



## DALYKO (MODULIO) APRAŠAS

| Dalyko (modulio) pavadinimas         | Kodas |
|--------------------------------------|-------|
| <b>DUOMENŲ GAVYBOS TECHNOLOGIJOS</b> |       |

| Dėstytojas (-ai)                                                          | Padalinys (-iai)                                                                                                  |
|---------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| <b>Koordinuojantis:</b> lekt. Paulius Danielius<br><br><b>Kitas (-i):</b> | Kauno fakultetas<br>Socialinių mokslų ir taikomosios<br>informatikos institutas Muitinės g.<br>8, LT-44280 Kaunas |

| Studijų pakopa | Dalyko (modulio) lygmuo | Dalyko (modulio) tipas |
|----------------|-------------------------|------------------------|
| Pirmoji        |                         | privalomasis           |

| Igyvendinimo forma | Vykdymo laikotarpis | Vykdymo kalba (-os) |
|--------------------|---------------------|---------------------|
| auditorinė         | 6 semestras         | lietuvių            |

| Reikalavimai studijuojančiajam                                        |                                            |
|-----------------------------------------------------------------------|--------------------------------------------|
| <b>Išankstiniai reikalavimai:</b><br>Aukštoji matematika, informatika | <b>Gretutiniai reikalavimai (jei yra):</b> |

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

| Dalyko (modulio) tikslas: studijų programos ugdomos kompetencijos                                                                                                                                                                                                                                                                                                                                                     |                                                                                                        |                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| 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 siekiniai                                                                                                                                                                                                                                                                                                                                                                                    | 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.<br>Kolokviumas: testas e. mokymo sistemoje.<br>Egzaminas. |
| Geba atlikti kvalifikuotą informatikos inžinerijos dalykinės srities problemos analizę, naudojant aktualius ir tinkamai parinktus empirinius duomenis, laikydamasis teorinių prieigų, taikydamas efektyvius metodus.                                                                                                                                                                                                  | Paskaitos, praktiniai darbai, aktyvaus mokymosi metodai (grupės diskusija, metodų lyginamoji analizė). |                                                                                              |

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

| Temos                                                                                                                                            | Kontaktinio darbo valandos |               |           |          |                       |          |                          | Savarankiškų studijų laikas ir užduotys |                                                                                                                                                                                                                                  |
|--------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|---------------|-----------|----------|-----------------------|----------|--------------------------|-----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                                  | Paskaitos                  | Konsultacijos | Egzaminas | Pratybos | Laboratoriniai darbai | Praktika | Visas kontaktinis darbas | Savarankiškas darbas                    | Užduotys                                                                                                                                                                                                                         |
| <b>1. Verslo analitikos sistemos</b>                                                                                                             | 2                          |               |           |          | 2                     |          | 4                        | 8                                       | 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. |
| Įvadas. Verslo analitikos (BI) sistemų raida ir BI sistemų klasifikacija. Duomenų gavybos įrankiai, Rapid Miner paketas. Duomenų vizualizavimas. | 2                          |               |           |          | 2                     |          | 4                        | 8                                       |                                                                                                                                                                                                                                  |
| <b>2. Duomenų sandėliavimo sistemos</b>                                                                                                          | 2                          |               |           |          |                       |          | 4                        | 8                                       |                                                                                                                                                                                                                                  |
| Duomenų integravimas, duomenų sandėliavimo sistemos ir technologijos, architektūros, ETL.                                                        | 2                          |               |           |          |                       |          | 4                        | 8                                       |                                                                                                                                                                                                                                  |
| <b>3. Duomenų gavyba</b>                                                                                                                         | 4                          |               |           |          | 6                     |          | 12                       | 24                                      |                                                                                                                                                                                                                                  |
| Duomenų gavybos gyvavimo ciklas. Žinių atradimas duomenų bazėse.                                                                                 | 2                          |               |           |          | 2                     |          | 4                        | 8                                       |                                                                                                                                                                                                                                  |
| Duomenų gavybos taikymas rinkodaros, švietimo ir kitose srityse.                                                                                 |                            |               |           |          | 2                     |          | 2                        | 8                                       |                                                                                                                                                                                                                                  |
| Duomenų analizės standartai: CRISP-DM. Duomenų gavybos uždavinių pavyzdžiai. Duomenų parengimo procesai.                                         | 2                          |               |           |          | 2                     |          | 6                        | 8                                       | Pasiruošimas koliokviumui.                                                                                                                                                                                                       |
| <b>Kolokviumas</b>                                                                                                                               | 2                          |               |           |          |                       |          | 2                        |                                         |                                                                                                                                                                                                                                  |
| <b>3. Duomenų gavybos algoritmai</b>                                                                                                             | 10                         |               |           |          | 20                    |          | 26                       | 38                                      | Laboratoriniai darbai.<br><br>Pasiruošimas egzaminui.                                                                                                                                                                            |
| Duomenų gavybos uždavinių tipai. Sprendimų medžių algoritmas.                                                                                    | 2                          |               |           |          | 4                     |          | 6                        | 8                                       |                                                                                                                                                                                                                                  |
| Asociacijų taisyklės.                                                                                                                            | 2                          |               |           |          | 4                     |          | 6                        | 8                                       |                                                                                                                                                                                                                                  |
| Teksto ir web duomenų analizė ir gavyba.                                                                                                         | 2                          |               |           |          | 4                     |          | 6                        | 8                                       |                                                                                                                                                                                                                                  |
| Neuroniniai tinklai.                                                                                                                             | 2                          |               |           |          | 4                     |          | 6                        | 8                                       |                                                                                                                                                                                                                                  |
| Kiti duomenų gavybos algoritmai. Regresinė analizė. Artimiausių kaimynų algoritmas. Klasterizavimo algoritmai. Genetiniai algoritmai.            | 2                          |               |           |          |                       |          | 2                        | 6                                       |                                                                                                                                                                                                                                  |
| Python programavimo kalbos priemonės duomenų gavybai.                                                                                            |                            |               |           |          | 4                     |          |                          |                                         |                                                                                                                                                                                                                                  |
| Konsultacija                                                                                                                                     |                            | 2             |           |          |                       |          | 2                        |                                         |                                                                                                                                                                                                                                  |
| Egzaminas                                                                                                                                        |                            |               | 2         |          |                       |          | 2                        |                                         |                                                                                                                                                                                                                                  |
| <b>Iš viso</b>                                                                                                                                   | <b>20</b>                  | <b>2</b>      | <b>2</b>  |          | <b>28</b>             |          | <b>52</b>                | <b>78</b>                               |                                                                                                                                                                                                                                  |

| Vertinimo strategija       | Svoris proc. | Atsiskaitymo laikas | Vertinimo kriterijai<br><i>Išsamiau apie vertinimo strategiją ir AI generatyvinių modelių naudojimą žr. po lentelę.</i>                                                                                                                                                                                                           |
|----------------------------|--------------|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Praktinės užduotys         | 40           | 2 – 16-a savaitės   | Praktinės užduotys atliekamos viso semestro metu, darbų ataskaitos turi būti įkeliamos į VMA (MS Word arba PDF formatu) pagal pateiktą atsiskaitymų grafiką. Kiekviena užduotis vertinama 10 balų skalėje pagal VU vertinimo kriterijus.<br>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 lektoriumi. Darbas turi būti pristatomas klasėje su pateiktimi pristatymams skirtu užsiėmimų laiku.                                                                                                   |
| 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.                                                                                                                                                                                                                 |

### Darbų vertinimo strategija ir AI (Artificial intelligence) generatyvinių modelių naudojimo gairės.

Individualus rašto darbas rašomas pasirinkus ir su dėstytoju suderinus temą iš pateikto sąrašo. Ruošdamas šį darbą bei praktinių darbų ataskaitas studentas turi naudoti patikimus interneto šaltinius ir mokslinius straipsnius bei gali naudoti AI generatyvinius modelius laikydamasis akademinio sąžiningumo principų (*Copy-Paste* principas laikomas plagiatu – turi būti naudojamas citavimas, **plačiau - žr. žemiau**). Jei rašto darbe naudota nepakankamai šaltinių, jų tarpe nėra mokslinių straipsnių, darbo vertinimas atitinkamai mažinamas.

Pristatymui studentai skaidres parengia kaip pagalbinę priemonę, jos nėra skirtos skaitymui, todėl studentai, siekiantys maksimalaus balo, kokybiškai pristato nurodytą temą (skaitymas nuo skaidrių vertinamas minimaliu balu).

**Išsamesni nurodymai rašto darbui ir pateikčiai studentams pateikiami renkantis temas.**

### AI generatyvinių modelių naudojimo pavyzdžiai

Geriausia tokias priemones naudoti:

- mokymosi proceso metu aiškinantis metodų, algoritmų bei technikų principus,
- sąvokų paaiškinimui,
- struktūros sudarymui,
- idėjų generavimui,
- konkrečių atvejų paieškai,
- santraukų generavimui (tolesniam darbui),
- didelių tekstų apdorojimui (tolesniam darbui).

Visa AI priemonėmis sugeneruota informacija privalo būti tikrinama, ir darbuose pateikiamos tai įrodančios nuorodos į išorinius šaltinius užtikrinant tinkamą citavimą (bet kuriuo atveju *Copy-Paste* principas laikomas plagiatu, jei necituojama). Taip pat svarbu suprasti, kad AI generatyviniai modeliai nėra darbo bendraautoriai. Plačiau apie VU akademinį sąžiningumą (ypač atkreipkite dėmesį į 19 punktą): [https://www.vu.lt/site\\_files/Veiklos\\_dokumentai/Akademines\\_etikos\\_kodeksas\\_VU\\_2018.pdf](https://www.vu.lt/site_files/Veiklos_dokumentai/Akademines_etikos_kodeksas_VU_2018.pdf).

### Kada AI generatyvinių modelių negalima naudoti šiame kurse?

Šių priemonių negalima naudoti:

- Rašto darbuose ir pateiktyse teksto pateikimui *Copy-Paste* principu be atitinkamo citavimo.
- Teksto pagražinimui (tai netaikoma mašininio vertimo priemonėms, tokioms kaip *DeepL*).
- Laikant įvertinimo testus semestro ir egzamino metu.

### Jei rengiant darbus buvo naudoti AI generatyviniai modeliai?

Jei rengiant rašto darbą idėjų generavimui buvo naudoti AI generatyviniai modeliai, rašto darbo pradžioje privalo būti aprašyta

- AI priemonių naudojimo strategija,
- kokie klausimai buvo užduodami,
- koks gautas rezultatas ir kiek procentų gauto rezultato modifikuota ir pritaikyta rengiant darbą.

Prieduose turi būti pateikiamos užklausos (pvz., Chat GPT užklausa: „...“) ir rezultatai (pvz., Chat GPT sugeneruotas atsakymas „...“) bei nurodoma generatyvinio modelio pavadinimas, versija ir naudojimo data. Plačiau apie citavimą <https://apastyle.apa.org/blog/how-to-cite-chatgpt>, <https://guides.library.uq.edu.au/referencing/chatgpt-and-generative-ai-tools>.

Taip pat turi būti aprašyta, kokia AI priemonių sugeneruoto teksto apimtis naudota darbe. Jei tekstas yra kopijuojamas iš generatyvinių modelių sistemų, privalo būti cituojamas, kaip ir bet kuris šaltinis. Rašto darbe citatų ir AI sugeneruotų tekstų negali būti daugiau nei 20% (pvz., plačiau <https://plagiarismcheck.org/blog/what-is-the-acceptable-percentage-of-plagiarism/>).

Naudojant AI generatyvinius modelius svarbu, kad studentai kritiškai vertintų pateikiamus atsakymus, laikytųsi etiškumo principo, informacija turi būti tiksli, o taip pat kiekvienas studentas turi užtikrinti skaidrumą kitų grupės narių atžvilgiu.

**Akademinio nesąžiningumo atveju:** dėstytojas, pastebėjęs plagijavimo požymių ar nustatęs, kad rašto darbe yra dirbtinio intelekto priemonių sugeneruotų tekstų blokų (t.y. įtariamas akademinis nesąžiningumas), informuoja administraciją. Tokiu atveju inicijuojamas procesas akademinio sąžiningumo įvertinimui.

| Autorius                                | Leidimo metai | Pavadinimas                                                                    | Periodinio leidinio Nr. ar leidinio tomas | Leidimo vieta ir leidykla ar internetinė nuoroda            |
|-----------------------------------------|---------------|--------------------------------------------------------------------------------|-------------------------------------------|-------------------------------------------------------------|
| <b>Privalomoji literatūra</b>           |               |                                                                                |                                           |                                                             |
| P. Danielius                            | 2024          | Moodle aplinka                                                                 |                                           | <a href="https://emokymai.vu.lt">https://emokymai.vu.lt</a> |
| Mohammed J. Zaki, Wagner Meira          | 2014          | Data mining and analysis: fundamental concepts and algorithms                  |                                           | New York [N.Y.] : Cambridge University Press                |
| Ian H. Witten, Eibe Frank, Mark A. Hall | 2011          | Data mining: practical machine learning tools and techniques                   |                                           | Burlington, MA : Morgan Kaufmann                            |
| G. Dzemyda, O. Kurasova, J. Zilinskas   | 2012          | Multidimensional Data Visualization Methods and Applications                   |                                           | Springer                                                    |
| <b>Papildoma literatūra</b>             |               |                                                                                |                                           |                                                             |
| Andreas C. Müller, Sarah Guido          | 2018          | Introduction to machine learning with Python: a guide for data scientists      |                                           | Sebastopol : O'Reilly                                       |
| Gabe Ignatow, Rada Mihalcea             | 2018          | An introduction to text mining: research design, data collection, and analysis |                                           | Los Angeles [etc.] : SAGE                                   |