SYLLABUS

1. Data about the program of study

1.1 Institution	The Technical University of Cluj-Napoca
1.2 Faculty	Faculty of Automation and Computer Science
1.3 Department	Computer Science
1.4 Field of study	Computer Science and Information Technology
1.5 Cycle of study	Master
1.6 Program of study/Qualification	Data Science / Master
1.7 Form of education	Full time
1.8 Subject code	8.1

2. Data about the subject

2.1 Subject name				***Data Analytics Pipeline***	
2.2 Course responsible/lecturer				Prof.Dr.Ing. Potolea Rodica – Rodica.Potolea@cs.utcluj.ro	
2.3 Teachers in charge of seminars/ laboratory/ project				Prof.Dr.Ing. Potolea Rodica – Rodica.Potolea@cs.utcluj.ro	
2.4 Year of study	1	2.5 Semester	2	2.6 Type of assessment (E - exam, C - colloquium, V - verification)	E
2.7 Subject category	Formative category: DA – advanced, DS – speciality, DC – complementary				DA
	Optionality: DI – imposed, DO – optional (alternative), DF – optional (free choice)				DO

3. Estimated total time

3.1 Number of hours per week	3	of which:	Course	1	Seminars	1	Laboratory	1	Project	0
3.2 Number of hours per semester	52	of which:	Course	14	Seminars	14	Laboratory	14	Project	0
3.3 Individual study:										
(a) Manual, lecture material and notes, bibliography										15
(b) Supplementary study in the library, online and in the field										15
(c) Preparation for seminars/laboratory works, homework, reports, portfolios, essays										15
(d) Tutoring										10
(e) Exams and tests										3
(f) Other activities:										

3.4 Total hours of individual study (suma (3.3(a)…3.3(f)))	58
3.5 Total hours per semester (3.2+3.4)	100
3.6 Number of credit points	5

4. Pre-requisites (where appropriate)

4.1 Curriculum	Introduction to Big Data, Machine Learning, Distributed System - bachelor
4.2 Competence	Operating with fundamental computer science concepts

5. Requirements (where appropriate)

5.1. For the course	Blackboard, Projector, PC MS Teams Platform
5.2. For the applications	PC, Specific Software

6. Specific competence

6.1 Professional competences	1 Working with advanced mathematical methods and models, engineering and computing specific techniques and technologies for data analysis. 2 Development of advanced techniques, methods and methodologies in the domains of software design, programming systems and environments and their applications . 3 Innovative design of artificial intelligence and related software and

	hardware using specific tools. 4 Contextual integration and exploitation of dedicated information systems. 5 Creative pooling of multidisciplinary knowledge in the field of computers and information technology for research, design, optimization, implementation and testing of theories, algorithms and original methods specific to artificial intelligence and computer vision systems. 6 Apply the fundamentals of data management and processing to a data science problem 7. Extract information from structured and unstructured data by considering their multivariate nature. 8 Identify machine learning and statistical modelling methods to use and apply them rigorously in order to solve a specific data science problem
6.2 Cross competences	1 Proof of knowledge for the economic, ethical, legal and social context associated with the profession, for correct task identification, schedule of activities, responsible decisions, with the final goal the design, preparation and presentation of a scientific paper. 2 Clear and concise description of professional activity flows, tasks and outcomes obtained by assuming the role of leader / project manager or as a member of a research team, as result of personal skills of domain specific information synthesis, global vision, communication skills with collaborators, ability of task stages identification. 3 Exercising the skill of continuous self-education and demonstrating critical, innovative and research abilities. 4 Capacity for managing the acquisition, the structuring, analysis and visualisation of data and information in the field of specialisation, and for critically assessing the results of this management.

7. Discipline objective (as results from the *key competences gained*)

7.1 General objective	Introducing students to data mining and data analysis processes; Knowledge of the techniques, algorithms and methods that can be used; Solutions identification and design for data analysis based on a given context
7.2 Specific objectives	Knowledge regarding data analysis methods and tools; Data processing; Determination and optimization of parameters; Methods for knowledge extraction

8. Contents

8.1 Lectures	Hours	Teaching methods	Notes
Introduction. What and why data analysis	1	Lectures using blackboard and projector; involving students in debate	
Advantages and challenges of data analytics	1		
Data analysis types and methods	1		
Data preparation. Dimensionality reduction	1		
Comparison, trend and ranking	1		
Variance, contribution and frequency	1		
Correlation and Pareto analysis	1		
Predictive data analysis	1		
Network analysis	1		
Data Visualisation	1		
Extracting knowledge through ML	1		
Time Series Analysis	1		
Framework for data analysis	1		
Realtime data	1		

Bibliography

- Jure Leskovec, Anand Rajaraman, Jeff Ullman. Mining of Massive Datasets. http://www.mmds.org/
- Hand, D.; Mannila, H.; Smyth, P. Principles of data mining. MIT Press, 2001. ISBN: 026208290X
- Spence, R, Pearson/Prentice Hall Information visualisation: design for interaction. 2007. ISBN: 9780132065504

- Domain articles

8.2 Applications – Seminars/Laboratory/Project	Hours	Teaching methods	Notes
Introduction to Data Analytics Tools	2	Oral presentation using slides, discutions (Q&A). Using multimedia tools, interactive teaching tools. Using specific software for data analysis	
Data preparation. Dimensionality reduction	2		
Comparison, trend and ranking	2		
Variance, contribution and frequency	2		
Correlation and Pareto analysis	2		
Predictive data analysis	2		
Network analysis	2		
Data Visualisation	2		
Extracting knowledge through ML	2		
Time Series Analysis	2		
Framework for data analysis	2		
Realtime data	2		
Business examples of applications	4		

Bibliography

- Python Data Science Handbook by Jake VanderPlas OReilly Media, Inc. 2nd edition, ISBN I978-1449369415
- Introduction to Machine Learning with Python: A Guide for Data Scientists by Andreas Müller, Sarah Guido, OReilly Media, Inc. 1st edition, ISBN 978-1449369415

** Se vor preciza, după caz: tematica seminariilor, lucrările de laborator, tematica și etapele proiectului.*

9. Bridging course contents with the expectations of the representatives of the community, professional associations and employers in the field

The class content was aligned with other similar classes from renowned universities and newest domain articles

10. Evaluation

Activity type	Assessment criteria	Assessment methods	Weight in the final grade
Course	Exam	Written exam	50%
Seminar	Presentation	Oral examination	25%
Laboratory	Exercises	Evaluation during the semester	25%
Project	-	-	-
Minimum standard of performance: Final grade > 5			

Date of filling in:	**Titulari**	**Titlu Prenume NUME**	**Semnătura**
	Curs	Prof.Dr.Ing. Potolea Rodica	
	Aplicații	Prof.Dr.Ing. Potolea Rodica	

Date of approval in the department
20.02.2024

Head of department
Prof.dr.ing. Rodica Potolea

Date of approval in the Faculty Council
22.02.2024

Dean
Prof.dr.ing. Mihaela Dinsoreanu