

UČNI NAČRT PREDMETA / COURSE SYLLABUS						
Predmet:	Strojno učenje za podatkovne vede I					
Course title:	Machine learning for data science 1					
Študijski program in stopnja Study programme and level	Študijska smer Study field			Letnik Academic year	Semester Semester	
Interdisciplinarni magistrski študijski program Računalništvo in matematika	ni smeri			1 ali 2	prvi ali drugi	
Interdisciplinary Masters study programme Computer Science and Mathematics	none			1 or 2	first or second	
Vrsta predmeta / Course type				izbirni		
Univerzitetna koda predmeta / University course code:				M2854		
Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
45		30			105	6
Nosilec predmeta / Lecturer:			Blaž Zupan			
Jeziki / Languages:	Predavanja / Lectures:	slovenski/Slovene, angleški/English				
	Vaje / Tutorial:	slovenski/Slovene, angleški/English				
Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:				Prerequisites:		
Vsebina:				Content (Syllabus outline):		

<p>Linearni modeli. Linearna regresija.</p> <p>Linearna diskriminantna analiza. Logistična regresija. Gradientni sestop. Stohastični gradientni sestop.</p> <p>Pristop strojnega učenja. Cenovna funkcija. Pristop z zmanjšanjem tveganja.</p> <p>Maksimizacija verjetja. Vrednotenje modelov. Prečno preverjanje.</p> <p>Izbor značilnk. Pristopi z iskanjem.</p> <p>Regularizacija.</p> <p>Drevesni modeli. Klasifikacijska in regresijska drevesa. Naključni gozd. Pristop bagging. Gradientni razvoj niza dreves.</p> <p>Gručenje. Metoda voditeljev. Algoritem EM.</p> <p>Nelinearna regresija. Bazne funkcije.</p> <p>Zlepki. Metoda podpornih vektorjev. Trik z jedri.</p>	<p>Linear models. Linear regression.</p> <p>Linear discriminant analysis. Logistic regression. Gradient descent.</p> <p>Stochastic gradient descent.</p> <p>The machine learning approach.</p> <p>Cost functions. Empirical risk minimization. Maximum likelihood estimation. Model evaluation. Crossvalidation.</p> <p>Feature selection. Search-based feature selection. Regularization.</p> <p>Tree-based models. Decision trees. Random forest. Bagging. Gradient tree boosting.</p> <p>Clustering. k-means. Expectation Maximization.</p> <p>Non-linear regression. Basis functions. Splines. Support vector machines. Kernel trick.</p>
---	---

<p>Nevronske mreže. Perceptron. Aktivacijske funkcije. Tehnika vzratnega razširjanja napak.</p>	<p>Neural networks. Perceptron. Activation functions. Backpropagation.</p>
---	--

Temeljni literatura in viri / Readings:

<p>James G, Witten D, Hastie T, Tibshirani T (2017) An Introduction to Statistical Learning, Springer.</p> <p>Hastie T, Tibshirani R, Friedman J (2003) The elements of statistical learning, Springer.</p>

Cilji in kompetence:

<p>Predmet je namenjen seznanitvi z matematičnimi in algoritimičnimi osnovami strojnega učenja ter vidiki uporabe strojnega učenja pri reševanju praktičnih problemov. Predmet pripravi študente na študij naprednejših metod iz strojnega učenja.</p>
--

Objectives and competences:

<p>The course aims at familiarizing the student with the fundamentals of machine learning, classical machine learning models, and the practicalities of applying machine learning to real-world problems. The course prepares students for the study of advanced machine learning methods.</p>
--

Predvideni študijski rezultati:

<p>Po uspešno zaključenem predmetu naj bi bili študentje zmožni:</p> <ul style="list-style-type: none"> - Uporabiti pristope strojnega učenja k podatkovni analitiki. - Evalvirati različne tehnike

Intended learning outcomes:

<p>After successfully completing the course, students should be able to:</p> <ul style="list-style-type: none"> - Apply the machine learning approach to data analysis. - Evaluate different types of models.

modeliranja.

- Izbrati ustrezno tehniko za dani problem in podatke.
- Interpretirati rezultate strojnega učenja.
- Prepoznati potencialne probleme.

- Choose the correct model for the problem at hand.
- Interpret machine learning results.
- Identify potential issues.

Metode poučevanja in učenja:

Predavanja, vaje, domače naloge in nabor manjših projektov.

Learning and teaching methods:

Lectures, , homework, and a set of smaller projects.

Načini ocenjevanja:

Delež (v %) /
Weight (in %)

Assessment:

Sprotno preverjanje (domače naloge, projekti)		Continuing (homework, projects) Final (written exam)
Končno preverjanje (pisni izpit)	50%	Grading: 6-10 pass, 5 fail
Ocene: 6-10 pozitivno, 5 negativno	50%	

Reference nosilca / Lecturer's references:

Blaž Zupan:

- Staric A, Demsar J, Zupan B (2015) Concurrent software architectures for exploratory data analysis. WIREs Data Mining and Knowledge Discovery 5(4):165-180. [COBISS.SI-ID 1536356035]
- Zitnik M, Nam EA, Dinh C, Kuspa A, Shaulsky G, Zupan B (2015) Gene prioritization

by compressive data fusion and chaining, PLoS Computational Biology

11(10):e1004552. [COBISS.SI-ID 1536572355]

– Stražar M, Žitnik M, Zupan B, Ule J, Curk T (2016) Orthogonal matrix factorization enables integrative analysis of multiple RNA binding proteins, *Bioinformatics* 32(10):1527-35. [COBISS.SI-ID 1537001923]

– Žitnik M, Zupan B (2016) Jumping across biomedical contexts using compressive data fusion, *Bioinformatics* 32(12):i90-i100. [COBISS.SI-ID 1537026243]

– Čopar A, Žitnik M, Zupan B (2017) Scalable non-negative matrix tri-factorization, *BioData Mining* 10:41. [COBISS.SI-ID 1537708995]