

UČNI NAČRT PREDMETA / COURSE SYLLABUS

Predmet:	Napredne statistične metode
Course title:	Advanced Statistical Methods

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Informatika v sodobni družbi, magistrski študijski program druge stopnje	-	Prvi ali drugi	Drugi ali četrty
Informatics in Contemporary Society, second cycle Masters Study Programme	-	First or second	Second or fourth

Vrsta predmeta / Course type

Izbirni / Elective

Univerzitetna koda predmeta / University course code:

1-ISD-MAG-IP-NSM-2016-10-1

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30	50	-	-	20	50	5

Nosilec predmeta / Lecturer:

**Jeziki /
Languages:**

**Predavanja /
Lectures:**

Slovenski, angleški / Slovene, English

Vaje / Tutorial:

Slovenski, angleški / Slovene, English

Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:

Za vključitev v predmet mora imeti študent/ka opravljen izpit iz Metod kvantitativne analize. Študent/študentka mora pred pristopom k izpitu pripraviti in zagovarjati seminarsko nalogo.

Prerequisites:

To join the course the student must have passed the exam in *Quantitative analysis*. Before taking the exam the student must successfully prepare and present the lab work task.

Vsebina:

- Statistično učenje: nadzorovano in nenadzorovano učenje, regresijski in klasifikacijski problemi, statistični model, ocenjevanje točnosti modela (kvaliteta prileganja modela podatkom, kompromis med pristranostjo in razpršenostjo).
- Linearna regresija: preprosta linearna regresija, večkratna linearna regresija,

Content (Syllabus outline):

- Statistical learning: supervised vs. unsupervised learning, regression vs. classification problems, statistical model, assessing model accuracy (measuring the quality of fit, the bias-variance trade-off).
- Linear regression: simple linear regression, multiple linear regression, assessing the accuracy of the model,

ocenjevanje kakovosti napovedi, kvalitativni prediktorji, razširitve linearnega modela.

- Klasifikacija: logistična regresija (logistični model, ocenjevanje regresijskih koeficientov, večkratna logistična regresija, logistična regresija za večrazredne probleme), linearna in kvadratična diskriminantna analiza, k-najbližjih sosedov, primerjava klasifikacijskih modelov.
- Samovzorčenje: križno preverjanje, zankanje.
- Drevesne metode: regresijska in klasifikacijska drevesa, bagging, slučajni gozdovi, boosting.
- Metoda podpornih vektorjev: klasifikator z maksimalnim robom, podporni vektorji, metoda podpornih vektorjev, večrazredna klasifikacija.
- Nenadzorovano učenje: analiza glavnih komponent, razvrščanje v skupine (metoda voditeljev, hierarhično razvrščanje).

qualitative predictors, extensions of the linear model.

- Classification: logistic regression (logistic model, estimating the regression coefficients, multiple logistic regression, multiclass logistic regression), linear and quadratic discriminant analysis, k-nearest neighbors.
- Resampling methods: cross-validation, bootstrap
- Tree-based methods: regression and classification trees, bagging, random forests, boosting.
- Support vector machines: maximal margin classifier, support vector classifier, support vector machines, multiclass classification.
- Unsupervised learning: principal component analysis, clustering methods (k-means clustering, hierarchical clustering).

Temeljni literatura in viri / Readings:

- HASTIE, TREVOR, TIBSHIRANI, ROBERT IN FRIEDMAN, JEROME H. (2013) The elements of statistical learning: data mining, inference, and prediction, New York, Springer.
- SUGIYAMA, MASASHI (2016) *Introduction to Statistical Machine Learning*, Waltham, MorganKauffman.
- GARETH JAMES, DANIELA WITTEN, TREVOR HASTIE, ROBERT TIBSHIRANI (2013) *An Introduction to Statistical Learning: with Applications in R*, Springer Science & Business Media.

Cilji in kompetence:

Učna enota prispeva k razvoju naslednjih splošnih in predmetno specifičnih kompetenc:

Splošne kompetence:

- poznavanje in razumevanje družbenih procesov in sposobnost njihove kompleksne analize
- uporaba metodoloških orodij, tj. izvajanje, koordiniranje in organiziranje raziskav, uporaba raznih raziskovalnih metod in tehnik
- uporaba in kombiniranje znanj za različnih disciplinarnih področij

Objectives and competences:

The module contributes to the following general and subject-specific competences:

General competences:

- knowledge and understanding of social processes and ability to perform complex analysis of these processes
- the use of methodological tools, ie. implementation, coordination and organization of research, application of various research methods and techniques
- the use and combining the

- sposobnost uporabe informacijsko-komunikacijskih tehnologij in sistemov na področju družboslovja
- sposobnost zbiranja, shranjevanja, analize in interpretacije velikih količin podatkov

*Predmetno-specifične kompetence:
Po opravljenem izpitu iz predmeta bo študent/ka pridobil/a kompetence:*

- poznavanje najpomembnejših metod v statističnem učenju
- poznavanje programskih orodij za analizo podatkov

- knowledge from different disciplines
- the ability to use information and communications technologies and systems in social science
- ability to collect, store, analyse and to interpret big data

*Subject-specific competences:
Upon completion of this course the student will have competences:*

- knowledge of most important methods in statistical learning
- knowledge of tools for data analysis

Predvideni študijski rezultati:

Znanje in razumevanje:

Študent/študentka:

- razume posebnosti statističnega učenja v primerjavi s klasičnimi podatkovnimi analizami
- spozna metode, ki so primerne za analize tovrstnih podatkov

Intended learning outcomes:

Knowledge and understanding:

The student:

- understands the specificity of statistical learning compared to classical data analysis
- she learns methods, designed for data analysis and statistical learning in particular

Metode poučevanja in učenja:

- *predavanja* z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov, predstavitve)
- vaje v računalniški učilnici
- *individualne in skupinske konzultacije* (diskusija, dodatna razlaga, obravnava specifičnih vprašanj)

Learning and teaching methods:

- *lectures* (explanation with discussions, questions, case-studies, presentations)
- *tutorials* in the computer classroom
- *individual and group consultations* (debate, additional explanations, considering specific issues)

Načini ocenjevanja:

Način:

- ustni izpit
- pisni izpit in seminarska naloga

Ustnega izpita ni potrebno opravljati, kadar študent s pisnim izpitom in seminarsko nalogo zbere vsaj 70% točk in je bil vsaj 50% na predavanjih.

Delež (v %) /
Weight (in %)

30
70 (45 + 25)

Assessment:

Type:

- oral exam
- written exam and seminar

Students having lecture attendance larger than 50 % and passing the written exam and the seminar work with at least 70 % of all possible points are exempted from the oral examination.