

Annual structure of DAMA
until the academic year 2025-2026

❖ **Course Structure**

1st year	
DAMA50 Mathematics for Machine Learning	(C ¹ , 30 ECTS)
DAMA51 Foundations in Computer Science	(C, 30 ECTS)
2nd year	
DAMA60 Algorithmic Techniques and Systems for Data Science and Machine Learning	(C, 30 ECTS)
DAMA61 Numerical and Computational Techniques for Data Science and Machine Learning	(C, 30 ECTS)
Note:	

C¹: Compulsory

❖ **Module Selection Guidelines**

The Programme is structured in two (2) academic years, which include four (4) compulsory modules. Students may choose one (1) to two (2) modules each academic year.

In the first year of studies if students wish to select one module; they can choose either the module DAMA50 or the module DAMA51. If they wish to select two modules they can choose both DAMA50 and DAMA51.

In the second year of studies if students wish to select one module; they can choose either the module DAMA60 or the module DAMA61. If they wish to select two modules they can choose both DAMA60 and DAMA61.

Each academic year students can choose from one (1) up to two (2) modules with full responsibilities.

During the module selection, all modules from previous academic years must be selected and consequently the selection of the modules of the next academic year can begin.

The minimum study duration of the Master's Degree Programme is two academic years.

❖ **Course Modules**

DAMA50 Mathematics for Machine Learning

Module code: DAMA50

ECTS Credit Points: 30

Module Type: Compulsory

Year: 1st

Language: English

Module general description: The students will learn the basic mathematical tools necessary for Machine Learning (ML). These include basic concepts from linear algebra such as vectors, matrices, measures and operations with vectors and matrices. From calculus students will be exposed to functions of many real variables and the basic concept of the gradient and directional derivative to be applied in backpropagation ML algorithms. Very basic tools of probability, statistics and optimisation will be also introduced. Overall, a student without prior knowledge of these mathematical areas will be able to form a background in order to

understand the ML techniques while and student with prior mathematical knowledge will be able to go much deeper in application of mathematics in ML. The mathematical study will be supplemented by computational software that will enable both analytical and numerical evaluations.

Learning Outcomes: After completing this module, students are expected to be able to:

- Recognize that the basic mathematical pillars for Machine Learning are Linear Algebra, Vector Calculus as well as Probability and Statistics and apply analytical and computational tools
- Formulate linear equations with many unknowns, detail matrix techniques for their solution.
- Summarize basic notions of vector spaces, use linear mappings for basis change.
- Use SageMath for solving linear algebra problems.
- Outline the concept of norm of a vector, inner products between two vectors and use it in obtaining the length of a vector.
- Explain what an orthonormal basis is in a vector space and describe the orthogonal complement of a subspace of the vector space.
- Outline the Gram-Schmidt orthogonalization procedure and derive an orthonormal basis for a vector space
- Use SageMath to perform basic vector manipulations and perform Gram-Schmidt orthogonalization.
- Recall the definition of the trace and the determinant of a matrix and be able to calculate both by hand for simple matrices;, explain the concepts of eigenvalues and eigenvectors of square matrices.
- Describe how a matrix can be decomposed through a Cholesky decomposition and Singular Value Decomposition (SVD), and apply these methods to simple matrices.
- Use SageMath to implement matrix decompositions.
- Outline the concept of the gradient of a function of many variables and describe its geometric significance.
- Summarize the gradient of matrices and its geometric significance and calculate it explicitly in specific cases.
- Summarize the concept of backpropagation and use it in simple models of neural networks.
- Use SageMath for evaluating gradients and derivatives.
- Summarize the properties of single variate and multivariate Gaussian distribution, find marginals and conditionals as well as transformations of the Gaussian function.
- Focus on the binomial Bernoulli distribution and detail the Beta distribution.
- Summarize the conjugate priors connected through Bayes theorem.
- Explain what is sufficient statistics and outline the exponential family of distributions.
- Perform a change of random variables and find the new distribution function.
- Recall how to find minima of a single variable function.
- Summarize the procedure to find the minimum of a multivariate function using the gradient descent algorithm.
- Explain how to perform stochastic gradient descent and what are its advantages and limitations compared to the gradient descent method.

- Describe what are the Lagrange multipliers and explain how they are used in constrained optimization.
- Describe convex optimization.
- Use SageMath to find the minimum of a multivariate function and t to minimize a function with constraints.

Subjects covered:

- Linear Algebra
- Calculus
- Statistics and Probabilities

Prerequisites: There are no prerequisites for this module.

Evaluation: Students are assigned to submit six (6) written assignments during the academic year. The average grade of the six (6) written assignments, weighted at 30%, is taken into consideration for the calculation of the final grade. The grade of written assignments is activated only with a score equal to or above the pass level (≥ 5) in the final or resit exams. The grade of the final or the resit exams shall be weighted at 70 % for the calculation of the final grade.

Students have the right to participate in the final/resit exams if (a) at least 50% of the potentially excellent grade has been obtained when adding the total of the six (6) assignments and (b) at least four (4) of the six (6) written assignments have been submitted.

Teaching Method: Distance education with five Contact Sessions held at weekends during the academic year.

DAMA51 Foundations in Computer Science

Module code: DAMA51

ECTS Credit Points: 30

Module Type: Compulsory

Year: 1st

Language: English

Module general description: The students will acquire a strong background as far as the algorithmic aspects and the computational requirements of data science and machine learning approaches are concerned. They will also develop an in depth understanding of the key technologies in data science and data analytics. After they will be presented with the fundamental concepts and principles that underlie the techniques for extracting knowledge from data, they will become acquainted with a number of practical considerations regarding the analysis and the interpretation of the data as well as assessing the quality of the input data and deriving insights from the results of mining the data. By the time the students will complete this module, they will be able to apply theory, languages algorithms and tools to solve real world problems while they will be proficient in interpreting and communicating findings to any kind of audience.

Learning Outcomes:

- Identify what data science is, and how it is related to statistics.
- List basic statistical analysis techniques.
- Indicate the roles and the skills of a data scientist.
- Name and recognize the steps of the data science process.
- Appraise the quality of input data.

- Accept that data science must be applied as an iterative process.
- Code simple algorithms in a high level programming language.
- Apply scraping and data munging.
- Use different data science software packages.
- Use appropriate methods of analysis to build highly understandable and accurate models.
- Assess the fit of a model to the data.
- Investigate potential issues in data and models.
- Interpret and communicate findings to an audience.
- Derive insights from the results of data analysis.
- Describe the terminology of the logic domain.
- Describe the terminology of the search domain.
- Describe the terminology of the decision trees domain.
- Understand the differences between propositional logic and predicate logic.
- Understand the differences between blind and informed search algorithms.
- Use natural language to explain sentences in propositional or predicate logic.
- Apply a search algorithm to a given search space, utilising constraints where applicable.
- Apply decision tree learning to various data sets using the WEKA software.
- Formulate search problems by defining state spaces, start and goal states, and state transition operators.
- Formulate game search problems.
- Analyze a data set identifying features/attributes and classes, towards applying a machine learning algorithm.
- Compare search algorithms in a set of problems.
- Evaluate a decision tree algorithm.
- Describe a world using propositional or predicate logic and apply inferences.
- Construct a heuristic function for a search problem.

Subjects covered:

- Data Structures, Algorithms and Databases
- Data Science Fundamentals
- Artificial Intelligence

Prerequisites: There are no prerequisites for this module.

Evaluation: Students are assigned to submit five (5) written assignments during the academic year. The average grade of the five (5) written assignments, weighted at 30%, is taken into consideration for the calculation of the final grade. The grade of written assignments is activated only with a score equal to or above the pass level (≥ 5) in the final or resit exams. The grade of the final or the resit exams shall be weighted at 70 % for the calculation of the final grade.

Students have the right to participate in the final/resit exams if (a) at least 50% of the potentially excellent grade has been obtained when adding the total of the five (5) assignments and (b) at least three (3) of the five (5) written assignments have been submitted.

Teaching Method: Distance education with five Contact Sessions held at weekends during the academic year.

DAMA60 Algorithmic Techniques and Systems for Data Science and Machine Learning

Module code: DAMA60

ECTS Credit Points: 30

Module Type: Compulsory

Year: 2nd

Language: English

Module general description: The students will acquire a strong background on the data structures, the algorithmic aspects and the computational requirements of data mining and machine learning approaches for analyzing very large volumes of data. Among other topics, the module will emphasize on tools for the parallelization of different machine learning algorithms such as Hadoop and Map Reduce, Recommender Systems, Dimensionality Reduction, Finding Nearest Neighbors and Similar Sets, Clustering, Link Analysis, Association Rules and Frequent Itemsets. The students are also expected to build on the basic programming skills that they acquired in DAMA50 and DAMA51 and enhance their understanding of how to apply these skills on a project where they will be asked to work on real data sets and computational infrastructure through R and/or Python and Azure and/or KNIME.

Learning Outcomes: After completing this module, students are expected to be able to:

- Define the importance of Big Data in Machine Learning Applications
- Identify key characteristics of Big Data
- Compare serial and parallel processing techniques for mining data
- Define scalability and fault tolerance for machine learning algorithms
- Apply Big Data tools for solving real life problems
- Identify the Big Data ecosystem
- Describe the benefits of Cloud Computing for Big Data Applications
- Recognize the importance of Distributed File Systems and Map-Reduce
- Create parallel algorithms for mining large volumes of data
- Apply similarity search by using MinHashing and Locality Sensitive Hashing
- Describe Data Stream Processing
- Apply specialized algorithms for dealing with stream data
- Recognize the technology underlying the principles of search engine operation
- Use frequent itemset mining through Apriori and its improvements
- Describe algorithms for clustering Big Data
- Identify key problems for mining data from Web Applications
- Describe algorithms for analyzing social network graphs
- Apply techniques for obtaining the important properties of large datasets
- Use Machine Learning algorithms for mining large datasets

Subjects covered:

- Machine Learning and Data Mining
- Big Data Analytics
- Distributed Learning

Prerequisites: There are no prerequisites for this module.

Evaluation: Students are assigned to submit five (5) written assignments during the academic year. The average grade of the five (5) written assignments, weighted at 30%, is taken into

consideration for the calculation of the final grade. The grade of written assignments is activated only with a score equal to or above the pass level (≥ 5) in the final or resit exams. The grade of the final or the resit exams shall be weighted at 70 % for the calculation of the final grade.

Students have the right to participate in the final/resit exams if (a) at least 50% of the potentially excellent grade has been obtained when adding the total of the five (5) assignments and (b) at least three (3) of the five (5) written assignments have been submitted.

Teaching Method: Distance education with five Contact Sessions held at weekends during the academic year.

DAMA61 Numerical and Computational Techniques for Data Science and Machine Learning

Module code: DAMA61

ECTS Credit Points: 30

Module Type: Compulsory

Year: 2nd

Language: English

Module general description: The students will be able to implement basic machine learning methods in Jupyter notebooks, use TensorFlow and Keras, write and execute python code, utilize linear and nonlinear regression, support vector machines, perform model regularization, implement decision trees and ensemble learning in the form of random forests. The students are expected to know how to perform dimensionality reduction and use principal component analysis. The module will also focus on neural network methods and deep learning including fully connected deep networks, convolutional neural networks and autoencoders. Use of recurrent neural networks, physics informed neural networks and restricted Boltzmann machines completes the material of the module. DAMA-61 builds heavily on DAMA-50 and after its completion the students will be able to use the mathematical tools acquired in the latter in real world data problems.

Learning Outcomes: After completing this module, students are expected to be able to:

- Recognize and execute Jupyter notebooks with machine learning procedures
- Implement TensorFlow and Keras
- Define Linear and/or nonlinear regression variables in supervised learning mode
- Implement support vector machines for data classification
- Identify decision boundaries
- Create decision trees and implement random forests
- Execute Lasso and alternative regularizations
- Perform dimensionality as well as principal component analysis
- Use TensorFlow/Keras to introduce fully connected neural networks
- Perform Deep Learning training and hyperparameter testing
- Execute code for recurrent neural networks
- Apply convolutional neural networks to specific data sets
- Perform unsupervised learning by implementing autoencoders
- Apply reinforcement learning and physics informed machine learning
- Describe restricted Boltzmann machines

Subjects covered:

- Supervised/Unsupervised Learning

- Neural Networks and Deep Learning
- Optimization

Prerequisites: There are no prerequisites for this module.

Evaluation: Students are assigned to submit six (6) written assignments during the academic year. The average grade of the six (6) written assignments, weighted at 30%, is taken into consideration for the calculation of the final grade. The grade of written assignments is activated only with a score equal to or above the pass level (≥ 5) in the final or resit exams. The grade of the final or the resit exams shall be weighted at 70 % for the calculation of the final grade.

Students have the right to participate in the final/resit exams if (a) at least 50% of the potentially excellent grade has been obtained when adding the total of the six (6) assignments and (b) at least four (4) of the six (6) written assignments have been submitted.

Teaching Method: Distance education with five Contact Sessions held at weekends during the academic year.

- - - - -