

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
1.7 Form of education	Full time

2. Data about the subject

2.1 Subject name	Mathematical Models for ML			Subject code	5.00
2.2 Course responsible / lecturer	Prof. Peter Ioan Radu -ioan.Radu.Peter@math.utcluj.ro				
2.3 Teachers in charge of seminars / Laboratory / project	Prof. Peter Ioan Radu				
2.4 Year of study	1	2.5 Semester	1	2.6 Type of assessment (E - exam, C - colloquium, V – verification)	E
2.7 Subject category	Formative category: DA – advanced, DS – speciality, DC – complementary				DS
	Optionality: DI – imposed, DO – optional (alternative), DF – optional (free choice)				DI

3. Estimated total time

3.1 Number of hours per week	2	of which:	Course	1	Seminars	1	Laboratory		Project	
3.2 Number of hours per semester	28	of which:	Course	14	Seminars	14	Laboratory		Project	
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										5
(e) Exams and tests										3
(f) Other activities:										5
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					4					

4. Pre-requisites (where appropriate)

4.1 Curriculum	- Undergraduate-level linear algebra (vector spaces, eigenvalues, orthogonality, projections). - Basic calculus (differentiation, multivariable optimization). - Familiarity with elementary statistics (mean, variance, correlation) is helpful but not required. - Some programming experience (e.g. Python/Matlab/R) for practical exercises.
4.2 Competence	- Ability to work with matrices and perform algebraic manipulations. - Understanding of optimization concepts (minimization/maximization problems). - Capacity to read and interpret mathematical notation and proofs.

	- Willingness to connect abstract mathematics to applied problems in machine learning.
--	--

5. Requirements (where appropriate)

5.1. For the course	Ability to solve basic linear algebra and optimization exercises.
5.2. For the applications	Familiarity with reproducible workflows (notebooks, scripts, or reports)

6. Specific competence

6.1 Professional competences	- Ability to apply linear algebra and optimization methods to formulate and solve machine learning problems. - Capacity to analyze, interpret, and evaluate algorithms such as regression, PCA, SVM, and EM from a mathematical perspective. - Competence in implementing data transformation, dimensionality reduction, and model building workflows for real datasets. - Skill in bridging abstract mathematical models with applications in data science (e.g. proteomics). - operations with mathematical methods and models, techniques and advanced specific engineering and IT technologies - develop creative ideas - use data processing techniques - use data analysis techniques
6.2 Cross competences	- Development of critical thinking in connecting theory to applications. - Problem-solving ability in interdisciplinary contexts. - Competence in teamwork, project-based learning, and communication of technical results. - Capacity for autonomous learning and continuous professional development in mathematical modeling and data science. The graduate: • develop an analytical approach • taking a proactive approach • developing strategies to solve problems • being open minded

7. Expected Learning Outcomes

Knowledge	- Define and explain the fundamental mathematical structures underlying machine learning, including linear algebra (QR, SVD, PCA), optimization (convexity, duality), and kernel methods. - Describe the role of deterministic modeling approaches in data science. - Identify and compare methods for data preprocessing, normalization, and missing data imputation. - Summarize the principles of GLM, PCA, SVM, and EM algorithms and interpret their applications in domains such as proteomics.
Skills	- Apply linear algebra and optimization techniques to formulate and solve machine learning problems. - Implement algorithms (regression, PCA, SVM, EM) in computational environments. - Analyze and evaluate the mathematical properties, assumptions, and limitations of core ML methods. - Design and develop small-scale data analysis pipelines, including preprocessing, dimensionality reduction, and model fitting. - Communicate results effectively, present findings clearly, and collaborate in interdisciplinary contexts.
Responsibilities and autonomy	- Work independently in solving mathematical modeling tasks and collaborate effectively in team projects. - Demonstrate responsibility in handling data ethically and reproducibly. - Engage in continuous learning, seeking new methods and tools beyond the course. - Transfer knowledge across disciplines (mathematics, computer science, applied sciences such as bioinformatics).

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

8.1 General objective	- To provide students with a solid understanding of the mathematical foundations of machine learning, emphasizing linear algebra, optimization, and deterministic modeling methods. - To enable learners to bridge abstract mathematical models with practical applications in data science and interdisciplinary fields such as proteomics.
8.2 Specific objectives	By the end of the course, students will be able to: - Explain core mathematical tools (matrix decompositions, regression, PCA, optimization, kernel methods). - Apply these tools to real datasets, implementing algorithms in computational environments. - Analyze and evaluate the mathematical properties and limitations of machine learning models. - Design small-scale data pipelines including preprocessing, transformation, and model building. - Integrate knowledge from mathematics with applications in data science and communicate results effectively.

	- Develop autonomy in extending mathematical knowledge and selecting appropriate methods for new problems.
--	---

9. Contents

9.1 Lectures	Hours	Teaching methods	Notes
Mathematical Problems in Data Science. Overview.	1		
Matrix Decompositions: QR, SVD. Their meaning and applications.	1		
Linear Regression. Qualitative and quantitative aspects.	1		
Generalized Linear Models (GLM).	1		
Optimization Problems I	1		
OptimizationProblems II (duality, Fenchel Transforms)	1		
Support Vector Machine (SVM)- mathematical aspects.	1		
Glimpse in Kernel Methods (SVM example).	1		
Dimensionality reduction - Principal Component Analysis	1		
Cleaning data. Data transformation, normalization. An overview.	1		
Missing data imputation- methods, approaches	1		
Applications of GLM and PCA - proteomic data analysis pipeline.	1		
Expectation maximization.	1		
What's next? Discussions on more advanced topics.	1		
Bibliography: [1] Christopher M Bishop. Pattern Recognition and Machine Learning. Springer, 2006. [2] Stephen Boyd and Lieven Vandenberghe. Convex Optimization. Cambridge University Press, 2004. [3] Nello Cristianini and John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000. [4] Gene H Golub and Charles F Van Loan. Matrix Computations. Johns Hopkins University Press, 2013. [5] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009. [6] Ian Jolliffe. Principal Component Analysis. Springer, 2002. [7] Yurii Nesterov. Introductory Lectures on Convex Optimization: A Basic Course. Springer, 2004. [8] Bernhard Schölkopf and Alexander J Smola. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, 2002. [9] Gilbert Strang. Linear Algebra and Learning from Data. Wellesley-Cambridge Press, 2019. [10] Lloyd N Trefethen and David Bau. Numerical Linear Algebra. SIAM, 1997.			
9.2 Applications - Seminars/Laboratory/Project	Hours	Teaching methods	Notes
Mathematical Problems in Data Science. Overview.	1		
Matrix Decompositions: QR, SVD. Their meaning and applications.	1		
Linear Regression. Qualitative and quantitative aspects.	1		
Generalized Linear Models (GLM).	1		
Optimization Problems I	1		
OptimizationProblems II (duality, Fenchel Transforms)	1		
Support Vector Machine (SVM)- mathematical aspects.	1		
Glimpse in Kernel Methods (SVM example).	1		
Dimensionality reduction - Principal Component Analysis	1		
Cleaning data. Data transformation, normalization. An overview.	1		
Missing data imputation- methods, approaches	1		
Applications of GLM and PCA - proteomic data analysis pipeline.	1		

Expectation maximization.	1		
What's next? Discussions on more advanced topics.	1		
Bibliography [1] Christopher M Bishop. Pattern Recognition and Machine Learning. Springer, 2006. [2] Stephen Boyd and Lieven Vandenberghe. Convex Optimization. Cambridge University Press, 2004. [3] Nello Cristianini and John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000. [4] Gene H Golub and Charles F Van Loan. Matrix Computations. Johns Hopkins University Press, 2013. [5] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009. [6] Ian Jolliffe. Principal Component Analysis. Springer, 2002. [7] Yurii Nesterov. Introductory Lectures on Convex Optimization: A Basic Course. Springer, 2004. [8] Bernhard Schölkopf and Alexander J Smola. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, 2002. [9] Gilbert Strang. Linear Algebra and Learning from Data. Wellesley-Cambridge Press, 2019.			

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

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

The course responds to the expectations of employers and professional associations by emphasizing:

- **Mathematical literacy** (linear algebra, optimization, probability) needed to justify and adapt methods.
- **Data handling skills** (normalization, transformation, imputation) required in real scientific and industrial datasets.
- **Conceptual clarity** – understanding *why* algorithms work, ensuring interpretability and reproducibility.
- **Adaptability** – preparing graduates to apply methods in diverse domains (IT, finance, healthcare, bioinformatics).

This ensures that students are not only tool users but competent data scientists, aligned with professional standards (ACM, IEEE, ASA) and industry needs.

11. Evaluation

Activity type	Assessment criteria	Assessment methods	Weight in the final grade
Course	Knowledge of theoretical principles and results. Problem solving skills	Exam	50%
Seminar	Problem solving skills. Activity	Exam	50%
Laboratory			
Project			
Minimum standard of performance:			

Date of filling in:	Responsible	Title First name Last name	Signature
---------------------	-------------	----------------------------	-----------

zz.II.aaaa	Course	Prof. Mat. Ioan Radu Peter	<i>Peter</i>
	Applications	Prof. Mat. Ioan Radu Peter	

Date of approval in the department	Head of department, Prof.dr.eng. Rodica Potolea
Date of approval in the Faculty Council	Dean, Prof.dr.eng. Vlad Mureșan