



STUDIJŲ DALYKO (MODULIO) APRAŠAS

Dalyko (modulio) pavadinimas	Kodas
Duomenų analizė Python aplinkoje	

Dėstytojas (-ai)	Padalinys (-iai)
Koordinuojantis: dr. Donatas Narbutis	Fizikos fakultetas

Studijų pakopa	Dalyko (modulio) tipas
Pirmoji	Pasirenkamasis

Igyvendinimo forma	Vykdyto laikotarpis	Vykdyto kalba (-os)
Auditorinė	2 semestras	Lietuvių, anglų

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	140	64	76

Dalyko (modulio) tikslas: studijų programos ugdomos kompetencijos
Suteikti žinių ir gebėjimų, reikalingų surinkti eksperimentų ir stebėjimų duomenis skaitmeniniu formatu, kurti skaitmeninį eksperimentą, duomenis apdoroti, pavaizduoti, parinkti metodus statistinei analizei ir ją atlikti, įvertinti rezultatų patikimumą, suformuluoti ir pagrįsti analizės išvadas.

Dalyko (modulio) studijų siekiniai	Studijų metodai	Vertinimo metodai
Gebės parengti analizei įvairiais formatais pateikiamus duomenis.	Paskaitos, laboratoriniai darbai, savarankiškas darbas.	Laboratoriniai darbai, egzaminas.
Gebės parinkti duomenų vizualizavimo būdą ir paaiškinti vizualizavimo klaidas.	Paskaitos, laboratoriniai darbai, savarankiškas darbas.	Laboratoriniai darbai, egzaminas.
Gebės parinkti metodus statistinei analizei.	Paskaitos, laboratoriniai darbai, savarankiškas darbas.	Laboratoriniai darbai, egzaminas.
Gebės įvertinti analizės rezultatų patikimumą, suformuluoti ir pagrįsti analizės išvadas.	Paskaitos, laboratoriniai darbai, savarankiškas darbas.	Laboratoriniai darbai, egzaminas.
Gebės sukurti supaprastintą skaitmeninį eksperimentą ir remiantis juo interpretuoti realaus eksperimento duomenis bei paaiškinti modelio parametrų fizikinę prasmę.	Paskaitos, laboratoriniai darbai, savarankiškas darbas.	Laboratoriniai darbai, egzaminas.

Temos	Kontaktinio darbo valandos	Savarankiškų studijų laikas ir užduotys
-------	----------------------------	---

	Paskaitos	Konsultacijos	Seminariai	Pratybos	Laboratoriniai darbai	Praktika	Visas kontaktinis darbas	Savarankiškas darbas	Užduotys
1. Python pagrindai. Anaconda distribucija. Jupyter aplinka. Git kodo versijavimas. Programavimo pradmenys duomenų analizei. Numpy ir matplotlib bibliotekos.	3				3			7	Literatūra, užduotys.
2. Skaitmeniniai duomenų formatai. Įvairių formatų failų nuskaitymas ir įrašymas. SQL duomenų bazių prieiga. Pandas biblioteka.	2				2			6	Literatūra, užduotys.
3. Duomenų paruošimas. Klaidos duomenyse, filtravimas, rūšiavimas. Elementarūs statistiniai skaičiavimai (vidurkis, mediana, moda, standartinis nuokrypis). Duomenų interpoliavimas ir ekstrapoliavimas.	3				3			7	Literatūra, užduotys.
4. Duomenų vizualizavimas. Plotly biblioteka. Histogramos, dvimatės diagramos. Blogos ir geros vizualizavimo praktikos. Daugiamačiai duomenys.	3				3			7	Literatūra, užduotys.
5. Skaitmeninis eksperimentas. Tiesiniai sąryšiai. Atsitiktiniai skaičiai ir jų generavimas. Gauso skirstinys, savybės ir taikymas. Matavimo paklaidų modeliavimas.	3				3			7	Literatūra, užduotys.
6. Modelio parametrų nustatymas. Atvirktinis uždavinys tiesiniam sąryšiui. Mažiausių kvadratų metodas. Matavimo skaičiaus ir paklaidų įtaka parametrų patikimumui.	3				3			7	Literatūra, užduotys.
7. Skaitmeninės nuotraukos. Nuotraukų registravimas, vaizdo elementai. Puasono skirstinys, jo savybės ir taikymas. Vaizdo distorsijos, kameros jautrumas. Vaizdų sutapatinimas.	3				3			7	Literatūra, užduotys.
8. Signalo aptikimas. Laikiniai sąryšiai, filtravimas, konvoliucija, koreliacija. Erdvinis filtravimas, objektų aptikimas nuotraukose.	3				3			7	Literatūra, užduotys.
9. Netiesiniai sąryšiai. Optimizavimo paketai ir metodai (Levenberg-Marquard, Markov Chain Monte Carlo, genetinis). Modelio suformulavimas. Modelio atitikimo duomenims gerumo kriterijus. Lokalūs ir globalūs minimumai. Modelio sudėtingumas ir parinkimo principai.	3				3			7	Literatūra, užduotys.
10. Duomenų klasifikavimas. Scikit-learn biblioteka. Mokymo ir testavimo duomenys. Sprendimų medžio klasifikatorius. Dimensijų sumažinimas, tiesinė ir logistinė regresija.	3				3			7	Literatūra, užduotys.

Klasterizacija. Modelių parinkimas.								
11. Vaizdų klasifikavimas. Keras biblioteka. Neuroninio tinklo konstrukcija ir pritaikymas vaizdų analizei. Duomenų parengimas. Mokymo procesas ir kontrolė. Veikimo patikimumas.	3				3			7
Iš viso	32				32			76

Vertinimo strategija	Svoris proc.	Atsiskaitymo laikas	Vertinimo kriterijai
Laboratoriniai darbai	50	1–15 semestro savaitės	<p>Studentai turi atlikti ir atsiskaityti 8 laboratorinio darbo projektus per semestrą. Laboratorinio darbo metu suformuluojamas projekto tikslas ir uždaviniai, išsiaiškinamos jų programinio įgyvendinimo galimybės. Darbas pabaigiamas savarankiškai; vienam laboratorinio darbo projektui skiriama 4 val. laboratorinio ir 9.5 val. savarankiško darbo laiko. Atsiskaityti už darbus reikia iki semestro pabaigos. Atsiskaitant pateikiamas „Jupyter notebook“ formate aprašymas su programiniu kodu, komentarais ir iliustracijomis. Darbų vertinimo kriterijai:</p> <ul style="list-style-type: none"> - Pasiektas tikslas, įgyvendinti uždaviniai, tvarkingas aprašas, programinis kodas ir iliustracijos, pagrįstos išvados – 10. - Pasiektas tikslas, tačiau yra bent viena esminė ar keletas neesminių pastabų dėl uždavinių įgyvendinimo, aprašo, programinio kodo, iliustracijų ar išvadų pagrįstumo – 9. - Iš dalies pasiektas tikslas, yra dvi esminės pastabos, kurias reikėtų taisyti – 8. - Iš dalies pasiektas tikslas, tačiau daugiau nei pusės uždavinių įgyvendinimą reikėtų taisyti – 7. - Atlikti bent du uždaviniai – 6. - Atliktas bent vienas uždavinys – 5. - Pateiktas „Jupyter notebook“ formate suformuluotos užduoties aprašymas su referato pobūdžio pasvarstymais apie jos atlikimui galimai panaudotinus metodus ir jų patikimumą – 4-1. - Niekas nepateikta – 0. <p>Suskaičiuojamas 8 laboratorinių projektinių darbų įvertinimo vidurkis (L).</p>
Egzaminas	50	16-a semestro savaitė	<p>Semestro metu rengiama kūrybinė tiriamoji užduotis, kurios pagrindu gali būti apjungti visi ar bent keletas laboratorinių darbų. Kūrybinės užduoties tema suderinama iki semestro vidurio. Užduoties aprašas atsiunčiamas iki egzamino PDF formatu. Egzamino kūrybinės užduoties aprašo vertinimui galioja tokie patys kriterijai kaip ir laboratoriniams darbams, kurie pateikti aukščiau. Egzamino metu apginami gauti rezultatai atsakant į 5 klausimus. Egzamino maksimalų vertinimą (E) sudaro kūrybinės užduoties aprašo vertinimas kai išsamiai atsakoma į bent 4 iš 5 klausimų ir atitinkamai mažesnis įvertinimas, jei atsakoma į mažiau klausimų. Jei neatsakoma į nei vieną klausimą, egzamino įvertinimas yra lygus kūrybinės užduoties aprašo įvertinimui minus 4 balai.</p> <p>Galutinis įvertinimas $P = 0.5 L + 0.5 E$.</p>

Autorius	Leidimo metai	Pavadinimas	Periodinio leidinio Nr. ar leidinio tomas	Leidimo vieta ir leidykla ar internetinė nuoroda
Privaloma literatūra				

Jake VanderPlas	2016	Python Data Science Handbook		Oreilly https://github.com/jakevdp/PythonDataScienceHandbook
Wes McKinney	2017	Python for Data Analysis	2nd Edition	Oreilly
Papildoma literatūra				
Scikit-learn developers	2018	Scikit-learn user guide		http://scikit-learn.org/dev/_downloads/scikit-learn-docs.pdf