VIDZEMES AUGSTSKOLA

FACULTY OF ENGINEERING
STUDY COURSE DESCRIPTION

Course Title:	**INTRODUCTION to DATA SCIENCE**				
Course code (VAIS):	*The course will be registered in the study administration system after accreditation*				
Study programme:	**Information technologies**				
Level of Study programme:	☐ 1st level professional higher education				
	☒ Professional Bachelor				
	☐ Professional Master				
	☐ PhD level				
Type of Study programme:	☒ Compulsory course (Part A)				
	☐ Professional specialization courses (Part B, compulsory)				
	☐ Professional specialization optional courses (Part B, optional)				
	☐ Elective courses (Part C)				
Course Workload:	**Credits**	**ECTS**	**Academic hours**	**Contact hours**	**Independent work hours**
FT (in LV: **PL**)	4	6	160	64	96
PT (in LV: **NL**)	4	6	160	20	140
Course Author/ Tutor:	**Kaspars Osis**				
	Assoc. Prof., Dr.sc.ing.				
	kaspars.osis@va.lv				
	Consultation: according to the schedule for each semester or per individual agreement.				
Study Form:	Full time (FT), Part time (PT)				
Study year, semester:	3rd year, 2nd semester				
Language:	Latvian, English				
Prerequisites for the Course:	Basic knowledge and experience in Python programming language (study course: Introduction to Python programming and data exploration); and study course "Statistics in Engineering"; or relevant courses; or agreement with study course instructor.				
Course Summary:	The study course provides knowledge about data science fundamental concepts; provides basic and in some areas more advanced knowledge and skills about applicable solutions, algorithms and tools to extract and generalize knowledge from data. Students will acquire an integrated set of skills encompassing data processing, supervised and unsupervised data processing models' learning solutions, visualization application, as well as understanding of the synthesis of these skills and ability to apply them for solving real-life problems. Wide spectrum of applicable solutions and algorithms is covered in this study course as well as real-life tasks related data sets are used.				
Course Methods:	Lectures, practical activities, group work, theory tests, final assessment (project work assignment) etc.				
Assessment:	Examination (project work assignment)				
Requirements for Credits:	1. Successful completion of workshops/practical work assignments (at least 60% points of totally available). 2. Passed theoretical tests. 3. Successful completion of project work assignment (at least 65% points of totally available). Final assessment consists of: workshops/practical work assignments, group work evaluations; theoretical tests; project work assignment and project work assignment presentation.				

VIDZEMES AUGSTSKOLA

All practical work assignments have to be accepted (i.e. at least with 60% evaluation) in order to get the final evaluation within this course. 360 points system is used to come up with final evaluation. Table below lists totally available points for each activity.

Work assignment or activity	**Points**
Practical work assignments	215
Theoretical tests	40
Active participation in study activities	20
Project work assignment (exam)	70
Project work assignment presentation (exam)	15
Total	**360**

Final course evaluation (mark) calculation based on 360 points system is done as it follows below:

>= 93% (334-points) = 10 >= 75% (270-points) = 6
>= 90% (324-points) = 9 >= 70% (252-points) = 5
>= 85% (306-points) = 8 >= 65% (234-points) = 4
>= 80% (288-points) = 7 < 65% (234-points) = 3

Missing practical work assignment deadline: each missed day counts for subtraction of 5% from totally available points. It is required to acquire at least 60% from totally available points (not counting potential delay) in order to accept practical work assignment as done. There is provided a template which must be used for documenting practical work assignments – otherwise practical work assignment is not accepted for evaluation!

Abiding by the Academic Ethics

Students must abide by the academic and research ethics, Vidzeme University of Applied Sciences Ethics Regulations, incl.:
- study papers must be independently developed;
- the study work should reference all statements, ideas and data used that have been authored by someone else;
- appropriate data acquisition methods should be used in the acquisition of data, the research ethics must be respected, empirical data must be collected independently and cannot be distorted or falsified;
- the examination must be carried out by the student independently, without the use of supporting materials and/or consultations with other students, unless the lecturer states otherwise.

In the event of non-compliance with the academic and research ethics, punishment is imposed in accordance with the ViA Ethics Regulations and the study course must be re-taken, unless the punishment is extramarital.

Learning Outcomes; the evaluation methods and criteria

Learning Outcomes	**The evaluation methods and criteria**
Knowledge	
Knowledge on data science application necessity and accordant cases.	Development of particular data science solution. Passed theoretical test.
Knowledge about data acquisition and processing within context of data science solutions.	Deploying data processing approaches within development of particular data science solution. Passed theoretical test.
Knowledge about visualization application within context of data science solutions development process.	Deploying visualization approach(es) within development of particular data science solution. Passed theoretical test.
Knowledge about supervised and unsupervised data processing models' learning solutions within context of data science.	Development of particular data science solution. Passed theoretical test.
Skills	

	To develop data science based solutions from several type of real-life tasks and related data sets perspective.	Developed practical group work.
	To develop entry level Python and MS Power BI based data science solutions.	Developed practical group work.
	To develop medium level Python and MS Power BI based data science solutions.	Developed practical group work.
	Competency	
	Use correct data science, Python and MS Power BI solutions terminology. To choose appropriate technological approaches for particular assignment implementation.	Course project development and presentation.
	Independently perform Python and MS Power BI based data science solutions design and development.	Course project development and presentation.
	To solve Python and MS Power BI based data science solutions basic issues.	Course project development and presentation.
Course Compulsory literature:	1. P. Deitel, H. Deitel. Intro to Python for Computer Science and Data Science: Learning to Program with AI, Big Data and the Cloud, Pearson Education, 2020. 2. G.Shmueli, P.C.Bruce, P.Dedeck, N.R.Patel. Data Mining for Business Analytics: Concepts, Techniques, and Applications in Python, Wiley-Blackwell, 2020. 3. P.Tan, M.Steinbach, A.Karpatne, V.Kumar. Introduction to Data Mining, 2ed, Global edition, Pearson, 2020.	
Course additional literature:	1. Ch. Shah. A Hands-On Introduction to Data Science, Cambridge University Press, 2020. 2. J. Roger-Salazar. Data Science and Analytics with Python, CRC Press, 2017. 3. J. Roger-Salazar. Advanced Data Science and Analytics with Python, CRC Press, 2020. 4. L. Igual, S. Segui. Introduction to Data Science. A Python Approach to Concepts, Techniques and Applications, Springer, 2017.	
Course confirmation date:	08.12.2022	
Date of course description update:		

Study Course Plan for Full Time Students:

Date	Theme	Academic hours		Study Form/ Organization of independent work of students and task description
		Contact hours	Independent work hours	
	1. Introduction to course. Introduction to data science. Python environment and data science.	4	2	Theoretical lecture. Work environment preparation.
	2. Data. Big data. Data Science and Ethics. Data preprocessing. Introduction to MS Power BI.	4	4	Theoretical lecture. Practical work. Group work.
	3. Data Science related Python Data structures.	4	4	Theoretical lecture. Practical work. Group work.

VIDZEMES AUGSTSKOLA

	4. Data Visualization.	4	4	Theoretical lecture. Practical work. Group work.
	5. Dimensionality Reduction.	4	5	Theoretical lecture. Practical work. Group work.
	6. Classification & Predictive Performance.	4	6	Theoretical lecture. Practical work. Group work.
	7. Classification with Naïve Bayes. Classification - Alternative Techniques: Rule-Based.	4	5	Theoretical lecture. Practical work. Group work.
	8. Linear Regression. Handling Time Series.	4	5	Theoretical lecture. Practical work. Group work.
	9. Association Analytics I: Association Rules.	4	5	Theoretical lecture. Practical work. Group work.
	10. Association Analytics II: Advanced Concepts. Association Rules (Advanced). Recommenders (Collaborative Filtering).	4	6	Theoretical lecture. Practical work. Group work.
	11. Association Analytics III: Advanced Concepts. Graphs. Subgraph Mining. Social Network Analytics.	4	7	Theoretical lecture. Practical work. Group work.
	12. Cluster Analytics: Basic Concepts and Algorithms.	4	6	Theoretical lecture. Practical work. Group work.
	13. Natural Language Processing. LPW.	4	7	Theoretical lecture. Practical work. Group work.
	14. Data analytics with MS Power BI and Python - I.	4	5	Theoretical lecture. Group work.
	15. Data analytics with MS Power BI and Python - II. Anomaly Detection.	4	5	Theoretical lecture. Group work.
	Final examination.	4	20	Course project development and presentation.
	Hours total:	64	96	

Note: lecturer keeps the rights to make changes in the course plan.

VIDZEMES AUGSTSKOLA

Study Course Plan for Part Time Students:

Date	Theme	Academic hours		Study Form/ Organization of independent work of students and task description
		Contact hours	Independent work hours	
	1. Introduction to course. Introduction to data science. Python environment and data science.	2	5	Theoretical lecture. Work environment preparation.
	2. Data. Big data. Data Science and Ethics. Data preprocessing. Introduction to MS Power BI.	2	15	Theoretical lecture. Practical work. Group work.
	3. Data Science related Python Data structures. Data Visualization. Dimensionality Reduction.	2	15	Theoretical lecture. Practical work. Group work.
	4. Classification & Predictive Performance. Classification with Naïve Bayes. Classification - Alternative Techniques: Rule-Based.	2	15	Theoretical lecture. Practical work. Group work.
	5. Linear Regression. Handling Time Series.	2	15	Theoretical lecture. Practical work. Group work.
	6. Association Analytics I: Association Rules. Association Analytics II: Advanced Concepts. Association Rules (Advanced). Recommenders (Collaborative Filtering).	2	15	Theoretical lecture. Practical work. Group work.
	7. Association Analytics III: Advanced Concepts. Graphs. Subgraph Mining. Social Network Analytics.	2	15	Theoretical lecture. Practical work. Group work.
	8. Cluster Analytics: Basic Concepts and Algorithms. Natural Language Processing. LPW.	2	15	Theoretical lecture. Practical work. Group work.
	9. Data analytics with MS Power BI and Python. Anomaly Detection.	2	10	Theoretical lecture. Group work.
	Final examination.	2	20	Course project development and presentation.
	Hours total:	20	140	

Note: lecturer keeps the rights to make changes in the course plan.