

**RTU studiju kurss "Mašīnmācīšanās un datizrace datu analīzei"**

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

**Vispārējā informācija**

Kods	DMI752
Nosaukums	Mašīnmācīšanās un datizrace datu analīzei
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles
Atbildīgais mācītbspēks	Inese Poļaka - Doktors, Asociētais profesors
Apjoms daļās un kredītpunktos	1 daļa, 7.5 kredītpunkti
Studiju kursa īstenošanas valodas	LV, EN
Anotācija	Katrā zinātniskā pētījumā nozīmīgu vietu ieņem dati un to analīze, lai nonāktu līdz pētījuma secinājumiem un tos pamatotu. Studiju kursā apskafīts formalizēts intelektuālās tehnoloģijas balstīts datu analīzes process, kas iekļauj tādus soļus kā uzdevuma formulēšana, datu iegūšana un priekšapstrāde, datu analīze ar mašīnmācīšanās/datizraces/statistikas metodēm, modeļu novērtēšana, modeļu un datu noslieču apzināšana, rezultātu interpretācija un vizualizācija. Pareiza pieeju un metožu izvēle katrā posmā var ietekmēt pētījuma rezultātu pamatotību un ticamību, tāpēc studiju kursā studentam ir iespēja izstrādāt pētījuma datu analīzes prototipu un diskutēt ar mācītbspēkiem un citiem studentiem to pilnveidot, lai sasniegtu sava pētījuma mērķus.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķis ir sniegt studentiem zināšanas par datu analīzes procesu un tā izstrādi pētījumam, ievērojot visus nepieciešamos soļus. Studiju kursa uzdevumi ir iepazīstināt ar populārākām metodēm, kas tiek izmantotas katrā posmā, lai sniegtu priekšstatu par procesu un atbalstītu datu analīzes procesa prototipa izstrādē, lai students spētu šo procesu izstrādāt savam pētījumam.
Patstāvīgais darbs, tā organizācija un uzdevumi	Studiju kursā studenti izstrādā sava (vai hipotētiskā) pētījuma datu analīzes procesa prototipu, t.sk. datu sagatavošanu, analīzes metožu izvēli, modeļu validāciju, rezultātu novērtēšanas pieeju. Studenti praktiski īsteno mašīnmācīšanās/datizraces/statistikos metožu pielietošanu patstāvīgajos darbos. Lai iepazītos ar datu analīzes aktualitātēm savā pētniecības jomā, studenti veic literatūras analīzi par izmantotajām metodēm un potenciālajām problēmām.
Literatūra	<b>Obligātā:</b> 1. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition New York, USA: Springer-Verlag, 2016 2. Witten I. H., Frank E., Hall M. A., Pal C. J.. Data Mining: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) 4th Edition Burlington, MA, USA: Morgan Kaufmann Publishers, 2016 <b>Papildu:</b> 1. EMC Education Services (Ed.). Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Hoboken, USA: Wiley 2015 2. Grus, J.. Data Science from Scratch, 2nd Edition. Newton, USA: O'Reilly Media, 2019 3. Bishop, C. M.. Pattern Recognition and Machine Learning (Information Science and Statistics) New York, USA: Springer-Verlag, 2006 4. Geron, A.. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition Newton, USA: O'Reilly Media, 2017
Nepieciešamās priekšzināšanas	Kopu teorijas pamata koncepti. Vairāku teorijas pamati. Mākslīgā intelekta pamati.

**Studiju kursa saturs**

Saturs	Pilna un nepilna laika klātienē studijas		Nepilna laika neklātienē studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Datu analīzes procesa apskats, posmi.	4	4	0	0
Uzdevuma definēšana un mašīnmācīšanās datu analīzes pieejas (uzraudzītā, neuzraudzītā, stimulētā mācīšanās).	4	8	0	0
Datu iegūšana un sagatavošana.	8	16	0	0
Literatūras apskats. Aktuālo mašīnmācīšanās/datizraces/statistikas metožu patstāvīga izpēte un salīdzinošā analīze, diskusija.	4	10	0	0
Datu analīzes metožu izvēle atbilstoši uzdevumam un implementācija.	4	20	0	0
Datu analīzes procesa prototipa izveide.	4	20	0	0
Mašīnmācīšanās modeļu noslieces un citi potenciālie trūkumi datu analīzes procesā.	4	6	0	0
Datu analīzes (mašīnmācīšanās/datizraces) rezultātu statistiskā analīze.	6	10	0	0
Vizualizācija.	4	6	0	0
Programmatūras rīku izmantošana.	8	20	0	0
Semināri un pārbaudījumi.	10	20	0	0
<b>Kopā:</b>	<b>60</b>	<b>140</b>	<b>0</b>	<b>0</b>

**Sasniedzamie studiju rezultāti un to vērtēšana**

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Pārzina datu analīzes procesu, tā posmus un potenciālās kļūdas.	Patstāvīgais darbs, izveidojot datu analīzes procesa prototipu. Eksāmens.
Spēj identificēt datizraces un mašīnmācīšanās metodes konkrētam analīzes uzdevumam un tās pielietot.	Mājas darbs, pielietojot metodes.
Izprot datu analīzes metožu algoritmus un to attīstību.	Literatūras izpētes atskaite par aktuālajām metodēm noteiktam pielietojumam.
Spēj novērtēt un interpretēt datu analīzes rezultātus.	Mājas darbs, pielietojot metodes. Eksāmens.

**Studiju rezultātu vērtēšanas kritēriji**

Kritērijs	% no kopējā vērtējuma
Datu analīzes procesa prototipa izveide	30
Literatūras analīze par aktuālām metodēm	10
Mājas darbs, pielietojot datizraces/mašīnmācīšanās metodes	30
Eksāmens	30
Kopā:	100

**Studiju kursa plānojums**

Daļa	KP	Stundas			Pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs
1.	7.5	2.0	5.0	0.0		*	