UNIVERSITY of NICOSIA

Course Syllabus

Course Code	Course Title	ECTS Credits
COMP-540DL	Mathematics for Data Science	10
Prerequisites	**Department**	**Semester**
None	Computer Science	Fall
Type of Course	**Field**	**Language of Instruction**
Required	Mathematics	English
Level of Course	**Lecturers**	**Year of Study**
2nd Cycle	George Portides and Nectarios Papanicolaou	1st
Mode of Delivery	**Work Placement**	**Corequisites**
Distance Learning	N/A	None

Course Objectives:

The main objectives of the course are to:

- Provide students with the necessary Mathematics, Probability and Statistics background for understanding concepts in Data Science and Machine Learning.
- Cover random variables and their probability distributions.
- Familiarize students with expectation and enable them to compute expectations of functions of random variables.
- Discuss Descriptive Statistics and provide students with the necessary knowledge to produce summary statistics and present data.
- Introduce students to the concepts and techniques necessary to interpret linear regression models.
- Cover the fundamentals of vector spaces and matrices.
- Discuss orthogonality and inner product spaces.
- Introduce eigenvalues, eigenvectors and the spectral theorem and discuss its connection with Data Science.
- Present selected matrix factorizations.

Learning Outcomes:

After completion of the course students are expected to be able to:

- Analyse the behaviour of random variables following the most common probability distributions.
- Compute expectations of functions of random variables.
- Summarize and present univariate data.
- Perform hypothesis tests and find confidence intervals for the mean.
- Interpret and calculate correlations between two variables and the coefficients of linear regression models.
- Compute vector inner products and norms.
- Apply the Gram-Schmidt process to construct orthonormal bases.
- Implement Gaussian elimination and related techniques to solve linear systems.
- Calculate the eigenvalues and associated eigenvectors of a matrix.
- Compute various matrix factorizations.
- Utilize Linear Algebra techniques to represent and analyse data.

Course Content:

1. Probability, Random Variables, Probability Distributions and Expectation
    - Discrete and continuous random variables and their probability distributions.
    - Expectation.
    - The Binomial, Poisson, Exponential and Normal distributions.
2. Exploratory Data Analysis
    - Discrete and continuous data, random samples, population.
    - Histograms, bar charts, pie charts, box plots.
    - Descriptive Statistics.
3. Statistical Inference about the Mean
    - Confidence Intervals.
    - Hypothesis testing for large and small sample sizes.
4. Linear Regression and Correlation
    - Scatterplots.
    - The regression model.
    - Least squares estimation.
5. Matrices
    - Special matrices (triangular, orthogonal, symmetric, diagonal).
    - Linear systems and Gaussian Elimination.
    - Determinants.
6. Vector Spaces
    - Linear Independence, Basis and Dimension.

- Subspaces.
- Vector Norms.
- Inner Products and Orthogonality – The Gram-Schmidt procedure.
- Application: The gradient of a multivariable function.

7. Eigenvalues and Eigenvectors
   - Eigendecomposition (diagonalization) of a matrix.
   - Algebraic and geometric multiplicity. Eigenspaces.
   - Matrix Powers and the Cayley-Hamilton theorem.
8. Matrix Factorizations
   - The Spectral decomposition.
   - The SVD decomposition and applications in data compression.
   - The QR factorization.
   - The pseudoinverse.

Learning Activities and Teaching Methods:

Online Presentations, Textbook Exercises, Online Exercises, Online Interactions (Forums and Chats).

Assessment Methods:

Final Exam
Assignments (2)

Required Textbooks / Readings:

Title	Author(s)	Publisher	Year	ISBN
Probability and Statistics	DeGroot M.H. and Schervish M.J.	Addison-Wesley	2012	978-0-321-50046-5
Introduction to Statistics and Data Analysis	Peck R., Olsen C. and Devore J. L.	Cengage Learning	2016	978-1-305-11534-7
Elementary Linear Algebra: Applications Version, 11th Ed.	Howard Anton and Chris Rorres	Wiley	2014	ES8-1-118-47350-4 (e-textbook)

Recommended Textbooks / Readings:

Title	Author(s)	Publisher	Year	ISBN
Introduction to Applied Linear Algebra: Vectors, Matrices and Least Squares, 1st Ed.	Stephen Boyd and Lieven Vandenberghe	Cambridge University Press	2018	978-1-58450-679-9