

DOKTORANTŪROS STUDIJŲ DALYKO SANDAS

Dalyko pavadinimas	Mokslo kryptis (šaka) kodas	Fakultetas	Institutas, katedra
Daugiamačių duomenų vizualizavimo metodai	Informatika, 09 P	MIF	Duomenų mokslo ir skaitmeninių technologijų institutas

Studijų būdas	Kreditų skaičius ECTS	Studijų būdas	Kreditų skaičius
paskaitos	1 (pavasario sem.)	konsultacijos	1
individualus	4	seminarai	1

Dalyko anotacija

Technikoje, medicinoje, ekonomikoje, ekologijoje ir daugelyje kitų sričių nuolatos susiduriama su daugiamačiais duomenimis. Faktiškai nėra žmonių veiklos srities, kur nebūtų kaupiami ir analizuojami tokie duomenys. O vystantis technologijoms, tobulėjant kompiuteriams ir programinei įrangai kaupiamų duomenų apimtys ypač sparčiai didėja. Auga ir poreikiai bei nauda, gaunama padarius teisingas išvadas. Daugiamačių duomenų vizualizavimas yra svarbus duomenų analizės įrankis, padedantis geriau suvokti daugiamačių duomenų struktūrą – susidariusias grupes (klasterius), labai išsiskiriančius objektus (taškus atsiskyrėlius), objektų tarpusavio artimumą jų visumos kontekste. Doktorantai įgys žinias apie daugiamačių duomenų vizualizavimo metodus bei gebėjimus juos taikyti sprendžiant duomenų analizės uždavinius, naudoti populiarias duomenų vizualizavimo sistemas.

Pagrindinės temos:

- Vizualizavimo tikslai ir jo nauda duomenų analizėje.
- Tiesioginio vizualizavimo metodai: sklaidos diagramų matrica, lygiagrečiosios koordinatės, Andrews kreivės, Černovo veidai ir kiti.
- Projektijos metodai: tiesinės ir netiesinės projektijos metodų skirtumai.
- Pagrindinių komponentų analizė (PCA): pagrindinių komponentų esmė, jų radimo principai, PCA taikymas vizualizuojant daugiamačius duomenis.
- Daugiamatės skalės (MDS): MDS principai, MDS tipai, panašumo matai, gautų rezultatų interpretavimas.
- Minimizavimo algoritmai daugiamatėms skalėms: gradientiniai metodai, šakų režių metodas ir kt.
- Dirbtiniai neuroniniai tinklai duomenims vizualizuoti: daugiasluoksnis perceptronas, autoasociatyvūs neuroniniai tinklai.
- Saviorganizuojantys neuroniniai tinklai: jų principai, mokymo algoritmas, papildomi vizualizavimo būdai, rezultatų interpretavimas.
- Vizualizavimo sistemos: Matlab įrankiai (*Dimensionality Reduction Toolbox*), Orange (<https://orange.biolab.si>), R paketas, Jupyter (<https://nbviewer.jupyter.org/>), DAMIS (<http://damis.lt/>), TABLEAU (<https://www.tableau.com>) ir kt.
- Vizualiosios analizės taikymai

Praktinė užduotis: pasirinkus duomenų aibę, ją vizualizuoti taikant įvairius vizualizavimo metodus; paruošti ataskaitą, kurioje aprašyti rezultatus, padaryti išvadas, ką leido gauti, pastebėti nagrinėjamuose duomenyse vizualizavimas.

Pagrindinė literatūra

Dzemyda, G.; Kurasova, O.; Žilinskas, J. Daugiamačių duomenų vizualizavimo metodai: vadovėlis informatikos krypties doktorantams ir magistrantams. Matematikos ir informatikos institutas. Vilnius:

Mokslo aidai, 2008. 204 p. ISBN 9789986680420.

<http://web.vu.lt/mii/j.zilinskas/DzemydaKurasovaZilinskasDDVM.pdf>

Borg, P. Groenen. Modern Multidimensional Scaling: Theory and Applications. Springer-Verlag, New York, USA. 2005.

Dzemyda, G.; Kurasova, O.; Žilinskas, J. Multidimensional Data Visualization: Methods and Applications. Springer, 2013, ISBN 978-1-4419-0235-1. doi:10.1007/978-1-4419-0236-8.

Han, J.; Pei, J.; Kamber, M. Data mining: concepts and techniques. Elsevier. 2011.

Yuk, M.; Stephanie D. Data visualization for dummies. John Wiley & Sons, 2014.

Konsultuojančiųjų dėstytojų vardas, pavardė	Mokslo laipsnis	Svarbiausieji darbai mokslo kryptyje (šakoje) paskelbti per pastaruosius 5 metus
Gintautas Dzemyda	habil. dr.	Bilinskas, M.J., Dzemyda, G., Trakymas, M. (2018). Approximation of the ribs-bounded contour in a Tomography scan slice. <i>International Journal of Information Technology & Decision Making</i> , 17 (1), 83-102. https://doi.org/10.1142/S0219622017500298
		Gėgžna, V., O. Kurasova, G. Dzemyda, R. Kurtinaitienė, I. Čiplys, J. V. Vaitkus, A. Vaitkuvienė (2018). The ROC-based analysis of spectroscopic signals from medical specimens. <i>Nonlinear Analysis: Modelling and Control</i> , 23 (3), 285-302.
		Evora, J., J. J. Hernandez, M. Hernandez, G. Dzemyda, O. Kurasova, E. Kremers (2015). Swarm intelligence for frequency management in smart grids. <i>Informatica</i> , 26 (3), 419-434.
Olga Kurasova	dr.	Medvedev, V.; Kurasova, O.; Bernatavičienė, J.; Treigys, P.; Marcinkevičius, V.; Dzemyda, G. A new web-based solution for modelling data mining processes. <i>Simulation Modelling Practice and Theory</i> . Amsterdam : Elsevier Science. ISSN 1569-190X. eISSN 1878-1462. 2017, Vol. 76, p. 34-46. DOI: 10.1016/j.simpat.2017.03.001.
		Bernatavičienė, J.; Dzemyda, G.; Kurasova, O.; Marcinkevičius, V.; Medvedev, V.; Treigys, P. Cloud computing approach for intelligent visualization of multidimensional data. <i>Advances in stochastic and deterministic global optimization</i> . Ser.: Springer optimization and its applications. ISSN 1931-6828. Vol. 107. Cham: Springer International Publishing Switzerland, 2016. ISBN 9783319299730. p. 73-85. DOI: 10.1007/978-3-319-29975-4_5.
		Orts, F.; Filatovas, E.; Ortega, G.; Kurasova, O.; Garzon, E. M. Improving the energy efficiency of SMACOF for multidimensional scaling on modern architectures. <i>Journal of Supercomputing</i> . Dordrecht: Springer. ISSN 0920-8542. eISSN 1573-0484. 2018, First published online, [13 p.]. DOI: 10.1007/s11227-018-2285-x.
Julius Žilinskas	(HP) dr.	Dzemyda, G.; Kurasova, O.; Žilinskas, J. Multidimensional Data Visualization: Methods and Applications. Springer Optimization and Its Applications, ISSN 1931-6828, Vol. 75, Springer, ISBN 978-1-4419-0235-1. 2013, doi:10.1007/978-1-4419-0236-8.

		<p>Varoneckas, A.; Žilinskas, A.; Žilinskas, J. Multi-objective optimization aided to allocation of vertices in aesthetic drawings of special graphs. <i>Nonlinear Analysis: Modelling and Control</i>, ISSN 1392-5113, 18(4), 476-492, 2013.</p>
		<p>Galiauskas, N.; Žilinskas, J. On multidimensional scaling with city-block distances. <i>Lecture Notes in Computer Science</i>, 2014, 8426, 82-87. ISSN 0302-9743. doi:10.1007/978-3-319-09584-4_9</p>
		<p>Galiauskas, N.; Žilinskas, J. Parallel branch and bound for multidimensional scaling with L1 distances formulated as quadratic programming with complementarity constraints. In: Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2013, ISBN 978-0-7695-5094-7, pp. 509-512. doi: 10.1109/3PGCIC.2013.87</p>