

RTU studiju kurss "Dabiskās valodas apstrādes pamati"

33000 Datorzinātnes, informācijas tehnoloģijas un enerģētikas fakultāte

Vispārējā informācija

Kods	DE1080
Nosaukums	Dabiskās valodas apstrādes pamati
Studiju kursa statuss programmā	Obligātais/Ierobežotās izvēles; Brīvās izvēles
Atbildīgais mācītājs	Gints Jēkabsons - Doktors, Docents
Apjoms daļās un kredītpunktos	1 daļa, 6.0 kredītpunkti
Studiju kursa īstenošanas valodas	LV, EN
Anotācija	Dabiskās valodas apstrāde ir starpdisciplināra zinātnes nozare, kas pēta metodes un modeļus, kuri ļauj veikt dabiskās valodas datorizētu apstrādi. Studiju kursā studējošie apgūst dabiskās valodas apstrādes pamatjēdzienus, zināšanas par dažādām metodēm un to praktiskiem pielietojumiem, kā arī prasmes ar tām iegūto rezultātu novērtēšanai un interpretēšanai. Praktisku uzdevumu risināšanai studējošie apgūst prasmes izmantot dabiskās valodas apstrādes programmatūru.
Mērķis un uzdevumi, izteikti kompetencēs un prasmēs	Studiju kursa mērķis ir sniegt studējošajiem teorētiskās un praktiskās zināšanas dabiskās valodas apstrādē un tās praktiskajos pielietojumos. Studiju kursa uzdevumi: - veicināt studējošo izpratni par dabiskās valodas apstrādes pamatjēdzieniem un tās risinātajiem uzdevumiem; - attīstīt studējošo prasmes identificēt un izvēlēties konkrētu uzdevumu risināšanai piemērotas dabiskās valodas apstrādes metodes; - uzlabot studējošo prasmes šīs metodes praktiski lietot, kā arī novērtēt un interpretēt ar tām iegūtos rezultātus; - attīstīt studējošo prasmes lietot konkrētām dabiskās valodas apstrādes problēmām atbilstošu programmatūru.
Patstāvīgais darbs, tā organizācija un uzdevumi	Patstāvīgais darbs ietver individuālu teorētiskas un praktiskas ievirzes uzdevumu risināšanu, izmantojot studiju kursā iegūtās zināšanas par dabiskās valodas apstrādi, tās metodēm un atbilstošiem programmatūras līdzekļiem.
Literatūra	Obligātā / Obligatory: 1. Steven Bird, Ewan Klein, Edward Loper. Natural Language Processing with Python (http://www.nltk.org/book/) 2. Jay Alamar, Maarten Grootendorst. Hands-On Large Language Models: Language Understanding and Generation, 2024, 425 p. (https://www.llm-book.com/) 3. Dan Jurafsky, James H. Martin. Speech and Language Processing 3rd ed. draft Prentice Hall, 2021, 1032 p. (http://web.stanford.edu/~jurafsky/slp3/) Papildu / Additional: 1. Al Sweigart. Automate the Boring Stuff with Python: Practical Programming for Total Beginners No Starch Press, 2015 (https://automatetheboringstuff.com/) 2. Alexander Clark, Chris Fox, Shalom Lappin (eds.). The Handbook of Computational Linguistics and Natural Language Processing Wiley-Blackwell, 2010, 802 p. 3. Anne O'Keeffe, Michael McCarthy (eds.). The Routledge Handbook of Corpus Linguistics London: Routledge, 2010, 682 p.
Nepieciešamās priekšzināšanas	Python programmēšanas pamati. Angļu valodas prasmes, lai lasītu obligāto literatūru.

Studiju kursa saturs

Saturs	Pilna un nepilna laika klātienes studijas		Nepilna laika neklātienes studijas	
	Kontakt stundas	Patstāv. darbs	Kontakt stundas	Patstāv. darbs
Dabiskās valodas apstrādes pamatjēdzieni.	4	8	0	0
Teksta priekšapstrāde. Regulārās izteiksmes, teksta segmentēšana, lemmatizēšana, celmošana, morfoloģiskā marķēšana u. c.	10	15	0	0
Dabiskās valodas apstrādes metodes un praktiskie pielietojumi, kas saistīti korpusa analīzi, tematu modelēšanu, teksta klasificēšanu, noskaņojuma analīzi u. c.	18	27	0	0
Lielie valodas modeļi un ar tiem saistītie praktiskie pielietojumi.	12	18	0	0
Dabiskās valodas apstrādes algoritmu realizēšana programmatūrā.	14	21	0	0
Iegūto rezultātu novērtēšana un interpretēšana.	6	9	0	0
Kopā:	64	98	0	0

Sasniedzamie studiju rezultāti un to vērtēšana

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

Pārzina dabiskās valodas apstrādes pamatus.	Individuālie uzdevumi, eksāmens. Kritēriji: students demonstrē izpratni par jomas terminoloģiju, kā arī jomas problēmu veidiem un to atbilstošajiem konceptuālajiem risinājumiem.
Spēj identificēt un izvēlēties piemērotas metodes konkrētas problēmas risināšanai dabiskās valodas apstrādē.	Individuālie uzdevumi, eksāmens. Kritēriji: students demonstrē kompetenci un prasmi risināt jomas tipiskas problēmas ar atbilstošām metodēm.
Prot lietot dabiskās valodas apstrādes metodes, kā arī novērtēt un interpretēt ar tām iegūtos rezultātus.	Individuālie uzdevumi, eksāmens. Kritēriji: students demonstrē izpratni par jomas tipisko metožu būtību un to atbilstošu lietošanu, kā arī demonstrē kompetenci iegūto rezultātu tālākā apstrādē.
Prot izstrādāt un lietot programmatūru problēmu risināšanai dabiskās valodas apstrādē.	Individuālie uzdevumi, eksāmens. Kritēriji: students demonstrē spējas izstrādāt un izmantot risināmajai problēmai atbilstošu programmatūru.

Studiju rezultātu vērtēšanas kritēriji

Kritērijs	% no kopējā vērtējuma
Individuālie uzdevumi	80
Eksāmens	20
Kopā:	100

Studiju kursa plānojums

Daļa	KP	Stundas			Pārbaudījumi			Brīvās izvēles pārbaudījumi		
		Lekcijas	Prakt d.	Laborat	Ieskaite	Eksām.	Darbs	Ieskaite	Eksām.	Darbs
1.	6.0	32.0	0.0	32.0		*			*	