COURSE OUTLINE

(1) GENERAL

SCHOOLS	**ENGINEERING, NATURAL SCIENCES**		
ACADEMIC UNIT/UNITS	**COMPUTER ENGINEERING AND INFORMATICS DEPARTMENT, DEPARTMENT OF MATHEMATICS**		
TITLE OF MASTER'S DEGREE	*MSC in Data Driven Computing and Decision Making*		
LEVEL OF STUDIES	Post graduate		
COURSE CODE	DDCD105	SEMESTER	First
COURSE TITLE	Matrix methods and tools in data driven science		

INDEPENDENT TEACHING ACTIVITIES *if credits are awarded for separate components of the course, e.g. lectures, laboratory exercises, etc. If the credits are awarded for the whole of the course, give the weekly teaching hours and the total credits*	WEEKLY TEACHING HOURS	CREDITS
	3	7.5
Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).	Total	7.5

COURSE TYPE *general background, special background, specialised general knowledge, skills development*	Special background, specialized general knowledge, skills development
PREREQUISITE COURSES:	None at the graduate level Linear algebra, numerical analysis (especially numerical linear algebra) and algorithms at the undergraduate level.
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	Greek but English is also possible
IS THE COURSE OFFERED TO ERASMUS STUDENTS	Yes
COURSE WEBSITE (URL)	https://eclass.upatras.gr/courses/CEID1164

(2) LEARNING OUTCOMES

Learning outcomes

The course learning outcomes, specific knowledge, skills and competences of an appropriate level, which the students will acquire with the successful completion of the course are described.

Consult Appendix A

- *Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area*
- *Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B*
- *Guidelines for writing Learning Outcomes*

ΓThe general objective of the course is the presentation and study of advanced methods and computational tools of linear algebra and matrix computations emphasizing the solution of problems in Data Science. The mathematical methods presented have as target the manipulation of the two fundamental objects of the field, namely graphs and matrices. At the end of the course, one would be familiarize with several state-of-the-art techniques and their application, they would know the strengths and weaknesses of these methods, and they would be able to select methods based on the problem characteristics. They would also be able to apply and combine these techniques and would be able to follow the rapidly evolving research literature on the topic.

General Competences

Taking into consideration the general competences that the degree-holder must acquire (as these appear in the Diploma Supplement and appear below), at which of the following does the course aim?

Search for, analysis and synthesis of data and information, with the use of the necessary technology	*Project planning and management* *Respect for difference and multiculturalism*

Adapting to new situations *Decision-making* *Working independently* *Team work* *Working in an international environment* *Working in an interdisciplinary environment* *Production of new research ideas*	*Respect for the natural environment* *Showing social, professional and ethical responsibility and sensitivity to gender issues* *Criticism and self-criticism* *Production of free, creative and inductive thinking* *……* *Others…* *…….*

(3) SYLLABUS

The course consists of 4 modules.

I. **Introduction:** Matrix computations as kernel in Data Analytics. Graphs, networks and matrices. Berkeley Dwarfs. Communication aware computational models. Fundamental problems in matrix computations. The BLAS and examples. Structure in matrix computations. Sparse matrix technology.

II. **Numerical Fundamentals:** Roundoff errors, numerical stability, well and ill posed problems, methods for error analysis. Elements of numerical optimization. Elements of iterative methods: Descent methods. Projection-based methods (Krylov subspace techniques) for linear systems. Gradient Descent techniques. Nonlinear problems and Newton's method. Projection methods for eigenvalue problems and the SVD. Computing matrix functions.

III. **HPC and the Software Stack:** The impact of high performance computing. The impact of communication and the concept of ``communication avoiding'' algorithms. Recursive design and cache-conscious vs. cache-oblivious algorithms. Software implementations and libraries. Scripting languages: From MATLAB to Julia.

IV. **Matrix methods in Data Driven applications:** Dimensionality reduction, clustering, retrieval. Matrix factorizations and low rank matrix approximation. Regularization. The LSI model. Clustering and representatives in matrix approximation methods. Elements of nonnegative matrix theory: Introduction to nonnegative matrices and Perron-Frobenius theory. Special matrices. Constrained optimization problems. Nonnegative rank factorization and approximation. Randomized Numerical Linear Algebra: Approximation methods for very large matrices and the CUR decomposition. The MATLAB TMG and Text Analytics Toolboxes. Tensors and elements of tensor methods. Methods in Web IR: Centrality indices. (PageRank and variants). Using matrix functions to compute matrix characteristics of importance to applications such as centrality indices and leverage scores.

TEACHING and LEARNING METHODS - EVALUATION

DELIVERY *Face-to-face, Distance learning, etc.*	Face-to-face
USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY *Use of ICT in teaching, laboratory education, communication with students*	All lectures and additional material are posted on the course's e-Class website.
TEACHING METHODS *The manner and methods of teaching are described in detail.* *Lectures, seminars, laboratory practice, fieldwork, study and analysis of bibliography,*	<table><tr><th>*Activity*</th><th>*Semester workload*</th></tr><tr><td>Lectures</td><td>36</td></tr><tr><td>Laboratory Project</td><td>30</td></tr><tr><td>Survey of the</td><td>86</td></tr></table>

tutorials, placements, clinical practice, art workshop, interactive teaching, educational visits, project, essay writing, artistic creativity, etc. *The student's study hours for each learning activity are given as well as the hours of non-directed study according to the principles of the ECTS*	relevant literature	
	Report writing	34
	Oral presentation	2
		188

STUDENT PERFORMANCE EVALUATION *Description of the evaluation procedure* *Language of evaluation, methods of evaluation, summative or conclusive, multiple choice questionnaires, short-answer questions, open-ended questions, problem solving, written work, essay/report, oral examination, public presentation, laboratory work, clinical examination of patient, art interpretation, other* *Specifically-defined evaluation criteria are given, and if and where they are accessible to students.*	Evaluatin is based on the following factors: 1) Class participation (attendance is mandatory). 2) Study and evaluation of papers selected from the recent literature in the form of a report. Each report must also include code and numerical experiments. 3) Oral presentation of the report. 4) Oral examination based on the report. The oral examination also tests the student’s familiarity with the topics discussed in class.

(4) ATTACHED BIBLIOGRAPHY

- G. Strang, Linear Algebra and Learning from Data, Wellesley-Cambridge Press, 2019.
- Gene Golub and Charles Van Loan, Matrix Computations, 4th edition, Johns Hopkins University Press, 2013.
- M. Mahoney, J. Duchi and A.C. Gilbert, eds., The Mathematics of Data, IAS/Park City Mathematics Series, vol. 25, 2018.
- Lars Eldén, Matrix Methods in Data Mining and Pattern Recognition, SIAM, 2007.
- D. Simovici, Linear Algebra Tools for Data Mining, World Scientific, 2012.
- J. Nocedal and S.J. Wright, Numerical Optimization, Springer, 2006.
- I. Goodfellow and Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.
- E. Gallopoulos, B. Philippe and A. Sameh, Parallelism in Matrix Computations, Springer, 2016.
- <u>Journals</u>: SIAM J. Matrix Analysis, SIAM J. Math. Data Sciences, SIAM J. Scientific Computing, IEEEE TKDE. Συνέδρια (Πρακτικά): IEEE Int’l Conf. Data Mining (ICDM), SIAM Int’l. Conf. Data Mining (SDM), SIAM Conf. Applied Linear Algebra,