MINI-COURSE

Workshop on Statistics and Applications, ISEG – Universidade de Lisboa, 4–5 December 2026

From Linear Models to Neural Networks: A Journey Through Machine Learning

Concepts, Methods and Applications with Python

Instructor: Jorge Caiado, PhD – jcaiado@iseg.ulisboa.pt

Course overview

Artificial Intelligence is reshaping how we analyse data, make forecasts and support decisions. At the heart of this shift is machine learning: a family of methods that let computers learn patterns, trends and relationships directly from data, without being explicitly programmed for each task. This mini-course offers a compact, rigorous and practical tour of that field, following the structure of the book *Machine Learning com Python* (Caiado, 2026).

The session is built around a problem-solving approach: rather than treating algorithms as isolated recipes, it walks participants through a complete data-science workflow, from importing and preparing data, through model training and evaluation, to the critical interpretation of results. Every topic is grounded in Python, using the ecosystem's most widely used libraries (NumPy, Pandas, Matplotlib, Seaborn and Scikit-learn), and illustrated with examples on real datasets.

The course is deliberately broad but practical: participants leave with a clear mental map of the main supervised and unsupervised methods, an understanding of when each is appropriate, and the vocabulary and workflow needed to apply them to their own research and professional problems.

Audience and prerequisites

Designed for graduate students and researchers in quantitative disciplines: statistics, econometrics, data science, economics, management, finance, engineering and the social sciences. Participants benefit from basic familiarity with statistics and some exposure to programming, but no prior machine-learning experience is assumed. A laptop with a working Python environment (Anaconda or equivalent) is recommended for the hands-on segment.

Learning objectives

By the end of the course, participants will be able to:

1. Structure a complete machine-learning project, from data preparation and exploration to model evaluation and interpretation.
2. Distinguish supervised from unsupervised learning and select appropriate methods for a given problem.
3. Explain and apply the core supervised algorithms: linear and logistic regression, decision trees, random forests, support vector machines, gradient boosting (XGBoost) and neural networks.
4. Apply unsupervised techniques for clustering, dimensionality reduction and anomaly/outlier

detection.

5. Evaluate models rigorously using cross-validation and the right performance metrics, while recognising the bias–variance trade-off and the risk of overfitting.
6. Implement and interpret these methods in Python with Scikit-learn and related libraries.

Course content

Part I: Foundations of machine learning

- Introduction to Python for data science
- Data preparation
- Training, validation and evaluation

Part II: Supervised learning

- Linear regression
- Logistic regression
- Decision trees
- Random forests
- Support vector machines
- XGBoost
- Neural networks

Part III: Unsupervised learning & anomaly detection

- Cluster analysis
- Principal component analysis (PCA)
- Multidimensional scaling (MDS)
- DBSCAN
- Isolation Forest

Reference book

Caiado, J. (2026). *Machine Learning com Python*. Lisboa: Edições Sílabo. ISBN 978-989-561-479-0.

About the instructor

Jorge Caiado

Jorge Caiado holds a PhD in Applied Mathematics to Economics and Management and is an Associate Professor with Habilitation at ISEG — Lisbon School of Economics and Management, University of Lisbon, where he coordinates the postgraduate programme in Applied Artificial

Intelligence and Machine Learning. He is an integrated researcher at ISEG Research and co-founder of GlobalSolver, a technology company specialising in artificial intelligence and big data. Over his career he has carried out research, consulting and training in data science, artificial intelligence, forecasting methods and statistical analysis, collaborating with universities, central banks, financial institutions, public bodies and national and international companies. He is the author of numerous scientific articles and technical books in statistics, artificial intelligence, data science, economics, finance and management — including *Inteligência Artificial: Guia para Iniciantes*, *Python e Exploração de Dados*, *Time Series Clustering and Classification*, *Métodos de Previsão em Gestão* and *Gestão de Instituições Financeiras*. He received first prize in the SICO2007 award from the Spanish IEEE Computational Intelligence Society, and is co-inventor of an international patent published in the United States in the area of optimisation, centralisation and negotiation of inter-company (B2B) purchasing.