

DOKTORANTŪROS STUDIJŲ DALYKO SANDAS

| | | | |
|--|-----------------------|----------------------------|--|
| Dalyko pavadinimas | Mokslų kryptis, kodas | Fakultetas | Katedra |
| Netiesiniai statistikos modeliai masinių duomenų analizėje | Informatika (N 009) | MIF | Duomenų mokslų ir skaitmeninių technologijų institutas |
| Studijų būdas | Kreditų skaičius ECTS | Studijų būdas | Kreditų skaičius |
| paskaitos (pavasario sem.) | 1 | konsultacijos | 1 |
| individualus | 4 | seminarai (pavasario sem.) | 1 |

| |
|--|
| <p>Dalyko anotacija</p> <p>Dalyko tikslas – papildyti studentų turimas žinias apie automatinį (mašininį) mokymąsi netiesinio statistinio modeliavimo žiniomis akcentuojant kritinį statistinį mąstymą.</p> <p>Kurso kontekste didieji duomenys (angl. <i>Big data</i>) suprantami, kaip masiškai be konkretaus tikslo surinkti duomenys ir yra vadinami masiniais duomenimis. Pastarieji duomenys, kaip taisyklė, yra heterogeniški nuo paprasčiausių mažų tekstinių įrašų iki akcijų informacijos kas minutę ar viso genomo duomenų. Šiuos duomenis galima analizuoti traktuojant juos kaip juodąją dėžę (angl. <i>Black box</i>) ir naudoti automatinio (mašininio) mokymosi metodus. O galima taikyti statistinį modeliavimą ir, atsižvelgiant į duomenų generavimo procesą, daryti prielaidas apie duomenų skirstinį ir jas testuoti. Kuo konkretnės (griežtesnės) prielaidos dera su duomenimis, tuo turiningesnė ir subtilesnė gautų rezultatų interpretacija ir išvados. Todėl automatinio (mašininio) mokymosi metodus prasminga derinti su netiesinio (parametriniais ir neparametriniais) statistinio modeliavimo metodais.</p> <p>Pagrindinės temos:</p> <ul style="list-style-type: none"> • Klasikiniai ir robastiniai tiesiniai metodai (Agresti) • Apibendrintieji tiesiniai (AT) modeliai, modelio pasirinkimas ir statistinės išvados (Agresti) • Pakartotiniai stebiniai, atsitiktiniai veiksniai ir ilgalaikių stebėjimų duomenys (Faraway) • Mišriųjų veiksnų modeliai, nenormalusis atsako kintamasis (Faraway) • Bajeso daugialygis (hierarchinis) modeliavimas, Markovo grandinių Monte Karlo metodai (Madigan) • AT modelio regularizavimas (Agresti) • Neparametrinė regresija (Faraway), apibendrintieji adityvūs modeliai ir glodinimo metodai (Agresti, Faraway) <p>Studentų gebėjimai ir atsiskaitymas: Kurso metu studentai, išanalizavę skirtą literatūrą, pasirinktiems duomenims realizuoja apibendrintuosius tiesinius ir neparametrinius regresijos modelius. Esant poreikiui, masinių duomenų analizei moka pritaikyti regularizacijos metodus, sudėtingų struktūrų tyrime – Bajeso metodus. Kursą išklauses studentas suvokia ir geba įvertinti duomenų ir jų analizės rezultatų neapibrėžtumą ir patikimumą. Metodai realizuojami R kalba arba kita studentui ir dėstytojui abipusei priimtina programine įranga. Kursas baigiamas pritaikytų metodų aprašymu (idealiu atveju moksline publikacija).</p> <p>Pagrindinė literatūra</p> <p>Agresti, A., 2015. Foundations of linear and generalized linear models. John Wiley & Sons. http://bayanbox.ir/view/7443147326514856944/Foundations-of-Linear-and-Generalized-Linear-Models-Wiley-Series-in-Probability-and-Statistics-Alan-Agresti-2015.pdf</p> <p>Faraway, J.J., 2016. Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models (Vol. 124). CRC press.</p> |
|--|

| |
|---|
| Madigan, D., Ridgeway G., 2002. Bayesian Data Analysis for Data Mining. https://pdfs.semanticscholar.org/58be/87292ce8c4a0eefca6dd5430368f4af4e177.pdf |
| Papildoma literatūra |
| Alpaydin, E., 2010. Introduction to Machine Learning. The MIT Press. ISBN-10: 0-262-01243-X, ISBN-13: 978-0-262-01243-0. http://cs.du.edu/~mitchell/mario_books/Introduction_to_Machine_Learning_-_2e_-_Ethem_Alpaydin.pdf |
| Han, J., Kamber, M., 2006. Data Mining Concepts and Techniques. 2nd ed. CA: Morgan Kaufmann Publishers is an imprint of Elsevier. http://ccs1.hnue.edu.vn/hungtd/DM2012/DataMining_BOOK.pdf https://pdfs.semanticscholar.org/02e0/bc77460469aefec5bd794ee6c4efc15e6adb.pdf |

| Konsultuojančiųjų dėstytojų vardas, pavardė | Mokslo laipsnis | Svarbiausieji darbai mokslo kryptyje, paskelbti per pastaruosius 5 metus |
|---|-----------------|---|
| Audronė Jakaitienė | dr. | http://www.elaba.mb.vu.lt/dmsti/?aut=Audronė+Jakaitienė |
| Marijus Radavičius | dr. | http://www.elaba.mb.vu.lt/dmsti/?aut=Marijus+Radavičius |
| Olga Kurasova | dr. | http://www.elaba.mb.vu.lt/dmsti/?aut=Olga+Kurasova |