STUDIJŲ DALYKO (MODULIO) APRAŠAS

Dalyko (modulio) pavadinimas	Kodas
Papildomi duomenų vizualizavimo skyriai	

Dėstytojas (-ai)	Padalinys (-iai)
Koordinuojantis: dr. Jolita Bernatavičienė **Kitas (-i):** dr. Viktor Medvedev	Matematikos ir informatikos fakultetas Duomenų mokslo ir skaitmeninių technologijų institutas

Studijų pakopa	Dalyko (modulio) tipas
Pirmoji	Pasirenkamas

Įgyvendinimo forma	Vykdymo laikotarpis	Vykdymo kalba (-os)
Auditorinė	6 (pavasario) semestras	Lietuvių

Reikalavimai studijuojančiajam	
Išankstiniai reikalavimai: nėra	**Gretutiniai reikalavimai (jei yra):** nėra

Dalyko (modulio) apimtis kreditais	Visas studento darbo krūvis	Kontaktinio darbo valandos	Savarankiško darbo valandos
5	125	48	77

Dalyko (modulio) tikslas: studijų programos ugdomos kompetencijos

Modulio tikslas – suteikti studentams papildomų žinių apie duomenų vizualizavimą, supažindinti su daugiamačių duomenų vizualizavimo galimybėmis, ugdyti praktinius gebėjimus taikyti duomenų vizualizavimo metodus, ugdyti gebėjimą vizualiai tirti ir interpretuoti duomenis.
Dalyko ugdomos studijų programos kompetencijos:

- Gebėjimas analizuoti, sisteminti, mokytis ir taikyti įgytas žinias praktikoje;
- Gebėjimas bendrauti ir bendradarbiauti valstybine bei užsienio kalba;
- Gebėjimas interpretuoti ir reprezentuoti daugiamačių duomenų analizės rezultatus.

Dalyko (modulio) studijų siekiniai	Studijų metodai	Vertinimo metodai
• gebės rinkti, analizuoti ir sisteminti informaciją kūrybiškai panaudodamas meta analizės rezultatus uždavinių sprendimui; • gebės bendrauti su įvairių sričių specialistais, pateikti ir interpretuoti gautus rezultatus tiek specialistui, tiek nespecialistui priimtina kalba; • gebės interpretuoti analizės rezultatus, išskirti prasmingą informaciją bei remiantis ja teikti siūlymus; • gebės savarankiškai rengti nedidelės apimties duomenų analize grindžiamus projektus, kurti nedidelius duomenų analizės ataskaitų įrankius.	Probleminis dėstymas, pavyzdžių nagrinėjimas, informacijos paieška, dalykinės literatūros studijavimas, laboratoriniai darbai	Egzaminas, laboratorinių darbų ataskaitų darbų gynimo vertinimas, darbas auditorijoje (diskusija), aktyvumas per užsiėmimus.

Temos	Kontaktinio darbo valandos							Savarankiškų studijų laikas ir užduotys	
	Paskaitos	Konsultacijos	Seminarai	Pratybos	Laboratoriniai darbai	Praktika	Visas kontaktinis darbas	Savarankiškas darbas	Užduotys
1. Daugiamačių duomenų vizualizavimo strategijos	**2**						**2**	**7**	Mokslinės literatūros skaitymas ir analizė, laboratorinių darbų atlikimas, ataskaitų rengimas, pasiruošimas laboratorinių darbų gynimui.
2. Daugiamačių duomenų dimensijos mažinimas	**2**				2		**4**	**10**	
3. Pagrindinės duomenų statistinės charakteristikos, pirminio duomenų apdorojimo metodai	**2**				2		**4**	**10**	
4. Tiesioginio duomenų vizualizavimo metodai	**4**				4		**8**	**10**	
5. Daugiamačių duomenų dimensijos mažinimo metodai: pagrindinių komponenčių analizė (PCA), daugiamatės skalės (MDS), saviorganizuojantys neuroniniai tinklai (SOM), lokaliai tiesinis vaizdavimas (LLE), t-SNE metodas ir kt.)	**6**				6		**12**	**20**	
6. Daugiamačių duomenų vizualizavimas klasterizavimo uždaviniuose	**4**				6		**10**	**10**	
7. Daugiamačių duomenų vizualizavimas klasifikavimo uždaviniuose	**4**				4		**8**	**10**	
Iš viso	**24**				**24**		**48**	**77**	

Vertinimo strategija	Svoris proc.	Atsiskaitymo laikas	Vertinimo kriterijai
Laboratoriniai darbai	40	Semestro metu	Studentai privalo atlikti 4 praktines užduotis, paruošti rašto darbus (ataskaitas) ir jas individualiai apginti atsakant į užduodamus klausimas.

		(pagal semestro pradžioje pateiktą atsiskaitymų tvarkaraštį)	• 9–10 balai – darbas tenkina visus būtinus reikalavimus: rašto darbo struktūra aiški ir logiška, yra visos reikiamos dalys (darbo tikslas, uždaviniai, eksperimentų vykdymo aprašas, rezultatai, išvados), darbas tinkamai suformatuotas, rezultatų analizė išsami, išvados pagrįstos, studentas atsako į užduotus klausimus; • 7–8 balai – yra visos reikiamos rašto darbo dalys (darbo tikslas, uždaviniai, eksperimentų vykdymo aprašas, rezultatai, išvados), tačiau yra trūkumų: ne visos darbo dalys tinkamai suformatuotos, yra loginių klaidų, rezultatų analizė nepakankamai išsami, ne visos išvados yra pagrįstos, studentas atsako ne į visus užduotus klausimus. • 5–6 balai – yra ne visos būtinos rašto darbo dalys, analizė yra paviršutiniška ir fragmentiška, ne visos išvados pagrįstos, yra esminių loginių klaidų, ne į visus esminius klausimus studentas geba atsakyti. • 1–4 balai – rašto darbas blogai struktūrizuotas, turi labai daug esminių klaidų, studentas neatsako į klausimus. • 0 balų – rašto darbas nėra pateiktas arba nėra apgintas. Yra laikoma, kad studentas yra atsiskaitęs už darbą, kai rašto darbas įkeltas į sistemą ir apgintas. Pavėlavus 2 savaites atsiskaityti, įvertinimas mažinamas dviem balais, 3 savaites – trimis balais. Vėliau darbas nebus vertinamas. Praktines užduotis būtina atsiskaityti ne vėliau kaip paskutinę praktinių užsiėmimų paskaitą
Grupiniai dimensijos mažinimo algoritmų pristatymai teorinių paskaitų metu	20	Semestro metu (pagal semestro pateiktą atsiskaitymų tvarkaraštį)	Vertinimas už algoritmų pristatymą 2 balai (vertinamas teorinis algoritmo pristaymas ir eksperimentinis tyrimas, atsakymai į klausimus)
Egzaminas raštu	40	Egzaminų sesijos metu	Dalyvavimas egzamine privalomas. Egzamino metu studentai atsako į 5 pateiktus klausimus. Dalyko egzamino laikymas eksternu negalimas.

Autorius	**Leidimo metai**	**Pavadinimas**	**Periodinio leidinio Nr. ar leidinio tomas**	**Leidimo vieta ir leidykla ar internetinė nuoroda**
Privaloma literatūra				
Jolita Bernatavičienė		Papildomi duomenų vizualizavimo skyriai: paskaitų konspektai		VMA
M.J.Zaki, W.Meira Jr.	2020	Data mining and analysis. fundamental concepts and algorithms, Second Edition		Cambridge University Press https://dataminingbook.info/book_html/
G.Dzemyda, O.Kurasova, J.Žilinskas	2008	Daugiamačių duomenų vizualizavimo metodai		Matematikos ir informatikos institutas
Papildoma literatūra				
G.Dzemyda, O.Kurasova, J. Žilinskas	2013	Multidimensional Data Visualization: Methods and Applications	Springer Optimization and Its Applications, Vol. 75	Springer https://www.researchgate.net/publication/266046175_Multidimensional_data_visualization_Methods_and_applications
H. Whitney	2013	Data insights: new ways to visualize and make sense of data	Elsevier/ Morgan Kaufman	VU library

COURSE UNIT (MODULE) DESCRIPTION

Course unit (module) title	Code
Supplementary Chapters in Data Visualization	

Lecturer(s)	Department(s) where the course unit (module) is delivered
Coordinator: Dr. Jolita Bernatavičienė **Other(s):Dr. Viktor Medvedev**	Institute of Data Science and Digital Technologies Faculty of Mathematics and Informatics

Study cycle	Type of the course unit (module)
1st (BA)	Compulsory

Mode of delivery	Period when the course unit (module) is delivered	Language(s) of instruction
Face-to-face	6th semester	Lithuanian

Requirements for students	
Prerequisites:	**Additional requirements (if any):**

Course (module) volume in credits	Total student's workload	Contact hours	Self-study hours
5	125	48	77

Purpose of the course unit (module): programme competences to be developed

The aim of the module is to provide students with additional knowledge of data visualisation, to introduce the possibilities of visualising multidimensional data, to develop practical skills in applying data visualisation techniques, and to develop the ability to visually explore and interpret data.
Purpose of the course unit:

- Ability to collect data from various data sources, to process and analyse the information;
- Ability to find the appropriate method for data analysis in solving various practical problems;
- Ability to interpret and represent the results of multidimensional data analysis.

Learning outcomes of the course unit (module)	Teaching and learning methods	Assessment methods
• Ability to select and apply various data visualization solutions, plan data collection, pre-processing activities to solve data mining task. • Apply methods of direct visualization, understand and interpret data visualized using these methods; • Perform linear and nonlinear projections of data, interpret results and apply them for data visualization; • Identify situations when data visualization helps perform data analysis more efficiently. • Ability to conduct a proper literature search and analysis with the view to depict the appropriate multidimensional data structures, to gather new knowledge and	Case studies, analysis of scientific literature, individual work, laboratory tasks, problem-oriented teaching	Exam, evaluation results of laboratory tasks, discussion

apply it for data visualization. • Ability to investigate the data dimensionality influence to data classification, clustering, interpret results. • Ability to apply existing data mining, machine learning tools for multidimensional data analysis, to perform statistical analysis, to solve data mining and data visualization tasks		

Content: breakdown of the topics	**Contact hours**							**Self-study work: time and assignments**	
	Lectures	Tutorials	Seminars	Exercises	Laboratory work	Internship/work placement	**Contact hours**	**Self-study hours**	**Assignments**
1. Strategies for multidimensional data vizualization	2						**2**	**7**	Work with scientific literature and its analysis, preparation of laboratory works and their reports, preparation for the defence of laboratory work.
2. Dimensionality reduction of multidimensional data.	2				2		**4**	**10**	
3. Basic statistical characteristics of the data, data pre-processing methods.	2				2		**4**	**10**	
4. Direct visualization methods.	4				4		**8**	**10**	
5. Dimensionality reduction methods: principal component analysis (PCA), multidimensional scaling (MDS), self-organizing maps (SOM), locally linear embedding (LLE), t-SNE	6				6		**12**	**20**	
6. Vizualization of multidimensional data in clustering tasks	4				6		**10**	**10**	
7. Vizualization of multidimensional data in classification tasks	4				4		**8**	**10**	
Total	**24**				**24**		**48**	**77**	

Assessment strategy	**Weight,%**	**Deadline**	**Assessment criteria**
Laboratory works	40%	During the semester (according to the timetable given at the beginning of the semester)	The students must perform 4 practical tasks, present the reports, and defend them individually by answering the questions asked. • 9-10 points – the report satisfies all the necessary requirements: the report structure is clear and logical, there are all the necessary parts (work aim, tasks, description of experiments, results, conclusions), the report is properly formatted, the analysis of the results is comprehensive, the conclusions are reasoned, the student answers all questions asked. • 7-8 points – all the required parts of the report are presented (work aim, tasks, description of experiments, results, conclusions), but there are shortcomings: not all parts of the report are properly formatted, there are logical errors, the analysis of the results is not sufficiently detailed, not all the conclusions are justified, and the student does not answer all the questions asked. • 5-6 points – not all the necessary parts of the report are presented, the analysis is superficial and fragmentary, not all the conclusions are justified, there are fundamental logical errors, the student is not able to answer all the essential questions. • 1-4 points – the report is poorly structured, it has a large number of fundamental errors, the student does not answer the questions. • 0 point – the report is not presented, or it is not defended. A student is considered to have completed the practical task when the report is uploaded to the system and defended. If the delay is 2 weeks, the grade is reduced by 2 points; if the delay is 3 weeks, the grade is reduced by 2 points. Thereafter the practical task will not be evaluated. Practical tasks must be submitted no later than the last of the practical lecture.
Group presentations of dimension	20%	During the	Presentation of theoretical and experimental investigation of the algorithm,

reduction techniques during theoretical lectures		semester (according to the timetable provided by the lecturer)	answers to questions are evaluated maximum up to 2 points.
Written exam	40%	During exam session	Participation in the written exam is mandatory In the exam, students answer 5 questions. External examination is not allowed.

Author	Year of publication	Title	Issue of a periodical or volume of a publication	Publishing place and house or web link
Compulsary reading				
1. Jolita Bernatavičienė		Supplementary Chapters in Data Visualization (a lecture notes)		VLE
2. M.J.Zaki, W.Meira Jr.	2020	Data mining and analysis. fundamental concepts and algorithms, Second Edition		Cambridge University Press https://dataminingbook.info/book_html/
Optional reading				
1. G. Dzemyda, O. Kurasova, J. Žilinskas	2013	Multidimensional Data Visualization	Springer	VU MIF library
2. H. Whitney	2013	Data insights: new ways to visualize and make sense of data	Elsevier/ Morgan Kaufman	VU library