Année universitaire 2024/2025

Digital Economics - 297 - 2nd year of master's degree

Crédits ECTS : 60

LES OBJECTIFS DE LA FORMATION

This academic track provides expert training in digital economics and methods for analyzing mass data. Harnessing this type of data requires new skills to be able to process high volumes of input and extract useful information. This track therefore aims to train quantitative economists in processing and modeling large, complex datasets to shed light on the decisions of businesses and institutional stakeholders. Employment opportunities are highly varied: data analyst, consultant, economic expert, etc.

Skills acquired:

- Make use of statistical and econometric tools to obtain reliable and robust answers, to shed light on businesses and public or semi-public institutions' options
- Learn about quantitative methods for processing massive databases
- Get trained in computer programming to process large and complex databases
- Report on the results of economic, statistical and/or econometric results to different audiences, orally and in writing

PRÉ-REQUIS OBLIGATOIRES

- Applicants should hold a postgraduate degree (Master degree equivalent to 60 ECTS), from either the Master 1 in Quantitative Economics or another Master at Université Paris-Dauphine-PSL, or another university in France or abroad; or an equivalent diploma from an Institute of Political Studies (IEP), a Grande Ecole in business or engineering, or an equivalent recognized Grand Etablissement, in France or abroad.
- Applicants should have an academic background in the following fields: Economics, mathematics applied to economics, computer science applied to economics, or any other educational program with a quantitative component
- B2-level mastery of English is required. This must be attested by a certificate of achievement from one of the following tests: TOEFL iBT (minimum score of 90), IELTS (minimum score of 6.5) or Cambridge certificate (C1). English-native candidates or students who have followed an international training in English of at least one year over the last two years and who have passed the corresponding exams are exempted.

Executive Education :

- Students with an 1st year of master's degree of at least 60 ECTS in economics, mathematics, mathematics applied to social sciences
- Professionals in the digital economy, data analysts, data scientists; executives, senior executives work in the field of the digital economy (Bac + 4)

PROGRAMME DE LA FORMATION

- Semester 3
  - Mandatory
    - Python for Data Science
    - Machine Learning

        - Data Science Project
        - Industrial Organization
        - Blockchain economics
        - Financial Data & Systemic risk
        - Time Series and Anomaly Detection
- Semester 4
    - Mandatory
        - NLP for economic decisions
        - Neural Networks
        - Private Cryptocurrencies
        - Solidity and smart contract development
        - Machine Learning for Economists
        - Business Cases
        - Internship
    - Optional - 3 ECTS
        - Anonymization, privacy
        - Computational social choice
        - Incremental learning, game theory and applications

DESCRIPTION DE CHAQUE ENSEIGNEMENT

Anonymization, privacy

ECTS : 3

Blockchain economics

ECTS : 6

Description du contenu de l'enseignement :

While this is a fairly recent technology, this class will take students through the fundamentals of blockchains as well as implications regarding financial, economic or social interactions. The class will start by some history needed to understand what lead to the creation of Bitcoin, the first blockchain, in 2009. We will then review the detailed functioning of a blockchain. We will continue by discussing important current developments in the industry as well as implications for the economic environment. Lastly, we will discuss potential future developments and how blockchains will impact a broad range of industries. Students will also be introduced to recent academic work related to blockchains.

Students will be asked to pick a blockchain project from a list and present it briefly during the presentation session in front of the class (group presentation). When reaching this presentation session, students will be expected to be able to assess the pros and cons of a given blockchain project, and have a critical opinion on this project.

Compétence à acquérir :

The objective of this class is to give students a deep theoretical overview of what a blockchain is. Nonetheless we will also use mock-blockchains, write smart contracts and interact with them, through some computer sessions. This will help solidify the knowledge learned and de-mystify the functioning of a blockchain.

Students will gain a deep knowledge of how a blockchain works internally. They will also be very aware of the different issues and perhaps they will be able to spot new use cases for a blockchain.

Mode de contrôle des connaissances :

The evaluation is composed of a group presentation (1/3), homework (1/3) and a final exam (1/3). Class participation can be highly rewarded especially for students who struggle with homework. Students are encouraged to actively interact during the class.

Bibliographie, lectures recommandées :

Melanie Swan, Blockchain: Blueprint for a new economy, O'Reilly, 2015

Andreas Antonopoulos, Mastering Bitcoin, 2nd edition, O'Reilly, 2017

Andreas Antonopoulos / Gavin Wood, Mastering Ethereum, 1st edition, O'Reilly, 2018

Primavera De Filippi/ Aaron Wright, Blockchain and the Law: The Rule of Code, Harvard University Press, 2018

Business Cases

ECTS : 0

Description du contenu de l'enseignement :

The lecture is a sequence of use-cases in the industry and in the service business performed by professionals.

Compétence à acquérir :

Knowing some of the most common use cases of data sciences in firms decision making business.

Mode de contrôle des connaissances :

Presentation of a resume of the lectures before a jury.

Computational social choice

ECTS : 3

Data Science Project

ECTS : 3

Financial Data & Systemic risk

ECTS : 3

Description du contenu de l'enseignement :

The course will equip students with the necessary knowledge to be able to undertake econometric analysis of the type commonly associated with modern financial econometrics research. Substantial emphasis will be placed on the development of programming skills in Python (or in MATLAB, especially for financial contagion and multivariate analysis).

Course outline:

1. Data collection (CRSP-Compustat, Yahoo-Finance, ECB data warehouse)
2. Market Risk Measurement (Value-at-Risk, Expected Shortfall) – ARCH/GARCH models – univariate time series
3. Backtesting tests for market-risk measurement (independence test, unconditional coverage test, conditional coverage test, super exception)
4. Systemic Risk and Macroprudential regulation (SIFIs identification, MES, SRISK, ?CoVaR) – multivariate time series
5. Principal Component Analysis (absorption ratio computation)
6. Contagion models (direct and indirect effects decomposition)

Compétence à acquérir :

The course provides a deep knowledge of the advanced time series techniques and their application to systemic risk. A technical presentation of these models will be given, before studying applications of these models to systemic risk.

Mode de contrôle des connaissances :

Individual homework assignment.

Bibliographie, lectures recommandées :

Benoit, S., Colliard, J.-E., Hurlin, C. and C. Pérignon (2017) Where the Risks Lie: A Survey on Systemic Risk, *Review of Finance*, 21(1), 109-152.

Benoit, S., Hurlin, C. and C. Pérignon (2019) Pitfalls in Systemic-Risk Scoring, *Journal of Financial Intermediation*, 38, 19-44.

Campbell, S. D. (2004) A Review of Backtesting and Backtesting Procedures, Working paper, Federal Reserve Board.

Christofferson, P. and Pelletier, D. (2004) Backtesting Value-at-Risk: A Duration-Based Approach, *Journal of Financial Econometrics*, 2(1), 84-108.

Du, Z. and J. C. Escanciano (2015) Backtesting Expected Shortfall: Accounting for Tail Risk, *Management Science*.

Diebold, F.X. and K. Yılmaz (2009) Measuring Financial Asset Returns and Volatility Spillovers, with Application to Global Equity Markets. *The Economic Journal*, 119(1), 158-171.

Diebold, F.X. and K. Yılmaz (2012) Better to Give than to Receive: Predictive Directional Measurement of Volatility Spillovers, *International Journal of Forecasting*, 28(1), 57-66.

Incremental learning, game theory and applications

ECTS : 3

Industrial Organization

ECTS : 6

Description du contenu de l'enseignement :

Theory and practice of competition in network industries; antitrust issues; theory of network and network effects; two-sided platforms and pricing

Compétence à acquérir :

Understanding of competition and regulation issues in network and digital economics

Mode de contrôle des connaissances :

Written exam

Bibliographie, lectures recommandées :

Belleflamme-Peitz, Industrial Organization

Internship

ECTS : 12

Description du contenu de l'enseignement :

Students must intern at a company (semester 4) for at least 4 months.

Compétence à acquérir :

The students should also complete an end-of-studies internship lasting at least 4 months. The curriculum includes guest lectures by visiting professionals on issues related to Big Data, providing another means of connecting with relevant business circles.

Mode de contrôle des connaissances :

The internship will conclude with a report to be reviewed by a committee.

Machine Learning

ECTS : 9

Description du contenu de l'enseignement :

The course gives a thorough presentation of the machine learning field and follows this outline:

1. general introduction to machine learning and to its focus on predictive performances (running example: k-nearest neighbours algorithm)
2. machine learning as automated program building from examples (running example: decision trees)
3. machine learning as optimization:
    1. empirical risk minimization
    2. links with maximum likelihood estimation
    3. surrogate losses and extended machine learning settings
    4. regularisation and kernel methods (support vector machines)
4. reliable estimation of performances:
    1. over fitting
    2. split samples
    3. resampling (leave-one-out, cross-validation and bootstrap)
    4. ROC curve, AUC and other advanced measures
5. combining models:
    1. ensemble techniques
    2. bagging and random forests

    3. boosting
6. unsupervised learning:
    1. clustering (hierarchical clustering, k-means and variants, mixture models, density clustering)
    2. outlier and anomaly detection

Compétence à acquérir :

After attending the course the students will

- have a good understanding of the algorithmic and statistical foundations of the main machine learning techniques
- be able to select machine learning techniques adapted to a particular task (exploratory analysis with clustering methods, predictive analysis, etc.)
- be able to design a model selection procedure adapted to a particular task
- report the results of a machine learning project with valid estimation of the performances of their model

Mode de contrôle des connaissances :

- quizzes and tests during the course
- machine learning project

Machine Learning for Economists

ECTS : 3

Description du contenu de l'enseignement :

Economic science has evolved over several decades toward greater emphasis on empirical work. Ever increasing mass of available data ('big data') in the past decade is likely to have a further and profound effect on economic research (Einav and Levin, 2014). Beyond economic research, governments and the industry are also increasingly seeking to use 'big data' to solve a variety of problems, usually making use of the toolbox from machine learning (ML).

The question we ask in this course is the following : What do we (not) learn from big data and ML as economists? Is ML merely applying standard techniques to novel and large datasets? If ML is a fundamentally new empirical tool, how does it fit with what we know? In particular, how does it fit with our tools for causal inference problems? As empirical economists, how can we use big data and ML? We'll discuss in detail how ML is useful to collect new data, for prediction in policy, and to provide new tools for estimation and inference.

Compétence à acquérir :

Course objectives:

1. Present a way of thinking about ML that gives it its own place in the econometric toolbox.

2 Develop an intuition of the problems to which it can be applied, and its limitations.

3. Think of unstructured data (text, image) as data we can use when economic outcomes are missing.

3 Specific focus on application of ML to social policies (health/labor/taxation/environment etc.).

Mode de contrôle des connaissances :

Grading:

1. In-class pairwise presentation of an academic paper (30% of overall grade).
2. Final exam (in-class written test). 70% of overall grade.

Bibliographie, lectures recommandées :

- Mullainathan, Sendhil and Jann Spiess (2017). "Machine learning: An applied econometric approach". In: Journal of Economic Perspective 31.2, pp. 87-106.
- Kleinberg, Jon et al. (2015). "Prediction policy problems". American Economic Review 105.5, pp. 491-495.
- Athey, S. (2017): "Beyond prediction: Using big data for policy problems", Science 355, 483–485.
- Athey, Susan, and Stefan Wager. 2021. "Policy Learning with Observational Data", Econometrica, 89(1): 133-161.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. and S. Mullainathan (2018): "Human Decisions and Machine Predictions", The Quarterly Journal of Economics, Volume 133, Issue 1, Pages 237–293.
- Susan Athey, Guido W. Imbens. 2019. Machine Learning Methods That Economists Should Know About. Annual Review of Economics 11:1, 685-725.
- Athey, Susan, and Guido Imbens. 2016. "Recursive Partitioning for Heterogeneous Causal Effects". PNAS 113(27): 7353–60.
- Belloni, A., V. Chernozhukov, S. Mullainathan and J. Spiess and C. Hansen.(2014): "High-Dimensional Methods and Inference on Structural and Treatment Effects" Journal of Economic Perspectives, Volume 28, Number 2 – Spring 2014,

NLP for economic decisions

ECTS : 3

Neural Networks

ECTS : 3

Private Cryptocurrencies

ECTS : 3

Python for Data Science

ECTS : 0

Solidity and smart contract development

ECTS : 3

Description du contenu de l'enseignement :

This course introduces all major uses cases of the blockchain industry from a technical perspective. The course begins with an introduction of Github and Solidity coding fundamentals before diving into smart contract development. Participants will learn the most common ERC standards for tokens and NFTs before building more complex contracts for DAOs. Finally, a deep dive into the EVM and an outlook into the future of Blockchain - L2s.

The course schedule is as follows:

Lecture 1 - Blockchain Basics and Development

Lecture 2 - Solidity Fundamentals

Lecture 3 - Contracts and Complex Data Structures

Lecture 4 - ERC20 Tokens and Tokenomics

Lecture 5 - Intro to DeFi

Lecture 6 - Further DeFi Applications

Lecture 7 - NFTs

Lecture 8 - ReFi and NFT applications (Guest Lecture)

Lecture 9 - SDLC, Security and Testing

Lecture 10 - DAOs and Governance

Lecture 11 - Assembly and Gas Optimization

Lecture 12 - Scaling the future of Ethereum: L2s

Compétence à acquérir :

At the conclusion of this course, participants will gain a solid foundation of Solidity programming and smart contract development, enough to be considered a junior blockchain developer. Participants will also gain an understanding of the open source philosophy and collaboration style.

Mode de contrôle des connaissances :

The level of mastery will be continuously assessed throughout the course by:

1. A weekly presentation on a topic more in depth than what is presented in the lecture material
2. Weekly homeworks
3. Final smart contract project with oral presentation

Bibliographie, lectures recommandées :

Mastering Ethereum by Andreas Antonopoulos - https://github.com/ethereumbook/ethereumbook

Time Series and Anomaly Detection

ECTS : 6

Description du contenu de l'enseignement :

This lecture is thought as an introduction to the analysis of complex data, and particularly to that having a temporal component. Methods aimed at exploring and modelling time series, longitudinal data and graphs with temporal components will be addressed. The issues of detecting patterns, breakpoints, changes of regimes, and anomalies will be at the core of the different approaches.

The first chapters will be devoted to hidden Markov models. After having briefly recalled some definitions and properties of Markov processes, we will define hidden Markov processes, illustrate them with several examples and give some of their properties. Inference techniques using the EM algorithm and Bayesian approaches will be presented and illustrated in practice. We will particularly focus on some specific models which are extremely useful for segmenting time series stemming from the economics field, such as autoregressive Markov switching models.

The second part of the lecture will tackle the issue of change-point detection methods. We will start by introducing the change-point detection issue. More specifically, we will consider several frameworks and derive inference procedures for computing and locating change-points : online vs. offline strategies, single vs. multiple change point detection, known vs. unknown number of change points, parametric vs. non-parametric approaches.

The third chapter will be aimed at introducing the issue of anomaly detection in the context of temporal data. After having defined what an anomaly is, we will start by assessing whether and how hidden-Markov models and change-point analysis may be useful for detecting anomalies. Then, we will compare these two approaches with other techniques, stemming either from the field of computational statistics, or from that of machine learning. During this chapter, we will also consider the questions of detecting patterns and clustering temporal data.

The fourth chapter will address data that can be modelled as a graph or a temporal graph. We will start by introduce some definitions and summaries for characterising the network (degree distribution, centrality indices, ...). Afterwards, we will tackle the questions of community detection and graph clustering. Eventually, we will address the issues of random networks and associated tests for randomness. Models and methods introduced in this lecture will be practiced using existing implementations in R and “real-life” datasets.

Compétence à acquérir :

Gain some background and perspective of time series analysis from a data science point of view.

Be able to handle temporal data subject to anomalies and change-points.

Have some basic knowledge about graphs and temporal graphs mining.

Mode de contrôle des connaissances :

Data challenge. A project to be done individually or by two, analysing real life data.