

RTU studiju kurss "Ievads datu analītikā"

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

Vispārējā informācija

Kods	DE0914
Nosaukums	Ievads datu analītikā
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles; Brīvās izvēles
Atbildīgais mācītbspēks	Ilze Birzniece - Doktors, Docents
Apjoms daļās un kredītpunktos	1 daļa, 3.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV
Anotācija	Datu analītika ir starpdisciplināra nozare, kas balstās uz matemātiku, statistiku, datorzinātni un datizrāci, lai no datiem iegūtu potenciāli vērtīgas zināšanas. Datizrāce (data mining) ir mašīnmācīšanās (machine learning) metožu lietošana lielām datu kopām, lai atklātu tajās sakarības un skaidrojumu vai prognozi uz vēsturisko datu pamata. Studiju kurss iepazīstina ar datizrāces iespējām datu analītikā un zināšanu ieguvē no datiem, apskatot arī citus būtiskus datu analītikas posmus – datu priekšapstrādi un iegūto rezultātu analīzi lēmumu pieņemšanā. Studiju kursā tiek apskatītas galvenās datizrāces pieejas pārraudzītajā (supervised) un nepārraudzītajā (unsupervised) apmācībā – regresija, klasifikācija, klasterēšana un asociatīvo likumu meklēšana, iepazīstinot ar populārākajām metodēm katrā no tām. Apgūtās metodes tiek praktiski izmantotas datizrāces uzdevumu veikšanai gan ar bezkoda rīkiem, gan programmējot.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķis ir attīstīt studējošo izpratni par datu analītikas iespējām un prasmes lietot konkrētam uzdevumam atbilstošas datizrāces pieejas un rīkus. Studiju kursa uzdevumi: - iepazīstināt studējošos ar mūsdienu datu analītikas iespējām, lietojot datizrāces pieejas; - veidot izpratni par datu ieguves un apstrādes procesiem, lai iegūtu datus balstītas zināšanas; - attīstīt prasmes darbā ar datizrāces metodēm un rīkiem; - veicināt analītiskās spējas, kritisko domāšanu un argumentāciju lēmumu pieņemšanas procesā.
Patstāvīgais darbs, tā organizācija un uzdevumi	Patstāvīgajā darbā studenti veic atsevišķus praktiskos darbus, kā arī apgūst daļu teorētiskā materiāla, pirms klātienē praktiskās nodarbības iepazīstoties ar videomateriāliem, literatūru un veicot pašpārbaudes testus.
Literatūra	Obligātā. / Obligatory: Witten, Ian H.. Data mining : practical machine learning tools and techniques /Ian H. Witten, Eibe Frank, Mark A. Hall. 2016, xxxiii, 629 lpp. : il. ; 24 cm. Datu ieguve : pamati/ A.Sukovs, L.Aleksejeva, K.Makejeva, A.Borisovs ; RTU. Datorzinātnes un informācijas tehnoloģijas fak. Informācijas tehnoloģijas inst. ; rec. I.Jackiva, Z.Markovičs. Rīga : Drukātava, 2007., 130 lpp. : il., tab. Papildu. / Additional: Albright, S. Christian., Business analytics : data analysis and decision making /S. Christian Albright, Wayne L. Winston. 2020, xxiii, 882 lpp. : ilustrācijas ; 29 cm Gordon, Mary Ellen. Business analytics : combining data, analysis & judgement to inform decisions /Mary Ellen Gordon. 2023, xiv, 219 lpp. : ilustrācijas ; 24 cm
Nepieciešamās priekšzināšanas	Pamatzināšanas datu apstrādē ar lietojumprogrammām, pamatzināšanas Python programmēšanas valodā.

Studiju kursa saturs

Saturs	Pilna un nepilna laika klātienē studijas		Nepilna laika neklātienē studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Mūsdienu datu analītika: statistika, mašīnmācīšanās, biznesa vajadzību noteikšana.	8	4	0	0
Datu priekšizpēte: datu tipi, vizualizācija, priekšapstrāde.	10	6	0	0
Datizrāces pieejas: regresija, klasifikācija, klasterēšana, asociatīvo likumu meklēšana. Rezultātu novērtēšanas metodes. Mūsdienīgi datizrāces lietojumi.	10	10	0	0
Datizrāces metožu praktiska apguve ar Weka.	8	8	0	0
Datizrāces metožu praktiska apguve ar Google Colab.	4	12	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
Prot veikt datu ielasi, nepieciešamo priekšapstrādi, vizualizāciju un statistikas izvadi, interpretējot iegūtos rezultātus.	Izpildīts praktiskais darbs. Students demonstrē izpratni par datu priekšapstrādes uzdevumiem un prasmes to izpildē ar atbilstošiem līdzekļiem.
Spēj risināt lineārās regresijas uzdevumu.	Izpildīts praktiskais darbs. Students prot atrisināt lineārās regresijas uzdevumu, izmantojot atbilstošus programmatūras līdzekļus.
Spēj risināt klasifikācijas uzdevumu ar dažādiem algoritmiem un rīkiem, novērtēt iegūtos rezultātus.	Izpildīti praktiskie darbi. Students prot atrisināt klasifikācijas uzdevumu, izmantojot atbilstošus programmatūras līdzekļus, kā arī demonstrē kompetenci iegūto rezultātu interpretēšanā.
Spēj risināt asociatīvo likumu atrašanu ar dažādiem rīkiem, novērtēt iegūtos rezultātus.	Izpildīti praktiskie darbi. Students prot atrast asociatīvos likumus, izmantojot atbilstošus programmatūras līdzekļus, kā arī demonstrē kompetenci iegūto rezultātu interpretēšanā.
Spēj risināt klasterēšanas uzdevumu ar dažādiem algoritmiem un rīkiem, novērtēt iegūtos rezultātus.	Izpildīti praktiskie darbi. Students prot atrisināt klasterēšanas uzdevumu, izmantojot atbilstošus programmatūras līdzekļus, kā arī demonstrē kompetenci iegūto rezultātu interpretēšanā.

Studiju rezultātu vērtēšanas kritēriji

Kritērijs	% no kopējā vērtējuma
Sekmīgi izpildīti vismaz 6 no 8 praktiskajiem darbiem	100
Kopā:	100

Studiju kursa plānojums

Daļa	KP	Stundas			Pārbaudījumi			Brīvās izvēles pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs	Ieskaite	Eksām.	Darbs
1.	3.0	20.0	20.0	0.0	*			*		