

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-2019-05-13

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30	-	20	-	-	100	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 uspešno izdelati 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 the seminar work.

Vsebina:

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

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,

<p>kvalitativni prediktorji, razširitve linearnega modela.</p> <ul style="list-style-type: none"> • Klasifikacija: logistična regresija (logistični model, ocenjevanje regresijskih koeficientov, večkratna logistična regresija, logistična regresija za večrazredne probleme), linearna in kvadratična diskriminantna analiza, <i>k</i>-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). 	<p>qualitative predictors, extensions of the linear model.</p> <ul style="list-style-type: none"> • Classification: logistic regression (logistic model, estimating the regression coefficients, multiple logistic regression, multiclass logistic regression), linear and quadratic discriminant analysis, <i>k</i>-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:

- 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 ter ocenitev njihove uporabnosti
- sposobnost sinteze izvirnih idej, konceptov in rešitev določenih problemov iz različnih disciplinarnih področij

Objectives and competences:

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

- knowledge and understanding of social processes and ability to perform complex analysis of these processes use of methodological tools, i.e. implementation, coordination and organisation of research, use of various research methods and techniques and to evaluate their usefulness
- competence to form original ideas, concepts and solutions for specific problems from different disciplines

- poznavanje in razumevanje širokega nabora aplikacij informacijsko komunikacijske tehnologije v sodobni družbi
- poznavanje konceptov in metodologij za analizo velikih količin podatkov
- poznavanje programskih orodij in metodologij za analizo podatkov ter simulacije diskretnih oziroma zveznih modelov

- knowledge and understanding of a wide range of applications of information communication technology in the modern society
- knowledge of the concepts and methodologies for the analysis of large amounts of data
- knowledge of most important methods in statistical learning
- knowledge of programming tools and methodologies for data analysis and simulation of discrete and continuous models

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, predstavitev)
- *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:

- pisni izpit
- seminarska naloga

Delež (v %) /

Weight (in %)

Assessment:

Type:

- written exam
- seminar work

30
70