Course guide
200250 - DSAF - Data Science for Finance

Last modified: 17/07/2025

Unit in charge: School of Mathematics and Statistics
Teaching unit: 723 - CS - Department of Computer Science.

Degree: BACHELOR'S DEGREE IN MATHEMATICS (Syllabus 2009). (Optional subject).

Academic year: 2025 **ECTS Credits:** 3.0 **Languages:** English

LECTURER

Coordinating lecturer: ARGIMIRO ALEJANDRO ARRATIA QUESADA

Others: Primer quadrimestre:
ARGIMIRO ALEJANDRO ARRATIA QUESADA - A

PRIOR SKILLS

Foundations of Machine Learning, Data Science. Basic knowledge of ML models such as neural networks, support vector regressors. Basic Statistics. Knowledge of R (preferable) or Python

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
GM-CE1. CE-1. Propose, analyze, validate and interpret simple models of real situations, using the mathematical tools most appropriate to the goals to be achieved.
GM-CE2. CE-2. Solve problems in Mathematics, through basic calculation skills, taking in account tools availability and the constraints of time and resources.
GM-CE3. CE-3. Have the knowledge of specific programming languages and software.
GM-CE4. CE-4. Have the ability to use computational tools as an aid to mathematical processes.
GM-CE5. Knowing how to use search tools of bibliographic resources in Mathematics.
GM-CE6. Ability to solve problems from academic, technical, financial and social fields through mathematical methods.

Generical:
GM-CB1. CB-1. Demonstrate knowledge and understanding in Mathematics that is founded upon and extends that typically associated with Bachelor's level, and that provides a basis for originality in developing and applying ideas, often within a research context.
GM-CB2. CB-2. Know how to apply their mathematical knowledge and understanding, and problem solving abilities in new or unfamiliar environments within broader or multidisciplinary contexts related to Mathematics.
GM-CB3. CB-3. Have the ability to integrate knowledge and handle complexity, and formulate judgements with incomplete or limited information, but that include reflecting on social and ethical responsibilities linked to the application of their knowledge and judgements.
GM-CB4. CB-4. Have the ability to communicate their conclusions, and the knowledge and rationale underpinning these to specialist and non-specialist audiences clearly and unambiguously.
GM-CG1. CG-1. Show knowledge and proficiency in the use of mathematical language.
GM-CG3. CG-3. Have the ability to define new mathematical objects in terms of others already know and ability to use these objects in different contexts.
GM-CG4. CG-4. Translate into mathematical terms problems stated in non-mathematical language, and take advantage of this translation to solve them.
GM-CG6. CG-6 Detect deficiencies in their own knowledge and pass them through critical reflection and choice of the best action to extend this knowledge.
GM-CB5. To have developed those learning skills necessary to undertake further interdisciplinary studies with a high degree of autonomy in scientific disciplines in which Mathematics have a significant role.

Transversal:
04 COE. EFFICIENT ORAL AND WRITTEN COMMUNICATION. Communicating verbally and in writing about learning outcomes, thought-building and decision-making. Taking part in debates about issues related to the own field of specialization.
05 TEQ. TEAMWORK. Being able to work as a team player, either as a member or as a leader. Contributing to projects pragmatically and responsibly, by reaching commitments in accordance to the resources that are available.
06 URI. EFFECTIVE USE OF INFORMATION RESOURCES. Managing the acquisition, structure, analysis and display of information from the own field of specialization. Taking a critical stance with regard to the results obtained.
07 AAT. SELF-DIRECTED LEARNING. Detecting gaps in one's knowledge and overcoming them through critical self-appraisal. Choosing the best path for broadening one's knowledge.
03 TLG. THIRD LANGUAGE. Learning a third language, preferably English, to a degree of oral and written fluency that fits in with the future needs of the graduates of each course.

TEACHING METHODOLOGY

Lectures combine theory and practice (R scripts for model exploration of time series and other examples will be provided). Attendance at classes is required, and submission of homeworks.

LEARNING OBJECTIVES OF THE SUBJECT

The course subjects of study range across themes from machine learning, mathematical finance, numerical methods and computer algorithms. There are two main objectives:1) To adquire knowledge about financial markets, their functioning and products, and in general to understand the behavior of financial time series, their statistical properties. 2) To learn the design and proper assessment of financial forecasting models and investment strategies based on supervised learning models or other models that use different types of information sets (quantitative and qualitative).

STUDY LOAD

Type	Hours	Percentage
Hours large group	15,0	20.00
Self study	45,0	60.00
Hours small group	15,0	20.00

Total learning time: 75 h

CONTENTS

1. An abridged introduction to finance and ML

Description:
Securities (bonds, stocks,derivatives); price and payoff; market indices; market jargon; financial markets zoo. Essential aspects of data exploration and learning. A review of software in R for ML.

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

2. Statistics of Financial Time Series

Description:
Descriptive statistics of financial time series. Returns. Stylized facts of financial returns. Forecasting (formal definition). Volatility.

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

3. Time series model adequacy methods.

Description:
Time series model adequacy. Causality. Independence. Correlation. The stationary bootstrap (for synthetic replication of time series). Feature selection through LASSO. (These are all useful techniques to select features for our time series forecasting models)

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

4. Financial time series models I

Description:
Financial Time Series Models. Econometric models: (linear) AutoRegressive Moving Averages models; (nonlinear) ARCH and GARCH. ML models: Feed forward (1-layer) Neural Networks (Nnet).
An introduction to Financial Fundamental Analysis (Graham and Dobb theory of investment)
Forecasting time series with Nnet and business fundamental indicators.

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

5. Financial time series models II

Description:
ML models II: Kernel machines. Support Vector Machines (SVM). Gaussian processes (GP).
Forecasting financial time series with GP, Nnet, SVM, with alternative data (e.g. News-based sentiment indicators). For this we shall review Sentiment Analysis and financial applications.

Full-or-part-time: 2h 50m
Theory classes: 1h 30m
Guided activities: 1h 20m

6. Algorithmic trading. Portfolio theory I

Description:
Algorithmic trading.
Portfolio optimization. Markowitz mean-variance model. Expected utility maximization theory.

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

7. Portfolio theory II.

Description:
Factor models of returns with alternative data. Robust portfolio optimization. The Machine Learning portfolio

Full-or-part-time: 2h 40m
Theory classes: 1h 30m
Guided activities: 1h 10m

8. Optimization Heuristics in Finance

Description:
Heuristic optimization. Simulated Annealing (SA). Genetic Programming (GenP).
Applications: Calibrating GARCH models. Optimization of portfolios with computationally hard constraints (with SA). Finding profitable trading rules with GenP.

Full-or-part-time: 3h
Theory classes: 1h 40m
Guided activities: 1h 20m

9. Option pricing models I.

Description:
Options. Type of options (European, American, Asian and other). The Black-Scholes Formula for Valuing European Options (a brief review). Monte-Carlo valuation of options I.

Full-or-part-time: 2h 50m
Theory classes: 1h 40m
Guided activities: 1h 10m

10. Option pricing models II. Research directions.

Description:
Monte Carlo valuation of options II. Valuing options with Gaussian processes. Other heuristics for valuing options. Review of further research topics in Machine Learning for Finance.

Full-or-part-time: 2h 50m
Theory classes: 1h 40m
Guided activities: 1h 10m

GRADING SYSTEM

There will be no written exam. The evaluation consists of take-home works (2), consisting of some R explorations and exercises to complement the theory.

BIBLIOGRAPHY

Basic:
- Arratia, Argimiro. Computational finance : an introductory course with R [on line]. Paris: Atlantis Press, cop. 2014 [Consultation: 28/06/2023]. Available on: https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=6312413. ISBN 9789462390690.