



STUDIJŲ DALYKO (MODULIO) APRAŠAS

Dalyko (modulio) pavadinimas	Kodas
DUOMENŲ GAVYBOS TECHNOLOGIJOS	

Dėstytojas / a (-ai)	Padalinys (-iai)
Koordinuojantis (-i): jaun. asist. Paulius Danielius	Kauno fakulteto Kalbų, literatūros ir vertimo studijų institutas <input type="checkbox"/> Socialinių mokslų ir taikomosios informatikos institutas <input checked="" type="checkbox"/>

Studijų pakopa	Dalyko (modulio) tipas
Pirmoji <input checked="" type="checkbox"/> Antroji <input type="checkbox"/>	Privalomasis dalykas <input checked="" type="checkbox"/> Pasirenkamasis dalykas <input type="checkbox"/> Bendrauniversitečių studijų dalykas <input type="checkbox"/> Individualiųjų studijų dalykas <input checked="" type="checkbox"/> Tarpkryptinis dalykas <input type="checkbox"/>

Igyvendinimo forma	Vykdyimo laikotarpis	Vykdyimo kalba (-os)
Auditorinė	1 semestras	Lietuvių

Reikalavimai studijuojančiajam	
Išankstiniai reikalavimai: Aukštoji matematika, informatika	Gretutiniai reikalavimai (jei yra):

Dalyko (modulio) apimtis kreditais	Visas studento darbo krūvis	Kontaktinio darbo valandos	Savarankiško darbo valandos
5	130	52	78

Dalyko (modulio) tikslas		
Ugdyti gebėjimą dirbti su duomenų gavybos ir duomenų analizės technologijomis. Pateikti studentams duomenų gavybos technologijų sampratą, supažindinti su duomenų analizės technologijomis, kryptimis, metodais. Ugdyti gebėjimą taikyti praktikoje pagrindinius duomenų analizės įrankius ir juose naudojamus metodus, algoritmus. Ugdyti gebėjimą apdoroti, interpretuoti ir analizuoti duomenų gavybos rezultatus.		
Dalyko (modulio) studijų rezultatai	Studijų metodai	Vertinimo metodai
Geba efektyviai bendrauti valstybine ir / arba anglų k. raštu ir žodžiu, tinkamai vartodamas dalykinę terminiją, pristatydamas siūlomas idėjas ir pagrįsdamas inovatyvius problemų sprendimo būdus tiek profesionalų, tiek IT paslaugų naudotojų auditorijai.	Aktyvaus mokymosi metodai (individualaus darbo pristatymas ir diskusija grupėje).	Individualaus rašto darbo pristatymas klasei.
Geba taikyti informatikos inžinerijos žinias, kuriant saugius ir kitus aktualius kriterijus atitinkančius informatikos taikomuosius sprendimus aktualioms visuomenei problemoms spręsti.	Paskaitos, atvejų analizė, praktiniai darbai.	Praktinių darbų ataskaitų gynimas. Kolokviumas. Egzaminas.
Geba atlikti kvalifikuotą informatikos inžinerijos dalykinės srities problemas analizę, naudojant aktualius ir tinkamai parinktus empirinius duomenis, laikydamasis teorinių priėgų, taikydamas efektyvius metodus.	Paskaitos, praktiniai darbai, aktyvaus mokymosi metodai (grupės diskusija, metodų lyginamoji analizė).	

Geba sisteminti, apibendrinti, interpretuoti ir kritiškai įvertinti tyrimo rezultatus, suformuluoti tyrimo išvadas ir pateikti praktinio pritaikymo rekomendacijas.	Praktiniai darbai, atvejų analizė.	
---------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------	--

Temos	Kontaktinio darbo valandos							Savarankiškų studijų laikas ir užduotys	
	Paskaitos	Konsultacijos	Seminarai	Pratybos	Laboratoriniai darbai	Praktika	Visas kontaktinis darbas	Savarankiškas darbas	Savarankiškai atliekamos užduotys
1. Įvadas į duomenų gavybą.									Praktinės užduotys.
Duomenų gavybos tikslas, uždaviniai, apimtis, taikymo sritys ir iššūkiai. Šiuolaikiniai duomenų šaltiniai. Žinių atradimas – nuo duomenų iki sprendimų.	2						2	6	Individuali užduotis – rašto darbas (mokslinė apžvalga pasirenkama tema iš pateikto sąrašo – duomenų gavybos taikymai skirtingose dalykinėse srityse), mokslinės literatūros analizė ir apibendrinimas. Rašto darbo pristatymas klasėje. Pasiruošimas koliokviumui.
Duomenų gavybos gyvavimo ciklas ir procesų modeliai. CRISP-DM ir kiti procesai.	2			2			4	8	
Duomenų prigimtis. Duomenų paruošimas ir pirminis apdorojimas.	2						2	6	
Verslo analitikos sistemos.	1						1		
Duomenų gavybos įrankiai. Programinės įrangos paketas <i>Rapid Miner (AI Studio)</i> . Duomenų vizualizavimas.	1			2			3	4	
2. Duomenų saugyklos.									
Duomenų sandėliavimo sistemos, technologijos ir architektūros.	2						2	4	
Kolokviumas							2		
3. Duomenų gavybos algoritmai ir metodai.									Praktinės užduotys.
Sprendimų medžiai.	2				4		6	6	
Asociacijų taisyklės.	2				4		6	6	
Teksto ir žiniatinklio duomenų analizė ir gavyba.	2				4		6	6	
Neuroniniai tinklai.	2				4		6	6	
Kiti duomenų gavybos algoritmai.	2			2	4		8	6	
<i>Python</i> programavimo kalbos priemonės duomenų gavybai.				2			2		
Konsultacija		2					2		
Egzaminas								20	Pasiruošimas egzaminui.

Iš viso	20	2	4	24	52	78
Pastaba: ne daugiau kaip 4 kontaktinio darbo val. gali būti pakeičiamos socialinių partnerių kviestinėmis paskaitomis arba edukacinėmis išvykomis pas socialinius partnerius.						

Vertinimo strategija	Svoris proc.	Atsiskaitymo laikas	Vertinimo kriterijai
Praktinės užduotys	40	2 – 16-a savaitės	Praktinės užduotys atliekamos viso semestro metu, darbų ataskaitos turi būti įkeliamos į VMA atitinkamai pateiktiems reikalavimams ir atsiskaitymų grafikui. Kiekviena užduotis vertinama 10 balų skalėje pagal VU vertinimo kriterijus. Praktinių užduočių įvertinimą (bendrą balą) sudaro visų užduočių įvertinimų vidurkis.
Rašto darbas su pristatymu	10	6-a savaitė	Rašto darbas pasirinkta tema iš pateikto sąrašo (duomenų gavybos taikymas skirtingose srityse). Tema turi būti suderinta su dėstytoju. Darbas turi būti pristatomas klasėje su pateiktimi pristatymams skirtu užsiėmimų laiku. Vertinimas: rašto darbas 50% + pateiktis ir pristatymas 50%. Vėluojantys pristatymai nepriimami ir neužskaitomi, jeigu tai nėra iš anksto suderinta su dėstytoju.
Kolokviumas	20	9-a savaitė	Kolokviumas apima išdėstytą teorinę ir praktinę medžiagą, vertinimas 10 balų skalėje pagal VU vertinimo kriterijus.
Egzaminas	30	Per sesiją	Egzaminas apima išdėstytą teorinę ir praktinę medžiagą, vertinimas 10 balų skalėje pagal VU vertinimo kriterijus.

DĖL DALYKO LAIKYMO EKSTERNU

Pažymėti <input checked="" type="checkbox"/>				Jei leidžiama, pateikti sąlygas, kt.
Neleidžiama	<input type="checkbox"/>	Leidžiama	<input checked="" type="checkbox"/>	Atsiskaitant dalyką eksternu, praktiniai darbai ir individualus rašto darbas su pateiktimi dėstytojui pateikiami el. paštu (arba įkeliami į e-mokymų aplinką) ne vėliau kaip prieš 5 d. d. iki egzamino. Vertinimui taikomi tokie patys kriterijai, kaip ir studijuojant dalyką nuosekliai.

DĖL GENERATYVINIO DIRBTINIO INTELEKTO (GDI) ĮRANKIŲ („CHATGPT“ AR KT.) NAUDOJIMO STUDIJUOJANT DALYKĄ:

Pažymėti <input checked="" type="checkbox"/>				Jei leidžiama, pateikti sąlygas, kt.
Neleidžiama	<input type="checkbox"/>	Leidžiama	<input checked="" type="checkbox"/>	<p>GDI galima naudoti aiškinantis metodu, algoritmų bei technikų principus, programų kodo pavyzdžius, idėjų generavimui, eksperimentuojant. Papildomi nurodymai, jeigu taikomi, pateikiami kiekvienos praktinės užduoties aprašyme.</p> <p>GDI negalima naudoti:</p> <ul style="list-style-type: none"> • Rašto darbuose ir pateiktyse teksto pateikimui <i>copy-paste</i> principu be atitinkamo citavimo, • Laikant įvertinimo testus semestro ir egzamino metu. <p>Jeigu GDI naudojamas ruošiant rašto darbą, studentas privalo aprašyti:</p> <ul style="list-style-type: none"> • GDI priemonių naudojimo strategiją, • kokie klausimai buvo užduodami, • koks gautas rezultatas ir kiek procentų gauto rezultato modifikuota ir pritaikyta rengiant darbą. <p><i>(Pristatant darbą pateiktis yra pagalbinė priemonė, todėl studentai, siekiantys maksimalaus balo, turi kokybiškai pristatyti nurodytą temą – skaitymas nuo skaidrių vertinamas minimaliu balu).</i></p>

			Visais atvejais studentai privalo laikytis rekomendacijų ir reikalavimų, pateiktų <i>Dirbtinio intelekto naudojimo Vilniaus universitete gairėse</i>
--	--	--	------------------------------------------------------------------------------------------------------------------------------------------------------

DĖL STUDIJŲ REZULTATŲ PASIEKIMO PAŽANGOS

Studentui, (1) nuosekliai semestro laikotarpiu per praktinius užsiėmimus (seminarus, pratybas, kt.) nedemonstruojančiam numatytų dalyko (modulio) studijų rezultatų pasiekimo pažangos ir (2) neįvykdžiusiam visų tarpinių atsiskaitymų reikalavimų ir užduočių dalyko apraše numatytu laiku, **neleidžiama** dalyvauti egzaminų sesijoje.

Autorius (-iai)	Leidimo metai	Pavadinimas	Periodinio leidinio Nr. ar leidinio tomas	Leidykla ar internetinė nuoroda
Privaloma literatūra				
P. Danielius	2025	Moodle aplinka		https://emokymai.vu.lt
Mohammed J. Zaki, Wagner Meira	2014	Data mining and analysis: fundamental concepts and algorithms		Cambridge University Press
Ian H. Witten, Eibe Frank, Mark A. Hall	2011	Data mining: practical machine learning tools and techniques		Morgan Kaufmann
G. Dzemyda, O. Kurasova, J. Žilinskas	2013	Multidimensional Data Visualization Methods and Applications		Springer
Papildoma literatūra				
Andreas C. Müller, Sarah Guido	2018	Introduction to machine learning with Python: a guide for data scientists		O'Reilly
Gabe Ignatow, Rada Mihalcea	2018	An introduction to text mining: research design, data collection, and analysis		SAGE