

**RTU studiju kurss "Datu zinātne un lielo datu apstrāde"****OR000 Rēzeknes akadēmija****Vispārējā informācija**

|                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kods                                                | RA0137                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Nosaukums                                           | Datu zinātne un lielo datu apstrāde                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Studiju kursa statuss programmā                     | Obligātais/Ierobežotās izvēles                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Atbildīgais mācītspēks                              | Iluta Arbidāne - Doktors, Profesors                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Mācītspēks                                          | Vaida Bartkute-Norkuniene - Profesors                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Apjoms daļās un kredītpunktos                       | 1 daļa, 6.0 kredītpunkti                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Studiju kursa īstenošanas valodas                   | LV, EN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| Anotācija                                           | Datu zinātne ir jauns, eksponenciāli augošs lauks, kas sastāv no instrumentu un paņēmieni kopuma, ko izmanto noderīgas informācijas iegūšanai no datiem. Datu zinātne un lielo datu analīze ir starpdisciplinārs, uz problēmu risināšanu orientēts priekšmets, kas iemācās pielietot zinātniskās metodes praktiskām problēmām. Kurss orientējas uz praktiskām nodarbībām un pašmācību datu kopu sagatavošanas un datu analīzes uzdevumu programmēšanas laikā. Šajā kursā tiks apskatīti lielo datu pamatjēdzieni, strukturētu un nestrukturētu datu analīzes metodoloģijas, uzsvāru liekot uz saistību starp datu zinātnieku un biznesa vajadzībām.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Mērķis un uzdevumi, izteikti kompetencēs un prasmēs | Šī kursa mērķis ir apgūt datu analīzes koncepcijas un paņēmienus, kas atvieglo lēmumu pieņemšanu no bagātīgas datu kopas. Studenti izpētīs datu jēdzienus, metadatu izveidošanu un interpretāciju, vispārējo lineāro metodi, klasteru analīzi un informācijas vizualizācijas pamatus. Studenti tiks pakļauti arī daudziem dažādiem datu zinātnes pieejas pielietojumiem.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Patstāvīgais darbs, tā organizācija un uzdevumi     | Mājas darbu izstrāde<br>Referāta izstrāde<br>Gatavošanās praktiskajām nodarbībām<br>Literatūras un gadījumu izpēte                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Literatūra                                          | <p>Obligātā/Obligatory:</p> <ol style="list-style-type: none"> <li>1.P. Bruce, A. Bruce (2017). Practical Statistics for Data Scientists. Published by O'Reilly Media, Inc.</li> <li>2.Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal. (2016). Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann; 4 edition</li> <li>3.R. Kitchin (2014). The Data Revolution. Big Data, Open Data, Data Infrastructures and Their Consequences</li> <li>4.J. Leskovek, A. Rajaraman and J. Ullman (2014). Mining of Massive Datasets. v2.1, Cambridge University Press. (free online: <a href="http://infolab.stanford.edu/~ullman/mmds/book.pdf">http://infolab.stanford.edu/~ullman/mmds/book.pdf</a>)</li> <li>5.F. Provost, T. Fawcett (2013). Data Science for business.</li> </ol> <p>Papildu/Additional:</p> <ol style="list-style-type: none"> <li>1.B. Steele, J. Chandler, S. Reddy. Algorithms for Data Science (2016). Springer; 1st ed. 2016 edition.</li> <li>2.V. Ankam (2016). Big Data Analytics. Publishing by Packt Publishing Ltd</li> <li>3.J. M. Stanton (2013). Introduction to Data Science. <a href="http://surface.syr.edu/istpub/165/">http://surface.syr.edu/istpub/165/</a></li> </ol> <p>Interneta resursi/Internet resources:</p> <ol style="list-style-type: none"> <li>1.100+ INTERESTING DATA SETS FOR STATISTICS, <a href="https://rs.io/100-interesting-data-sets-for-statistics/">HTTPS://RS.IO/100-INTERESTING-DATA-SETS-FOR-STATISTICS/</a></li> <li>2.19 FREE PUBLIC DATA SETS FOR YOUR DATA SCIENCE PROJECT, <a href="https://www.springboard.com/blog/free-public-data-sets-data-science-project/">HTTPS://WWW.SPRINGBOARD.COM/BLOG/FREE-PUBLIC-DATA-SETS-DATA-SCIENCE-PROJECT/</a></li> <li>3.BIG DATA TOOLS AND TECHNIQUES, <a href="http://web.stanford.edu/class/CS102/datasets.htm">HTTP://WEB.STANFORD.EDU/CLASS/CS102/DATASETS.HTM</a></li> <li>4.TUTORIAL - LEARN PYTHON IN 10 MINUTES, <a href="https://www.learnpython.org/">HTTPS://WWW.LEARNPYTHON.ORG/</a>, <a href="https://www.stavros.io/tutorials/python/r-for-data-science/">HTTPS://WWW.STAVROS.IO/TUTORIALS/PYTHON/R FOR DATA SCIENCE</a>, <a href="https://r4ds.had.co.nz/">HTTPS://R4DS.HAD.CO.NZ/</a></li> </ol> |
| Nepieciešamās priekšzināšanas                       | Nepieciešamas labas matemātiskās zināšanas un programmēšanas prasmes, kas ir pietiekamas jaunu valodu un programmatūras apguvei. Pamatzināšanas par statistiku, lineārā algebra būs kā papildus plus.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

**Studiju kursa saturs**

| Saturs                                                                                                                    | Pilna un nepilna laika klātienēs studijas |                | Nepilna laika neklātienēs studijas |                |
|---------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|----------------|------------------------------------|----------------|
|                                                                                                                           | Kontakt stundas                           | Patstāv. darbs | Kontakt stundas                    | Patstāv. darbs |
| Ievads datu zinātnē                                                                                                       | 4                                         | 14             | 0                                  | 0              |
| Pētījuma datu analīze un datu zinātnes process                                                                            | 6                                         | 12             | 0                                  | 0              |
| Ievads mašīnmācībā (pamata mašīnmācīšanās algoritmi: lineārā regresija; k-tuvākie kaimiņi (k-NN); k-līdzekļi)             | 6                                         | 14             | 0                                  | 0              |
| Objektu ģenerēšana un funkciju izvēle (nozīmes iegūšana no datiem)                                                        | 6                                         | 10             | 0                                  | 0              |
| Relāciju datu bāzes (datu bāzes un relāciju pieeja, ievads SQL vaicājumu valodā un pamata analītika, datu bāzu algoritmi) | 6                                         | 18             | 0                                  | 0              |

|                                                                                                                               |    |     |   |   |
|-------------------------------------------------------------------------------------------------------------------------------|----|-----|---|---|
| Lielie dati un to ietekme uz analītiku. Gadījumu izpēte finanšu pakalpojumu nozarē. Programmēšana Python, rīki datu analīzei. | 6  | 14  | 0 | 0 |
| Lielo datu glabāšana un izguve: noSQL, GraphDB. Lielo datu izkļaidēšana skaitļošanā: karte - samaziniet, dzirksteļojiet       | 6  | 14  | 0 | 0 |
| Datu vizualizācija                                                                                                            | 6  | 14  | 0 | 0 |
| Kopā:                                                                                                                         | 46 | 110 | 0 | 0 |

#### ***Sasniedzamie studiju rezultāti un to vērtēšana***

| Sasniedzamie studiju rezultāti                                                                                                                                                                                                                                                                                                                                                                      | Rezultātu vērtēšanas metodes                                                 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| <b>Zināšanas:</b><br>- Studenti zinās, kādas statistiskās analīzes metodes izvēlēties, ņemot vērā īpašās secinājumu prasības un pieejamos datus.<br>- Students iegūs praktiskās iemaņas, kas vajadzīgas mūsdienu datu analītikā.<br>- Students iegūs lietišķo pieredzi ar datu zinātnes programmatūru, programmēšanu, lietojumprogrammām un procesiem.                                              | Mājas darbi<br>Laboratorijas darbi<br>Eksāmens                               |
| <b>Prasmes:</b><br>- Students spēj pielietot un attīstīt klasterizācijas, klasifikācijas un mašīnmācīšanās pētījumu metodes.<br>- Students iegūs praktisku pieredzi ar reālās pasaules datu analīzi.<br>- Students spēj aprakstīt reālās pasaules problēmas datu zinātnes jomā.<br>- Students iegūs nepieciešamās zināšanas datu zinātnē, kas ir pietiekamas jaunu metožu izstrādei citās zinātnēs. | Mājas darbi<br>Laboratorijas darbi<br>Eksāmens                               |
| <b>Kompetence:</b><br>Students spēj identificēt informāciju un matemātiskos aspektus finanšu pētījumos; novērtēt izmantoto metožu pareizību un pielietojamību katrā pašreizējā situācijā.                                                                                                                                                                                                           | Mājas darbi<br>Laboratorijas darbi<br>Gadījumu izpētes vingrinājumi Eksāmens |

#### ***Studiju rezultātu vērtēšanas kritēriji***

| Kritērijs                       | % no kopējā vērtējuma |
|---------------------------------|-----------------------|
| Mājasdarbi.                     | 20                    |
| Referāts.                       | 20                    |
| Dalība praktiskajās nodarbībās. | 30                    |
| Eksāmens.                       | 30                    |
| Kopā:                           | 100                   |

#### ***Studiju kursa plānojums***

| Daļa | KP  | Stundas  |          |         | Pārbaudījumi |        |       |
|------|-----|----------|----------|---------|--------------|--------|-------|
|      |     | Lekcijas | Prakt d. | Laborat | Ieskaite     | Eksām. | Darbs |
| 1.   | 6.0 | 24.0     | 24.0     | 0.0     |              | *      |       |