

UČNI NAČRT PREDMETA / COURSE SYLLABUS

Predmet: Analiza podatkov s strojnim učenjem
Course title: Data analysis with machine learning

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Informacijska družba, doktorski študijski program tretje stopnje	-	Prvi	Prvi
Information Society, third cycle Doctoral Study Programme	-	First	First

Vrsta predmeta / Course type

Izbirni/ Optional

Univerzitetna koda predmeta / University course code:

1-ID-DR-IP-APSU-2021-02-04

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
20	/	20	/	/	410	15

Nosilec predmeta / Lecturer: doc. dr. Bernard Ženko

Jeziki / Languages:

Predavanja / Lectures: Slovenski / Slovenian, Angleški / English
Vaje / Tutorial: Slovenski / Slovenian, Angleški / English

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

Vpis v prvi letnik študija.

Prerequisites:

Enrolment in the first year of studies.

Vsebina:

V okviru predmeta bo delo usmerjeno k nadgradnji že pridobljenega metodološkega znanja s spoznavanjem osnovnih metod strojnega učenja. Poleg predstavitve teoretskega metodološkega okvira se bo izvajalo tudi samostojno raziskovalno delo študentov z uporabo različnih metod podatkovnega rudarjenja (v okviru vaj in seminarske naloge).

Predvideni so naslednji tematski sklopi:

- postopek analize podatkov;
- osnovne metode strojnega učenja (izbiranje značilk, odločitvena drevesa in pravila, k-najbližjih

Content (Syllabus outline):

The course is focused on upgrading the existing methodology knowledge with basic principles of machine learning. In addition to the presentation of theoretical principles, individual student research work with practical application of different data mining techniques will be required (laboratory and individual seminar work).

The course contains the following themes:

- data analysis procedure;
- basic machine learning methods (feature selection, decision rules and trees, k-nearest neighbors, naive Bayes, ensemble methods, association rules);

- sosedov, naivni Bayes, ansambelske metode, asociacijska pravila);
- vrednotenje modelov.

- model evaluation.

Temeljni literatura in viri / Readings:

- Witten I.H., Eibe F., Hall M.A.: Data mining: Practical machine learning tools and techniques, Fourth Edition, Morgan Kaufman Publishers, 2016.
- Hastie T., Tibshirani R., Friedman J.: The elements of statistical learning: Data Mining, Inference, and Prediction, Springer.
- Kononenko I.: Strojno učenje, Založba FE in FRI, Ljubljana, 2005.
- Mitchell T.: Machine Learning. McGraw-Hill, 1997

Cilji in kompetence:

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

- sposobnost identificiranja danega raziskovalnega problema, njegove analize ter možnih rešitev
- ustvarjanje novega znanja, ki pomeni relevanten prispevek k razvoju znanosti
- sposobnost obvladanja standardnih metod, postopkov in procesov raziskovalnega dela na različnih znanstvenih področjih
- sposobnost za reševanje konkretnih raziskovalnih problemov na posameznih področjih družbenih in ostalih ved
- razvoj veščin in spretnosti v uporabi znanja na raziskovalnem področju doktorske disertacije
- sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih znanj in zmožnost interpretacije v kontekstu družboslovja in ostalih ved
- sposobnost oblikovanja in implementacije izvirnih znanstvenih rešitev danih družbenih problemov

Objectives and competences:

Learning unit contributes to the development of the following general and subject-specific competences:

- the ability to identify, analyze and construct solution for a given research problem
- the creation of new knowledge and contribution to the development of science
- mastery of standard methods, approaches and processes of scientific research in various scientific fields
- skills and abilities for solving concrete research problems in various fields of social and other sciences
- development of skills and abilities in usage of knowledge in doctoral research
- the ability to extract, select, evaluate and insert new knowledges and the competence of interpretation in the context of social and other sciences
- ability of designing and implementing novel scientific solutions for given social problems

Predvideni študijski rezultati:

Znanje in razumevanje:

Študent/študentka:

- demonstrira poznavanje ključnih pojmov strojnega učenja,

Intended learning outcomes:

Knowledge and understanding:

The student:

- demonstrates knowledge of basic machine learning concepts,

- demonstrira praktično obvladovanje osnovnih metod strojnega učenja,
- demonstrira sposobnost samostojnega raziskovalnega dela z uporabo metod strojnega učenja,
- demonstrira sposobnost prezentacije svojih raziskovalnih rezultatov na znanstvenih srečanjih in v znanstvenih publikacijah.

- demonstrates the ability of practical application of basic machine learning methods,
- demonstrates the ability of individual and independent research work with help of machine learning tools,
- demonstrates the ability of presenting the acquired research results in scientific meetings and publications.

Metode poučevanja in učenja:

- *Predavanja* z aktivno udeležbo študentov; razlaga, diskusija, vprašanja, primeri, reševanje problemov.
- *Vaje* s praktičnim izvajanjem predstavljenih metod na konkretnih podatkih.
- *Individualno delo* študentov; samostojni študij znanstvene in strokovne literature, izdelava seminarske naloge in njena ustna predstavitev.

Learning and teaching methods:

- *Lectures* with active participation of students; explanation, discussion, questions, examples, problem solving.
- *Lab work* with practical applications of the presented methods on specific data.
- *Individual work* of students; independent study of scientific and professional literature, writing of the seminar work and its presentation.

Delež (v %) /

Weight (in %)

Načini ocenjevanja:

Assessment:

<p>Način (pisni izpit, ustno izpraševanje, naloge, projekt):</p> <ul style="list-style-type: none"> • Projektna naloga 	<p>100</p>	<p>Type (examination, oral, coursework, project):</p> <ul style="list-style-type: none"> • Project assignment
---	------------	--

Reference nosilca / Lecturer's references:

- STROJNIK, Lidija, STOPAR, Matej, ZLATIĆ, Emil, KOKALJ, Doris, NAGLIČ GRIL, Mateja, ŽENKO, Bernard, ŽNIDARŠIČ, Martin, BOHANEK, Marko, MILEVA BOSKOSKA, Biljana, LUŠTREK, Mitja, GRADIŠEK, Anton, POTOČNIK, Doris, OGRINC, Nives. Authentication of key aroma compounds in apple using stable isotope approach. *Food chemistry*, 2019, vol. 277, str. 766-773, doi: 10.1016/j.foodchem.2018.10.140.
- SIMIDJIEVSKI, Nikola, TANEVSKI, Jovan, ŽENKO, Bernard, LEVNAJIĆ, Zoran, TODOROVSKI, Ljupčo, DŽEROSKI, Sašo. Decoupling approximation robustly reconstructs directed dynamical networks. *New journal of physics*, 2018, 29 str., doi: 10.1088/1367-2630/aae94.
- TUŠAR, Tea, GANTAR, Klemen, KOBLAR, Valentin, ŽENKO, Bernard, FILIPIČ, Bogdan. A study of overfitting in optimization of a manufacturing quality control procedure. *Applied soft computing*, 2017, vol. 59, str. 77-87, doi: 10.1016/j.asoc.2017.05.027.
- GAMBERGER, Dragan, ŽENKO, Bernard, MITELPUNKT, Alexis, SHACHAR, Netta, LAVRAČ, Nada. Clusters of male and female Alzheimer's disease patients in the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. *Brain informatics*, 2016, vol. 3, no. 3, str. 169-179, doi: 10.1007/s40708-016-0035-5.
- NOVAK BABIČ, Monika, ZALAR, Polona, ŽENKO, Bernard, DŽEROSKI, Sašo, GUNDE-CIMERMAN, Nina. Yeasts and yeast-like fungi in tap water and

groundwater, and their transmission to household appliances. *Fungal ecology*, 2016, vol. 20, str. 30-39, doi: 10.1016/j.funeco.2015.10.00.

- GAMBERGER, Dragan, ŽENKO, Bernard, MITELPUNKT, Alexis, LAVRAČ, Nada. Identification of gender specific biomarkers for Alzheimer's disease. V: *Brain informatics and health : 8th International Conference, BIH 2015, London, UK, August 30 - September 2, 2015 : proceedings*, Springer. 2015, LNAI 9250, str. 57-66. http://link.springer.com/chapter/10.1007%2F978-3-319-23344-4_6.
- NOVAK BABIČ, Monika, ZALAR, Polona, ŽENKO, Bernard, SCHROERS, Hans-Josef, DŽEROSKI, Sašo, GUNDE-CIMERMAN, Nina. *Candida and Fusarium species known as opportunistic human pathogens from customer-accessible parts of residential washing machines*. *Fungal biology*, 2015, vol. 119, iss. 2/3, str. 95-113, doi: 10.1016/j.funbio.2014.10.007.
- DEBELJAK, Marko, POLJANEC, Aleš, ŽENKO, Bernard. *Modelling forest growing stock from inventory data : a data mining approach*. *Ecological indicators*, 2014, vol. 41, str. 30-39.
- ŠKRABAN, Jure, DŽEROSKI, Sašo, ŽENKO, Bernard, MONGUS, Domen, GANGL, Simon, RUPNIK, Maja. *Gut microbiota patterns associated with colonization of different clostridium difficile ribotypes*. *PloS one*, 2013, vol. 8, iss. 2, str. e58005-1-e58005-13.
- CAROTENUTO, Marianeve, DŽEROSKI, Sašo, ŽENKO, Bernard, SLAVKOV, Ivica, et al. *Neuroblastoma tumorigenesis is regulated through the Nm23-H1/h-Prune C-terminal interaction*. *Scientific reports*, 2013, vol. 3, str. 1351-1-1351-11.
- AHO, Timo, ŽENKO, Bernard, DŽEROSKI, Sašo, ELOMAA, Tapio. *Multi-target regression with rule ensembles*. *Journal of machine learning research*, 2012, vol. 13, str. 2367-2407.
- ŽENKO, Bernard, DŽEROSKI, Sašo. *Learning classification rules for multiple target attributes*. *Lecture notes in computer science*. Berlin; New York: Springer, cop. 2008, vol. 5012, str. 454-465.
- DŽEROSKI, Sašo, ŽENKO, Bernard. *Is combining classifiers with stacking better than selecting the best one?*. *Machine learning*, 2004, vol. 54, str. 255-273.