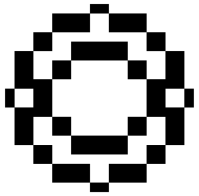
Machine Learning

Master in Data Science and Advanced Analytics
BA and DS specializations

Roberto Henriques

Ricardo Santos

NOVA IMS INFORMATION MANAGEMENT SCHOOL

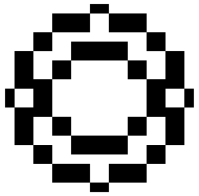


OBJECTIVES

In this course, students will gain a strong foundation in the core concepts, applications, and tools for machine learning.

The general Learning Objectives are:

- Understand the significant machine learning approaches, including supervised learning and unsupervised learning.
- Use Python to implement popular machine learning algorithms like linear regression, logistic regression, neural networks, support vector machines, decision trees, and k-means clustering.
- Evaluate and compare machine learning models using proper evaluation metrics and techniques like train/test splits, cross-validation, confusion matrices, and classification reports.
- Gain experience with the whole machine learning workflow, including data exploration, data cleaning and preprocessing, feature engineering, model optimisation, and deployment.
- Apply machine learning to solve real-world problems through hands-on projects and assignments using datasets from domains like computer vision, natural language processing, and recommender systems.
- Develop proper techniques to avoid overfitting, handle missing data, and perform feature selection and dimensionality reduction.
- By completing this course, students will gain valued machine-learning skills to drive innovations and technologies powered by artificial intelligence.



LEARNING OUTCOMES

By the end of the course, the student should:

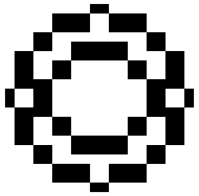
LO1. Understand the fundamental concepts, algorithms, and machine learning techniques, including supervised and unsupervised learning.

LO2. Gain proficiency in using Python and machine learning libraries like Scikit-Learn, Pandas, and NumPy to implement machine learning workflows, including data preprocessing, model training, evaluation, and prediction.

LO3. Apply popular machine learning algorithms such as linear regression, logistic regression, decision trees, support vector machines, and neural networks to solve real-world classification and regression problems.

LO4. Develop skills in data preparation techniques for machine learning, including data cleaning, feature engineering, feature selection, scaling, and dimensionality reduction. Be able to transform raw datasets into formats suitable for ML algorithms.

LO5. Rigorously evaluate machine learning models using appropriate evaluation metrics, visualize model performance, and optimize models through grid search, cross-validation, and hyperparameter tuning techniques. Avoid issues like overfitting and underfitting.



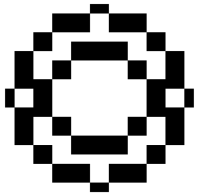
Syllabus

- LU1. Introduction to machine learning fundamentals.
  - a. Overview of machine learning and key terminology
  - b. Categories of machine learning: supervised, unsupervised, reinforcement
  - c. Applications and examples of machine learning
- LU2. Data Preprocessing and feature selection
  - a. Data preprocessing techniques: cleaning, transformations, feature engineering
  - b. Types of feature selection
- LU3. Model Selection and Evaluation
  - a. Train/test splits, cross-validation
  - b. Classification evaluation metrics and techniques
  - c. Regression evaluation metrics and techniques
  - d. Overfitting and underfitting
- LU4. Introduction Supervised Learning Algorithms
  - a. Linear and Logistic regression
  - b. Probability-based learning
  - c. Similarity-based learning
  - d. Regression and classification trees
  - e. Ensemble classifiers
  - f. Neural networks
  - g. Support Vector Machines
- LU5. Machine Learning Projects
  - a. End-to-end machine learning project walkthrough
  - b. Deploying machine learning models to applications

#	Type	Week		Details
1	T	8-Sep-25	Introduction to machine learning. Motivation & Applications of Machine Learning.	RH
	TP		No practical class	
2	T	15-Sep-25	Machine Learning Tribes	RH
	TP		No practical class	
3	T	22-Sep-25	Feature selection	RH
	TP		The concept of Supervised and Unsupervised Learning applied to the real world	RS
4	T	29-Sep-25	Model selection and evaluation	RH
	TP		Data transformation, feature engineering and feature selection in Python	LD
5	T	6-Oct-25	Linear and Logistic Regression	RH
	TP		How to evaluate predictive models using different metrics in Python	LD
6	T	13-Oct-25	Probability based Learning	RH
	TP		Linear and Logistic Regression in Python	LD
7	T	20-Oct-25	Similarity based learning	RH
	TP		Probability based Learning algorithms in Python	RS
		27-Oct-25	**No class week for trimester courses exams**	
			Handout: November 3th at 18:00h	
8	T	3-Nov-25	Neural Networks I	RH
	TP		Similarity based learning algorithms in Python	RS
9	T	10-Nov-25	Neural Networks II	RH
	TP		Neural Networks in Python.	RS
10	T	17-Nov-25	Regression and classification trees I	RH
	TP		Neural Networks in Python. The concept of GridSearch.	RS
11	T	24-Nov-25	Regression and classification trees II	RH
	TP		Regression Trees in Python	RS
12	T	1-Dec-25	Ensemble classifiers I	RH
	TP		Classification Trees in Python	RS
13	T	8-Dec-25	Ensemble classifiers II	RH
	TP		Ensemble classifiers in Python I	RS
14	T	15-Dec-25	Support Vector Machines	RH
	TP		Ensemble classifiers in Python II	RS
15	TP	Extra	SVM in Python	RS
16	TP	to be scheduled	Project Support	
			Project Deadline: December 22nd at 18:00h	

Teaching methodologies

- The course is based on theoretical and practical lessons.
- As presented in the curricular unit planning table, various effective teaching methodologies are applied:
  - **Lectures** - covering theoretical concepts and algorithms. Use of slides and diagrams to convey key ideas.
  - **In-class coding demos** – demonstrating programming examples and machine learning workflows on notebooks.
  - **Hands-on labs** – providing structured labs to implement models and algorithms with clear instructions.
  - **Project-based learning** - Assign sizable projects for students to work on applying multiple models and techniques end-to-end.

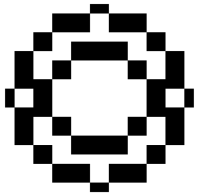


EVALUATION

- Regular (1st examination period)
  - Exam (50%)
  - Practical handout (10%)
  - Final Project (40%)

- Resit (2nd examination period)
  - Exam (60%)
  - Final Project (40%)

- A minimum grade of 8.0 (in 20) for the exam
- A minimum grade of 5.0 (in 20) for the projects



REFERENCES

- Pattern Recognition and Machine Learning, Christopher M. Bishop
- Machine Learning Yearning; technical strategy for AI engineers in the era of deep learning. Andrew NG. available here
- Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies (The MIT Press), John D. Kelleher, Brian Mac Namee , Aoife D`arcy
- Mastering Machine Learning with Python in Six Steps. A Practical Implementation Guide to Predictive Data Analytics Using Python. Swamynathan, Manohar

Questions?