

| OSNOVNI PODACI O PREDMETU   |  |         |
|---|--|---------|
| Naziv predmeta  | Primijenjena multivarijatna analiza podataka za informatičare                  |         |
| Studijski program   | Diplomski studij Informatika   |         |
| Status predmeta   | obvezatan za modul IIS   |         |
| Semestar  | 1.   |         |
| Bodovna vrijednost i način izvođenja nastave  | ECTS koeficijent opterećenosti studenata                                       | 6       |
|   | Broj sati (P+V+S)  | 30+30+0 |
| Nositelj predmeta   | Prof. dr. sc. Maja Matetić   |         |
| E-mail  | <a href="mailto:majam@uniri.hr">majam@uniri.hr</a>                             |         |
| Ured  | Radmile Matejčić 2, kabinet O-407  |         |
| Vrijeme konzultacija  | uz prethodni dogovor e-mailom, četvrtkom 11.00.-12.00.                         |         |
| Asistent  | Dejan Ljubobratović, mag. educ.  |         |
| E-mail  | <a href="mailto:dejan.ljubobratovic@uniri.hr">dejan.ljubobratovic@uniri.hr</a> |         |
| Ured  | Radmile Matejčić 2, kabinet O-522  |         |
| Vrijeme konzultacija  | Petkom od 11:30 do 12:30 ili po dogovoru e-mailom                              |         |
| DETALJNI OPIS PREDMETA  |  |         |
| <i>Ciljevi predmeta</i>   |  |         |
| Ciljevi predmeta uključuju upoznavanje postupaka sumiranja i vizualizacije različitih tipova podataka i identifikacije prikladnih metoda podatkovne analitike, razumijevanje temeljnih mehanizama multivarijatnih modela i njihovu evaluaciju i interpretaciju, uporabu analitičkih alata i suvremenu programsku podršku u praksi. Cilj predmeta je dodatno ispitati matematičke temelje numeričkih algoritama koji se koriste i istražiti njihovu upotrebu putem praktičnih primjera u raznim domenama primjene.   |  |         |
| <i>Uvjeti za upis predmeta</i>  |  |         |
| Nema uvjeta za upis predmeta.   |  |         |
| <i>Očekivani ishodi učenja za predmet</i>   |  |         |
| Očekuje se da nakon izvršavanja svih programom predviđenih obveza studenti budu sposobni:   |  |         |
| <ol style="list-style-type: none"> <li>1. Izvesti multivarijatnu istraživačku analizu podataka</li> <li>12. Objasniti karakteristike i svojstva multivarijatne normalne distribucije. Opisati koncept međusobnih odnosa višedimenzionalnih podataka putem korelacije</li> <li>13. Oblikovati jednostavne skripte za procesiranje, analizu i vizualizaciju podataka uporabom alata za podatkovnu analizu</li> <li>14. Provjeriti preduvjete, postaviti hipoteze, izvesti i interpretirati rezultate primjene sljedećih analitičkih metoda na višedimenzionalnim podacima: višestruka regresija, analiza glavnih komponenti, faktorska analiza, diskriminantna analiza, višedimenzionalno skaliranje, grupiranje, analiza mreža</li> <li>15. Objasniti i analizirati problem nestabilnosti numeričkog izračuna</li> </ol> |  |         |

- 16. Analizirati složenost i točnost algoritama za rješavanje problema numeričke analize kao što su rješavanje nelinearne jednačbe, traženje ekstrema nelinearne realne funkcije više varijabli, traženje interpolacijskog polinoma, efikasno računanje vrijednosti polinoma, numeričko deriviranje i integriranje
- 17. Implementirati zadane numeričke algoritme iz područja multivarijatne statistike i umjetne inteligencije u programskom jeziku (npr. R, Python).
- 18. Kritički interpretirati rezultate multivarijatne analize podataka i provesti individualni projekt multivarijatne analize podataka u prikladnom računalnom statističkom okruženju

*Sadržaj predmeta*

Sadržaj predmeta čine teme:

- Multivarijatni podaci i multivarijatne statističke metode. Multivarijatne vizualizacije.
- Multivarijatna normalna distribucija. Svojstva distribucija.
- Koeficijenti korelacije i primjena. Analiza varijance. Višestruka i multivarijatna regresija.
- Redukcija dimenzija. Analiza glavnih komponenti. Faktorska i diskriminativna analiza. Klasifikacija.
- Grupiranje. Višedimenzionalno skaliranje.
- Aritmetika pomičnog zareza. Stabilnost algoritama.
- Direktne metode za rješavanje linearnih sustava. Polinomijalna interpolacija.
- Numeričko integriranje i deriviranje. Linearni problem najmanjih kvadrata.
- Numeričko rješavanje nelinearnih jednačbi.

|                                |  |   |
|--------------------------------|--|---|
| <i>Vrsta izvođenja nastave</i> | <input checked="" type="checkbox"/> predavanja           | <input checked="" type="checkbox"/> samostalni zadaci |
|                                | <input checked="" type="checkbox"/> seminari i radionice | <input type="checkbox"/> multimedija i mreža          |
|                                | <input checked="" type="checkbox"/> vježbe               | <input type="checkbox"/> laboratorij                  |
|                                | <input type="checkbox"/> obrazovanje na daljinu          | <input type="checkbox"/> mentorski rad                |
|                                | <input type="checkbox"/> terenska nastava                | <input type="checkbox"/> ostalo                       |

*Komentari*

*Obavezna literatura (u trenutku prijave prijedloga studijskog programa)*

1. Härdle, Wolfgang Karl, and Léopold Simar. Applied multivariate statistical analysis. Springer Nature, 2019.
2. James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning with applications in R. Vol. 112. New York: Springer, 2021.
3. Bloomfield, Victor A. Using R for numerical analysis in science and engineering. Chapman and Hall/CRC, 2018.
4. Greenbaum, Anne, and Tim P. Chartier. Numerical methods: design, analysis, and computer implementation of algorithms. Princeton University Press, 2012.

*Dopunska literatura (u trenutku prijave prijedloga studijskog programa)*

1. Everitt, Brian, and Torsten Hothorn. An introduction to applied multivariate analysis with R. Springer Science & Business Media, 2011.
2. Johnson, Richard Arnold, and Dean W. Wichern. Applied multivariate statistical analysis. Vol. 6. London, UK:: Pearson, 2014.
3. Denis, Daniel J. Univariate, Bivariate, and Multivariate Statistics Using R: Quantitative Tools for Data Analysis and Data Science. John Wiley & Sons, 2020.

|  |     |
|--|-----|
| 4. G. Tabachnick, L.S. Fidell, Using multivariate statistics, 6th Edition, Pearson, 2018.  |     |
| 5. Hair J.F. et al. Multivariate Data Analysis, 7th Edition, Pearson, 2014.  |     |
| 6. Wickham, Hadley, and Garrett Grolemond. R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc.", 2017.  |     |
| <i>Načini praćenja kvalitete koji osiguravaju stjecanje izlaznih znanja, vještina i kompetencija</i>   |     |
| Predviđa se periodičko provođenje evaluacije s ciljem osiguranja i kontinuiranog unapređenja kvalitete nastave i studijskog programa (u okviru aktivnosti Odbora za upravljanje i unapređenje kvalitete Fakulteta informatike i digitalnih tehnologija). U zadnjem tjednu nastave provodit će se anonimna evaluacija kvalitete održane nastave od strane studenata. Provest će se i analiza uspješnosti studenata na predmetu (postotak studenata koji su položili predmet i prosjek njihovih ocjena). |     |
| <i>Mogućnost izvođenja na stranom jeziku</i>   | Da. |

### OBVEZE, PRAĆENJE RADA I VREDNOVANJE STUDENATA

| VRSTA AKTIVNOSTI                                   | ECTS     | ECTS - PRAKTIČNI RAD | ISHODI UČENJA | SPECIFIČNA AKTIVNOST   | METODA PROCJENJIVANJA   | BODOVI MAX. |
|--|----------|----------------------|---------------|--|---|-------------|
| Pohađanje nastave i aktivnosti u nastavi           | 2        | 0.5                  | 1-8           | Aktivnost na satu  | Sudjelovanje u rješavanju problema  | 5           |
| Kontinuirana provjera znanja - kvizovi             | 1        | 0.25                 | 1-8           | Dva kviza (provjere na računalu: teorija + oblikovanje programskog koda) | 2 kviza: (12+13) bodova ovisno o stupnju točnosti i potpunosti                    | 25          |
| Kontinuirana provjera znanja - kolokviji           | 1        | 1                    | 1-8           | 2 praktična kolokvija (rješavanje zadataka na računalu)                  | 2 praktična kolokvija: (20+20) bodova ovisno o stupnju točnosti i potpunosti      | 40          |
| Seminarski rad i projektni zadatak (završni ispit) | 2        | 2                    | 1-8           | Istraživanje uporabom odgovarajućih alata                                | Prezentacija seminarskog rada i projekta (na kraju semestra ili na ispitnom roku) | 30          |
| <b>UKUPNO</b>                                      | <b>6</b> | <b>3.75</b>          |               |  |   | <b>100</b>  |

#### Obveze i vrednovanje studenata

##### 1. Pohađanje nastave i aktivnosti u nastavi

Nastava se odvija prema mješovitom modelu u kombinaciji klasične nastave u učionici i *online* nastave uz pomoć sustava za e-učenje prema rasporedu koji je prikazan je tablicom u nastavku. Studenti su dužni koristiti sustav za e-učenje Merlin (<https://moodle.srce.hr/>) gdje će se objavljivati informacije o predmetu, materijali za učenje, zadaci za vježbu, zadaci za domaće zadaće te obavijesti vezane za izvođenje nastave (putem foruma Obavijesti).

##### 2. Kontinuirana provjera znanja – teorija

Tijekom semestra pisat će se u učionici uz nadzor nastavnika dva kviza putem sustava Merlin koji uključuje teorijska pitanja i praktične zadatke. Na ovoj aktivnosti studenti će moći skupiti ukupno maksimalno 25 bodova (definiran je ukupni prag od 6 bodova za zbroj bodova kvizova).

### 3. Kontinuirana provjera znanja – praktični dio

U okviru vježbi održavaju se dva praktična kolokvija (rješavanje zadataka na računalu) koji se boduju sa ukupno 40 bodova (20+20 bodova, definiran je ukupni prag od 10 bodova za zbroj bodova kolokvija).

### 4. Završni ispit (Seminarski rad i projekt)

Tijekom semestra studentima će biti ponuđene teme za projektni zadatak. Obrana projekta predstavlja završni ispit. Studenti za koje se utvrdi da su u seminarskom radu koristili veće dijelove doslovno kopiranog teksta dobit će 0 bodova.

Seminarski rad i/ili projektni zadatak studenti mogu rješavati samostalno i/ili u projektnom timu. Hodogram aktivnosti studenta:

- a) Priprema podataka za multivarijatnu analizu podataka
- b) Odabir metoda, kratak opis i objašnjenje metoda
- c) Izvođenje eksperimenata
- d) Prikaz rezultata, interpretacija rezultata
- e) Izrada izvješća projekta (svaki član piše vlastito izvješće u par rečenica o tome koji su mu bili zadaci na projektu i kako je tim funkcionirao)
- f) Predaja zajedničkog seminara i izvješća na kraju semestra (prva verzija)
- g) Cijeli tim zajednički predstavlja rezultate projekta na kraju semestra ili na ispitnom roku. Svaki član tima izlaže dio kojim se bavio na projektu. Bar dva dana prije obrane potrebno je na Merlin predati materijale uz seminarski rad, izvješća i prezentaciju.

Seminarski rad i projekt se izlaže po dogovoru na kraju semestra ili u okviru termina završnog ispita i nosi ukupno 30 bodova, a smatra se pozitivno ocijenjenim samo ako student postigne minimalno 50%-ni uspjeh (15-30 bodova).

### 5. Nadoknada aktivnosti

Nadoknadi kolokvija 27.01.2023. student ima pravo pristupiti samo u slučaju opravdanog izostanka s kviza/kolokvija uz predočenje ispričnice asistentu u roku 10 dana nakon izostanka.

### Ocjenjivanje

Kontinuiranim radom tijekom semestra na prethodno opisani način studenti mogu ostvariti najviše 70 ocjenskih bodova, a da bi mogli pristupiti završnom ispitu moraju ostvariti 50% i više bodova (minimalno 35).

Završni ispit nosi udio od maksimalno 30 ocjenskih bodova, a smatra se položenim samo ako na njemu student postigne minimalno 50%-ni uspjeh (ispitni prag je 50% uspješno riješenih zadataka).

Ukoliko je završni ispit prolazan, skupljeni bodovi će se pribrojati prethodnima i prema ukupnom rezultatu formirati će se pripadajuća ocjena. U suprotnom, student ima pravo pristupa završnom ispitu još 2 puta (ukupno do 3 puta).

### Konačna ocjena iz predmeta

Donosi se na osnovu zbroja svih bodova prikupljenih tijekom izvođenja nastave prema sljedećoj skali:

|                 |                            |
|-----------------|----------------------------|
| A – 90% - 100%  | (ekvivalent: izvrstan 5)   |
| B – 75% - 89,9% | (ekvivalent: vrlo dobar 4) |
| C – 60% - 74,9% | (ekvivalent: dobar 3)      |
| D – 50% - 59,9% | (ekvivalent: dovoljan 2)   |
| F – 0% - 49,9%  | (ekvivalent: nedovoljan 1) |

### **Ispitni rokovi**

Redoviti:

10.02.2023. i 24.02.2023.

Izvanredni:

10.03.2023. i 08.09.2023.

**RASPORED NASTAVE – zimski (I.) semestar akademske godine 2022./2023.**

Nastava će se na predmetu odvijati u zimskom semestru prema sljedećem rasporedu:

predavanja: petkom od 8.15 do 09.45 sati u prostoriji O-365

vježbe: petkom od 10.00 do 11.30 sati u prostoriji O-365

| Tj. | Datum  | Vrijeme     | Prostor | Tema  | Nastava | Izvođač   |
|-----|--------|-------------|---------|---|---------|---|
| 1.  | 07.10. | 8:15-9:45   | 365     | Uvod u primijenjenu statistiku. Statističko učenje.                       | P       | prof. dr. sc. Maja Matetić                        |
| 1.  | 07.10. | 10:00-11:30 | 365     | Uvod u predmet. Markdown jezik.   | V       | Dejan Ljubobratović                               |
| 2.  | 14.10. | 8:15-9:45   | 365     | Statistički testovi. Snaga statistike.                                    | P       | prof. dr. sc. Maja Matetić                        |
| 2.  | 14.10. | 10:00-11:30 | 365     | Jupyter bilježnice (notebooks)  | V       | Dejan Ljubobratović                               |
| 3.  | 21.10. | 8:15-9:45   | 365     | Analiza varijance.  | P       | prof. dr. sc. Maja Matetić                        |
| 3.  | 21.10. | 10:00-11:30 | 365     | R markdown jezik.   | V       | Dejan Ljubobratović                               |
| 4.  | 28.10. | 8:15-9:45   | 365     | Regresijski modeli.   | P       | prof. dr. sc. Maja Matetić                        |
| 4.  | 28.10. | 10:00-11:30 | 365     | Analiza varijance - ANOVA   | V       | Dejan Ljubobratović                               |
| 5.  | 4.11.  | 8:15-9:45   | 365     | GLM. Diskriminantna analiza.  | P       | prof. dr. sc. Maja Matetić                        |
| 5.  | 4.11.  | 10:00-11:30 | 365     | GLM, logistička regresija   | V       | Dejan Ljubobratović                               |
| 6.  | 11.11. | 8:15-9:45   | 365     | <b>1. Kviz teorije</b>  | P       | prof. dr. sc. Maja Matetić                        |
| 6.  | 11.11. | 10:00-11:30 | 365     | Logistička regresija  | V       | Dejan Ljubobratović                               |
| 7.  | 18.11. | 8:15-9:45   | online  | Analiza glavnih komponenti.   | P       | prof. dr. sc. Maja Matetić                        |
| 7.  | 18.11. | 10:00-11:30 | online  | Diskriminantna analiza  | V       | Dejan Ljubobratović                               |
| 8.  | 25.11. | 8:15-9:45   | 365     | Faktorska analiza. Odabir modela.   | P       | prof. dr. sc. Maja Matetić                        |
| 8.  | 25.11. | 10:00-11:30 | 365     | Analiza glavnih komponenti (PCA) i Analiza linearne diskriminacije (LDA). | V       | Dejan Ljubobratović                               |
| 9.  | 2.12.  | 8:15-9:45   | 365     | Regularizacija.   | P       | prof. dr. sc. Maja Matetić                        |
| 9.  | 2.12.  | 10:00-11:30 | 365     | <b>1. kolokvij</b>  | V       | Dejan Ljubobratović                               |
| 10. | 9.12.  | 8:15-9:45   | 365     | Unakrsno vrednovanje (Cross validation), bootstrap.                       | P       | prof. dr. sc. Maja Matetić                        |
| 10. | 9.12.  | 10:00-11:30 | 365     | Unakrsno vrednovanje (Cross validation), bootstrap.                       | V       | Dejan Ljubobratović                               |
| 11. | 16.12. | 8:15-9:45   | 365     | Metode temeljene na stablima.   | P       | prof. dr. sc. Maja Matetić                        |
| 11. | 16.12. | 10:00-11:30 | 365     | Stabla odlučivanja  | V       | Dejan Ljubobratović                               |
| 12. | 23.12. | 8:15-9:45   | 365     | Metoda potpornih vektora (SVM - support vector machines). Grupiranje.     | P       | prof. dr. sc. Maja Matetić                        |
| 12. | 23.12. | 10:00-11:30 | 365     | Algoritam slučajne šume (RF - Random forest)                              | V       | Dejan Ljubobratović                               |
| 13. | 13.01. | 8:15-9:45   | 365     | <b>2. Kviz teorije</b>  | P       | prof. dr. sc. Maja Matetić                        |
| 13. | 13.01. | 10:00-11:30 | 365     | Metoda potpornih vektora (SVM - support vector machines).                 | V       | Dejan Ljubobratović                               |
| 14. | 20.01. | 8:15-9:45   | 365     | Analiza preživljavanja. Numeričke metode.                                 | P       | prof. dr. sc. Maja Matetić                        |
| 14. | 20.01. | 10:00-11:30 | 365     | <b>1. Kolokvij</b>  | V       | Dejan Ljubobratović                               |
| 15. | 27.01  |             | 365     | Nadoknade   | P i V   | prof. dr. sc. Maja Matetić<br>Dejan Ljubobratović |

P – predavanja

V – vježbe

**Napomena:** Moguće su manje izmjene rasporeda nastave. Za nove verzije rasporeda potrebno je pratiti obavijesti u e kolegiju.