streaming data // International journal of computers communications &amp; control. Oradea: Universitatea Agora. ISSN 1841-9836. 2015, Vol. 10, no. 1, p. 8-21.</p>
<p>Vytautas Valaitis</p>	<p>dr.</p>	<p>Valaitis, V. et al., 2015 Minimizing hexapod robot foot deviations using multilayer perceptron. *International Journal of Advanced Robotic Systems*. Vol. 12.</p> <p>Valaitis, V. et al., 2015, Piezoelectric force sensors for hexapod transportation platform. Transport: special issue on smart and sustainable transport. Vol. 30, no. 3.</p> <p>Valaitis, V. et al., 2015, A Price we Pay for Inexact Dimensionality Reduction. Bioinformatics and Biomedical Engineering*. p. 289-300.</p>