

RTU studiju kurss "Datu zinātne un lielo datu tehnoloģijas"

0R000 Rēzeknes akadēmija

Vispārējā informācija

Kods	RA0755
Nosaukums	Datu zinātne un lielo datu tehnoloģijas
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles
Atbildīgais mācītbspēks	Pēteris Grabusts - Doktors, Profesors
Apjoms daļās un kredītpunktos	1 daļa, 6.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV
Anotācija	Kursā tiek dotas zināšanas par lielo datu analītikas iespējām, datu analītikas nozīmīgumu un pamatprincipiem. Apskatītas populārākās datizraces metodes, to iespējas datu ieguvē un tālākajā analizē. Dots priekšstats par datu vizualizāciju un lielo datu glabāšanas un apstrādes rīkiem.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Kursa mērķis ir sniegt informāciju un praktisko pieredzi par izmantotajiem algoritmiem, metodēm un rīkiem lielo datu apstrādē. Iegūtās kompetences: prot pielietot lēmumu analīzes metodes praktiskā datu apstrādē. Iegūtās prasmes: spēja noteikt un izmantot nepieciešamos rīkus konkrētas darbības veikšanai datu analītikā un datu organizēšanā.
Patstāvīgais darbs, tā organizācija un uzdevumi	Referāta izstrāde par jebkādu lielo datu apstrādes aspektu; Azure rīka apgūšana un praktiskā uzdevuma realizēšana; Apache programmatūras komplekta Hadoop iepēju apguve.
Literatūra	Obligātā literatūra 1. Tom White. "Hadoop: The Definitive Guide", 2015 2. Mahmoud Parsian. "Data Algorithms: Recipes for Scaling Up with Hadoop and Spark", 2015 3. Matei Zaharia, Holden Karau, Andy Konwinski, Patrick Wendell. "Learning Spark", 2015 4. Sandy Ryza, Uri Laserson. "Advanced Analytics with Spark: Patterns for Learning from Data at Scale", 2015 5. Lars George. "HBase: The Definitive Guide", 2011 6. https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html Papildliteratūra 1. Viktor Mayer-Schönberger, Kenneth Cukier. "Big Data: A Revolution That Will Transform How We Live, Work, and Think", 2014 2. Nathan Martz, James Warren. "Big Data: Principles and best practices of scalable realtime data systems", 2015 3. Jeroen Janssens. "Data Science at the Command Line", 2014
Nepieciešamās priekšzināšanas	Datu bāzes. Mākslīgā intelekta pamati.

Studiju kursa saturs

Saturs	Pilna un nepilna laika klātienēs studijas		Nepilna laika neklātienēs studijas	
	Kontakst stundas	Patstāv. darbs	Kontakst stundas	Patstāv. darbs
Lielo datu analītikas iespējas	6	7	0	0
Datizraces metodes un paņēmieni	6	7	0	0
Dziļās mašīnmācīšanās metodes	5	8	0	0
Mākslīgo neironu tīkli, Azure rīks neironu tīklu aprakstīšanai	5	8	0	0
Lielo datu glabāšanas un apstrādes ietvars Hadoop	5	8	0	0
Datizraces metožu un algoritmu pielietošana Spark	5	8	0	0
Lielu grafu datu analīzes pakete Giraph	5	8	0	0
Grafu analīzes valoda. Grafu analīzes piemēri	5	8	0	0
Procesu pārvaldības rīks Zookeeper	5	8	0	0
Zookeeper uzbūve. Ziņojumu apmaiņas sadalītos procesos implementācija ar Zookeeper	5	8	0	0
Lielo datu vizualizācija.	5	8	0	0
Datu vizualizācijas nozīmīgums. Datu vizualizācijas metodes.	5	8	0	0
Kopā:	62	94	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

Sasniedzamie studiju rezultāti	Rezultātu vērtēšanas metodes
--------------------------------	------------------------------

Zināšanas Zināšanas par lielo datu īpašībām, veidiem, to iegūšanu, vākšanu, datu analītikas mehānismiem, kā arī par datu pārvaldību, datu stratēģijas īstenošanu un pārskatīšanu Zināšanas par dažādām lielo datu iegūšanas (datizraces) metodēm Zināšanas par datu pārvaldību un tās dažādām aplikācijām Pārzina dažādas datu vizualizācijas metodes Sasniegta izpratne datu analīzes nozīmīgumu, datu analītikas pamatprincipiem	Izstrādāts un prezentēts referāts. Nokārtots eksāmens.
Prasmes Spēja noteikt un izmantot nepieciešamos rīkus konkrētas darbības veikšanai datu analītikā un datu organizēšanā. Prasme pielietot dažādas lielo datu iegūšanas (datizraces) metodes Iegūta prasme pielietot dažādas datu vizualizācijas metodes	Apgūts Azure rīks, realizēts praktiskais uzdevums. sekmīgi nokārtots eksāmens.
Kompetence Prot pielietot lēmumu analīzes metodes praktiskā lielo datu apstrādē Pārzina dažādas datu (t. sk. lielo datu) iegūšanas, uzglabāšanas, apstrādes, analīzes un vizualizācijas koncepcijas un teorijas, veidus, formas un modeļus, kā arī attiecīgos datu apstrādes instrumentus un to lietošanas iespējas	Apache programmatūras komplekta Hadoop iepēju apguve. Nokārtots eksāmens.

Studiju rezultātu vērtēšanas kritēriji

Kritērijs	% no kopējā vērtējuma
Dalība un aktivitātes lekcijās	60
Eksāmens	40
Kopā:	100

Studiju kursa plānojums

Daļa	KP	Stundas			Pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs
1.	6.0	31.0	31.0	0.0		*	