centro adscrito a:

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

TEACHING GUIDE TO DATA ANALYTICS, ARTIFICIAL INTELLIGENCE, AND MACHINE LEARNING IN FINANCE
2025- 2026

GENERAL DATA

Name:	Data Analytics, Artificial Intelligence, and Machine Learning in Finance
Code:	
Course:	2025-26
Titration:	Master's Degree in Financial Innovation and Fintech
Number of credits (ECTS):	5
Location in the curriculum:	1st year 2nd semester
Department:	
Head of department:	
Date of last revision:	April 2025
Teaching staff:	Gerard Albà (Mathematician, CIO MoraBanc), Eric Risco (Computer Scientist, AndorraTelecom), Paul Gantheil (MSc Finance, Financial Analyst Creand, Superprof)

1. OVERVIEW

Data science is at the core of any growing modern business. The business information obtained from the use and analysis of data allows us to improve the effectiveness, quality and efficiency of business activity. Its importance is increasingly strategic for any business model.

Likewise, industry and many sectors are transforming, turning data into productive knowledge. Two major forces drive this evolution. On the one hand, programmatic access to varied and massive sources of data (*Big Data*), with increasing processing speeds. And, on the other hand, the second major force, which will continue to be dominant in the coming years: the development of machine learning algorithms and the growing importance of generative artificial intelligence (*GenAI).*

This new data-driven process allows us to build systems that automatically learn from experience. And they can be applied to many types of industry-specific problems, as we will see during the course, by modeling complex relationships or classification tasks, where traditional models are not suitable.

We will explore in the sessions, use cases in depth (we will also see their corresponding implementation in code), including the main categories of tools that data science covers: supervised and unsupervised machine learning, reinforcement learning, deep learning and natural language processing (NLP), large language models (LLMs) and generative artificial intelligence, and accelerated computing (HPC).

2. OBJECTIVES

- Understand the fundamentals of data science applied to finance and its impact on decision-making in banking, insurance, and markets.
- Know the key stages of a financial analytics project, from understanding the business and objectives to the processing, analysis, and visualization of financial data.
- Master the practical application of the most widely used data tools and technologies in the financial sector, such as Python, R and their main libraries for numerical analysis, visualization and machine learning (Pandas, NumPy, Scikit-learn, TensorFlow, etc.).
- Apply supervised and unsupervised machine learning techniques (regression, classification, clustering, association) to real problems in the financial environment.
- Explore advanced methods such as Reinforcement Learning, with algorithms such as Q-Learning, SARSA and Deep Q-Learning applied to financial decisions.
- Introduce yourself to the use of generative artificial intelligence and large language models (LLMs), including concepts such as transformer architecture, embeddings, agents, and RAG, with their applications in finance.
- Integrate data science into key areas of finance:
    - Valuation of assets and derivatives.
    - Multi-asset portfolio management.
    - Personalization of financial services.
    - Credit and actuarial risk assessment.
    - Algorithmic trading.
    - Economic analysis with nowcasting.
    - Analysis of cryptoassets and blockchain.
- Apply knowledge to real practical cases, developing analytical and technical skills in simulated or real industry scenarios.

3. LEARNING OUTCOMES

At the end of the course, the student will be able to:

K1.2: Describe the main technological trends that are redefining the financial sector.
K6.1: Identify the different types of crypto assets, the mining process, the platforms, and the associated risks and benefits.

K6.2: Identify the tokenization process, the different types and platforms and main use cases.
K9.1: Identify the different elements of the business model and define the competitive advantages of a technological nature that make it differential and sustainable.
S1.1: Communicate effectively orally, in writing and graphically with other people about learning, thinking and decision-making, and participate in debates, making use of interpersonal skills, such as active listening and empathy, which favour teamwork.
S2.1: Develop the capacity to contribute to innovation in new or existing business institutions and organizations, through participation in creative projects and have the ability to apply skills and knowledge on technology-based business sales, organization and development.
S3.1: Understand advanced digital technologies, so that they can be applied with a critical perspective, in diverse contexts, in academic, professional, social or personal situations.
S4.1: Differentiate the different types of technology that are being applied to the financial sector based on disintermediation.
S9.1. Differentiate the different types of cryptoassets and digital assets.
S9.2: Differentiate the use cases of tokens, the Ethereum platform, and decentralized applications.
C1.1: Integrate the values of sustainability, understanding the complexity of systems, in order to undertake or promote actions that restore and maintain the health of ecosystems and improve justice, generating diverse visions for sustainable futures.
C2.1: Identify and analyse problems that require autonomous, informed and reasoned decision-making, in order to act with social responsibility, in accordance with ethical values and principles.
C3.1: Develop the capacity to assess gender and gender inequalities and to design solutions.
C4.1: Apply financial decentralization technologies in applications or business models that enable cost reduction and improved profitability, considering
the current ethical and legal framework.

C9.1: Apply new technologies and application of banking digitalisation to new development projects.

C9.2: Apply success stories in technology and financial disintermediation to new business outlets.

We can highlight:

- Interpret and apply key data science concepts in the financial context.

- Design and develop financial data analytics projects, from data preparation to visualization of results.
- Use programming tools and languages (such as Python or R) for data analysis and predictive modeling in Finance problems.
- Apply supervised and unsupervised machine learning techniques to real financial problems.
- Implement reinforcement learning models in financial scenarios.
- Understand and apply generative AI models and language models (LLMs) for tasks such as textual analysis, prediction, and automation in finance.
- Analyze and solve real cases in areas such as asset valuation, portfolio management, risk analysis, algorithmic trading, and blockchain and cryptoasset analysis.
- Critically evaluate the impact of AI and big data technologies on the financial sector, with an ethical and strategic approach.

4. CONTENTS

TOPIC 1: INTRODUCTION TO DATA SCIENCE APPLIED TO FINANCE

Specific learning outcomes:

- Understand the fundamental principles of data science applied to finance, including its strategic and transformative potential.
- Recognize representative use cases where data science brings value in banking, insurance, investments, and other financial sectors.

Contents:

1.1. Change of model: from theory-based finance to data-based finance.

1.2. Application of AI to the different areas of the financial sector.

TOPIC 2: ELEMENTS AND STAGES OF A DATA SCIENCE PROJECT IN FINANCE

Specific learning outcomes:

- Understand the financial context of the project. Translating business needs into analytical objectives.
- Prepare and adapt data for use in exploratory analytics (EDA) and machine learning models.

- Apply statistical and visual techniques to explore datasets. Identify patterns, correlations, outliers, and relationships between variables. Use EDA as a preliminary step for the selection of appropriate algorithms.
- Know different data structures used in finance and data science: tables, graphs, trees, etc.
- Evaluate model performance and adjust parameters according to the business objective.
- Know and apply machine learning algorithms Supervised.
- Know and apply Unsupervised machine learning algorithms.
- Know and apply Reinforcement learning algorithms.
- Understand the functioning **of Generative Artificial Intelligence (*GenAI):*** *Large Language Models* (LLMs) and their application in finance.
- Know the most used programming languages and libraries in data science for finance: Python and R.

Contents:

2.1 Introduction.

2.2 Elements and stages of financial analytics. Tools, technological infrastructure and programming languages most used in the financial industry.

2.2.1 Understand the Business and the objectives of the project. Examples in Finance.

2.2.2 Data.

2.2.2.1 Financial data: sources, cleaning, pre-processing visibility. Data architecture.

2.2.2.2 Data Structures.

2.2.2.3 Data Processing.

2.2.2.4 *Exploratory Data Analysis (EDA*).

2.2.3 Algorithms.

2.2.3.1 Predictive models.

2.2.3.2 Machine *Learning*: Supervised, Unsupervised, Reinforcement *Learning*.

2.2.3.3 Supervised Learning: Regression and Classification. Linear and logistic regression; Naïve Bayes; Support Vetor Machines (SVM); Decision trees; Random forests; k-NN; Gradient *boosting*; Neural networks.

2.2.3.4 Unsupervised learning. Clustering: K-Means, hierarchical, Gaussian models; Association: Apriori algorithm to discover relationships between variables; Dimensionality Reduction: Principal Component Analysis (PCA).

2.2.3.5 *Reinforcement Learning.*

2.2.3.5.1 Q-Learning.

2.2.3.5.2 SARSA-Learning.

2.2.3.5.3 Deep Q-Learning.

2.2.4 Generative AI and Language Models: *Large Language Models* (LLMs).

2.2.4.1 Transformer Architecture.

2.2.4.2 Tokens y *Embeddings.*

2.2.4.3 Agents.

2.2.4.4 Retrieval-Augmented Generation (RAG).

2.2.4 Programming languages: Python and *R*. Use key libraries for analysis and modeling:

- Pandas: data manipulation.
- NumPy and SciPy: scientific and numerical calculations.
- Matplotlib: Generation of graphs and visualizations.
- Scikit-learn: traditional machine learning models.
- TensorFlow and Keras: development of neural networks.
- PyTorch: Advanced modeling in *deep learning*.

TOPIC 3: HOW DATA SCIENCE IS USED IN FINANCE

Specific learning outcomes:

- Understand the role of data science in market analysis and quantitative decision-making.
- Model financial time series to identify trends and seasonalities.
- Machine learning and AI models applied to financial derivatives. Value options, futures, and swaps using numerical and analytical valuation models. Assess and manage the risk associated with derivatives using measures.
- Design and optimize diversified portfolios of financial assets following mean-of-variance criteria and other modern approaches.
- Extract data from unstructured sources (news, social media, reports).

- Apply NLP techniques to measure sentiment and qualitative variables that improve predictive models.
- Application of machine learning models to Financial risks. Develop models to predict delinquency, fraud and operational risks. Implement anomaly detection algorithms to monitor critical events. Estimate risk metrics (credit, market, liquidity) and validate their robustness.
- Segment customers using clustering and scoring algorithms. Design recommendation systems and personalized marketing strategies.
- Predict customer behaviors and needs through predictive models.

Contents:

3.1 Financial Markets and Quantitative Finance.

3.1.1 The prices of financial assets.

- Apply data analysis techniques to model and forecast the behavior of asset prices.
- Use time series, stochastic models, and simulations to study market movements.

3.1.2 Valuation and management of derivative instruments.

- Employ quantitative tools based on machine learning to value derivatives such as options, futures, and swaps.
- Assess the risk associated with these products using mathematical and computational models based on data science.

3.1.3 Multi-asset portfolios and applied machine learning.

- Optimize diversified portfolios using optimization models.
- Analyze the correlation between assets and adjust the data-driven asset allocation.

3.1.4 Unstructured feeling and information.

- Extract value from unstructured data (news, social media, reports).
- Apply natural language processing (NLP) to gauge market sentiment.
- Integrate this information into predictive and risk management models.

3.2 Risk Control and Management in Banking and Insurance.

- Apply models to predict delinquency, fraud and operational risks.
- Use classification and anomaly detection algorithms to prevent underwriting losses and risk origination.

- Estimate credit, market, and liquidity risk with advanced data analysis tools.

3.3 Customer Personalization and Analytics.

- Use customer segmentation techniques (*clustering*) to personalize financial services.
- Analyze customer behavior to improve products, marketing, and retention.
- Predict needs and behaviors using machine learning and advanced analytics.

TOPIC 4: CASE STUDIES

Specific learning outcomes:

- Apply customer segmentation methods and recommendation algorithms to create customized investment solutions.
- Develop predictive credit risk models to estimate defaults and manage credit exposure.
- Apply actuarial methods and machine learning techniques to manage risk and determine prices in insurance products.
- Use *nowcasting* models to forecast economic indicators with real-time data.
- Design and optimize automated trading algorithms to execute trades in financial markets.
- Use data science techniques to analyze blockchain and crypto assets, improve decision-making, and forecast market behavior.

Contents:

4.1 Personalized Financial Advice.

4.2 Financial Risk Management: Credit.

4.3 Actuarial Risk Management in Insurance.

4.4 Economy and Nowcasting.

4.5 Algorithmic Trading.

4.6 Data Science for Blockchains and Cryptoassets.

5. TEACHING AND LEARNING METHODOLOGY

Teaching is carried out through a series of face-to-face sessions, whose spirit is to combine theory and practice in a balanced way, complemented by a series of didactic materials (manual and presentation), exercises and cases for their resolution, provided online and with *feedback*, also online, using the program's own platform as a digital support.

The face-to-face activities will be carried out through:

- Presentation by the teacher of the practical application of the theoretical contents of the different topics of the subject. It is essential that the student reads beforehand (by the student) the reference manual for each topic that is available in the virtual classroom.
- Discussion of content.
- Application of concepts and methodologies to examples/case studies

Learning will be consolidated through the resolution of the exercises and/or cases that will be provided on the virtual campus of the subject.

6. EVALUATION

In accordance with the Bologna Plan, the model rewards the constant and continuous effort of students. 60% of the grade is obtained from the continuous evaluation of the directed activities and the remaining 40% from the final face-to-face exam. The final exam has two sittings.

The final grade of the subject (NF) will be calculated based on the following formula:

- **NF = Final Exam Grade x 40% + Continuous Evaluation Grade x 60%**
- The minimum grade of the final exam to calculate the NF will be 40 points out of 100.
- The subject is approved with an NF equal to or greater than 50 points out of 100.

Type of activity	Description	% Continuous evaluation	
Questionnaires:			**60%**
Questionnaire 1	Test Topic 1	20%	
Questionnaire 2	Test Topic 2	20%	
Questionnaire 3	Test Topic 3	20%	
Final examination			**40%**
	Final examination	100%	

7. BIBLIOGRAPHY

7.1. BASIC BIBLIOGRAPHY

Albà et al, G. (2023). *La tecnología cuantitativa de los roboadvisors.* IEF-Instituto de Estudios Financieros.

AQR. (2024). *Machine Learning: Why Finance Is Different.* Obtenido de https://www.aqr.com/Learning-Center/Machine-Learning/Machine-Learning-Why-Finance-Is-Different

BlackRockEngineering. (2022). *I feel the need, the need for speed.* Obtenido de https://engineering.blackrock.com/i-feel-the-need-the-need-for-speed-92afe9e64452

Damon, L. (2020). *Build Your Own AI Investor.*

Dixon et al, M. (2020). *Machine Learning in Finance From Theory to Practice.* Springer.

Giles, M. (2024). *Oxford Mathematical Institute.* Obtenido de https://people.maths.ox.ac.uk/~gilesm/

Hilpisch, Y. (2020). *Artificial Intelligence in Finance.* O'Reilly.

HSLU. (2023). *GPT for Financial Advice.* Institute of Financial Services Zug IFZ.

Hull, I. (2021). *Machine Learning for Economics and Finance in TensorFlow 2 Deep Learning Models for Research and Industry.* Apress.

Isaiah, H. (2021). Decision Trees. En H. Isaiah, *Machine Learning for Economics and Finance in TensorFlow 2.* Apress.

Jansen, S. (2020). *Machine Learning for Algorithmic Trading.* Packt.

JPMorgan. (2017). *(JPMorgan, Cross Asset Portfolios of Tradable Risk Premia Indices Hierarchical Risk Parity: Enhancing Returns at Target Volatility, 2017).* Global Quantitative & Derivatives Strategy.

Kaabar, S. (2024). *Deep Learning for Finance.* O'Reilly.

Karasan, A. (2022). *Machine Learning for Financial Risk Management with Python.* O'Reilly.

Mazars. (2019). *Deep hedging Application of deep learning to hedge financial derivatives.*

NVIDIA. (2007). *Computational Finance in GPUs.* Obtenido de https://developer.download.nvidia.com/compute/cuda/1.1-Beta/x86_website/Computational_Finance.html

Tatsat et al, H. (2020). Blueprint for Building a Trading Strategy Based on Sentiment Analysis. En H. Tatsat et al, *Machine Learning and Data Science Blueprints for Finance.*

Tatsat et al, H. (2020). Blueprint for Creating a Machine Learning Model for Predicting Loan Default Probability. En H. Tatsat et al, *Machine Learning and Data Science Blueprints for Finance.* O'Reilly.

Tatsat et al, H. (2020). Blueprint for Using Classification Models to Determine Whether a Transaction Is Fraudulent. En H. Tatsat et al, *Machine Learning and Data Science Blueprints for Finance.* O'Reilly.

Tatsat et al, H. (2020). Blueprint for Using Clustering to Implement Hierarchical Risk Parity. En *Machine Learning and Data Science Blueprints for Finance.* O'Reilly.

Tatsat et al, H. (2020). Case Study 2: Chatbot Digital Assistant. En H. Tatsat et al, *Machine Learning and Data Science Blueprints for Finance.*

Tatsat et al, H. (2020). *Machine Learning and Data Science Blueprints for Finance.* O'Reilly.

Tatsat et al, H. (s.f.). Blueprint for Using Clustering for Grouping Investors. En H. Tatsat et al, *Machine Learning and Data Science Blueprints for Finance.* O'Reilly.

7.2 BIBLIOGRAFÍA COMPLEMENTARIA

Alexander, C. (2008). *Market Risk Analysis, Quantitative Methods in Finance.* Wiley.

Alonso, A. (. (2021). *Understanding the performance of machine learning models to predict credit default: a novel approach for supervisory evaluation.* Banco de España.

Atlanta, F. R. (s.f.). *GDPNow*. Obtenido de https://www.atlantafed.org/cqer/research/gdpnow

Brownlee, J. (2018). *Deep Learning for Time Series Forecasting Predict the Future with MLPs, CNNs and LSTMs in Python.* Machine Learning Mastery.

Castillo, G. (2023). *Data Science for Web3: A comprehensive guide to decoding blockchain data with data analysis basics and machine learning cases.* Packt.

Cesa et al, M. (2022). JP Morgan testing deep hedging of exotics. *Risk magazine*.

ChenDataBytes. (2024). *Fraud Detection and Anomaly Detection Start from supervised learning to unsupervised learning*. Obtenido de https://medium.com/@chenycy/fraud-detection-and-anomaly-detection-e65dd11a3146

Citi. (2023). *Unleashing AI.* Citi GPS: Global Perspectives & Solutions.

DeepMind, G. (2015). *From Pixels to Actions: Human-level control thorugh Deep Reinforcement Learning*. Obtenido de https://research.google/blog/from-pixels-to-actions-human-level-control-through-deep-reinforcement-learning/

Denuit et al, M. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems.* Wiley.

Domashova et al, J. (2021). Usage of machine learning methods for early detection of money laundering schemes. *Procedia Computer Science*, 184-192.

Fama et al, E. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*.

Giannone et al, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 665-676.

HUANG, A. (2021). *Stock recommandation via machine learning algorithms.* Obtenido de https://github.com/WideSu/Stock-recommandation-via-machine-learning-algorithm.

JPMorgan. (2017). *(JPMorgan, Cross Asset Portfolios of Tradable Risk Premia Indices Hierarchical Risk Parity: Enhancing Returns at Target Volatility, 2017).* Global Quantitative & Derivatives Strategy.

Kaggle. (2019). *Anomaly Detection with Unsupervised Learning*. Obtenido de Anomaly Detection with Unsupervised Learning: https://www.kaggle.com/code/jiedong00/anomaly-detection-with-unsupervised-learning

Kaggle. (2021). *Credit Card Fraud Detection*. Obtenido de https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Kaggle. (s.f.). *Kaggle*. Obtenido de Kaggle.com

Karasan, A. (2022). *Machine Learning for Financial Risk Management with Python.* O'Reilly.

Krikorian, M. (2021). *Fraud Detection applying Unsupervised Learning techniques*. Obtenido de https://medium.com/southworks/fraud-detection-applying-unsupervised-learning-techniques-4ae6f71b266f

Le Borgne et al, Y. (2022). *Reproducible machine learning for credit fraud detection - practical handbbook*. Obtenido de https://fraud-detection-handbook.github.io/fraud-detection-handbook/Foreword.html#authors

Mass, C. (2011). Nowcasting: The Next Revolution in Weather Prediction. *Bulletin of the American Meteorological Society*.

Mazars. (2019). *Deep hedging Application of deep learning to hedge financial derivatives.*

Mazars. (s.f.). *Deep hedging Application of deep learning to hedge financial derivatives.*

Microsoft. (2024). *wiss Re Launches Swiss Re Life Guide Scout, a Generative AI-Powered Underwriting Assistant*. Obtenido de https://news.microsoft.com/de-ch/2024/04/22/swiss-re-launches-swiss-re-life-guide-scout-a-generative-ai-powered-underwriting-assistant/

OpenBB. (2023). *Creating an AI powered financial analyst*. Obtenido de OpenBB: https://openbb.co/blog/creating-an-ai-powered-financial-analyst

Oxford, U. o. (2018). *Deep learning approach to hedging.* University of Oxford .

Rachev et al, S. (2006). *Financial Econometrics: From Basics to Advanced Modeling Techniques.* Wiley.

Reacfin. (2017). *Machine learning aplications to non-life pricing.*

Reese et al, R. (2017). Fraud detection of insurance claims. En R. Reese et al, *Machine Learning: End-to-End guide for Java developers.* Packt.

Santander, B. (2016). *Santander Product Recommendation Can you pair products with people?* Obtenido de https://www.kaggle.com/competitions/santander-product-recommendation/overview/description

Tegmark, M. (2017). *Life 3.0.* Penguin Books.

Tibshirani, R. (2002). Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *Statistical Methodology* , 411-423.

York, F. R. (2017). Macroeconomic Nowcasting and Forecasting with Big Data.

Zeng et al, Q. (2023). *Sentiment Analysis in Finance*. Obtenido de https://github.com/Apricot0/Sentiment-Analysis-in-Finance

Zuckerman, G. (2019). *The Man Who Solved the Market.* Penguin.