

Master 2 Data Science, IP Paris

Journée de pré-rentree/Kickoff meeting, Year 2023-2024

September 4, 2023, École Polytechnique



Organization of the meeting

Morning (10:00-13:00)

10:00 Presentation of the M2DS

10:30 Alumni club of M2DS

10:45 Courses presentations

11:30 Short pause (15 min)

11:45 Courses presentations

Lunch break (13:00-14:30)

13:00 Free token at Magnan

Afternoon (14:30-17:00)

14:30 Courses presentations

15:00 Discussions and questions

M2 Data Science



Objectives

- Become expert in the field of Data Science, machine learning and artificial intelligence.
- Skills in mathematics of statistics, optimization, machine learning, computer science, big data infrastructure.
- Fast moving domain: learn to adapt.
- Multidisciplinary competency: not one good profile.

Facts

- IP Paris and Hi! Paris are major world actors in the domain.
- ≈ 120 students, 50-60% from IP Paris engineer schools.
- 25% of students pursue with a PhD, the rest pursue industry.

Administrative requirement and information

Mandatory registrations:

- **Administrative registration** In your institution of origin (ENSAE,ENSTA, UP Saclay, Telecom paris, Telecom SudParis) or École Polytechnique for new students of IP Paris.
- **Pedagogical registration at École Polytechnique**
 - Contact : `master-admission@ip-paris.fr`
 - Contact for international students : `internationalstudents@ip-paris.fr`

Student Card

- Access to buildings (update on terminals at Télécom/ENSAE/Magnan).
- Access and payment at Magnan restaurant.
- Access to library.

Contact : administrative questions

- Main email : `staffmasterdatascienceipparis@polytechnique.fr`
- Stéphanie Clevenot : `stephanie.clevenot@polytechnique.edu`

Pedagogical contract

Validation of the Master DS

- 42 ECTS from courses :
 - 3 ECTS Data camp (mandatory)
 - 3 ECTS : Master classes and Hi!ckathon (mandatory, evenings +1 weekend)
 - 6 ECTS : CAPSTONE Project or ML research seminar (choice).
 - 30 ECTS from selection of courses.
- 18 ECTS internship (public or private sector).

Rules

- Pedagogical contract must be signed at the beginning of the year.
- Courses and internship are validated for a grade greater or equal to 10/20.
- You can validate two courses in other masters (max 10 ECTS) but it needs agreement from one of the M2 coordinators (and of the other institution).
- ENSAE and ENSTA student can count 10 ECTS from the engineer school.
- M2 year validated when 60 ECTS (42+18) have been validated.

Courses organization

CALENDAR 2023-24										M2 DATA SCIENCE							28/08/2024
September	October	November	December	January	February	March	April	May	June	July	Semester 1		Semester 2				
1/9	1/10	1/11	1/12	1/1	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10	1/11	1/12	1/1	
2/9	2/10	2/11	2/12	2/1	2/2	2/3	2/4	2/5	2/6	2/7	2/8	2/9	2/10	2/11	2/12	2/1	
3/9	3/10	3/11	3/12	3/1	3/2	3/3	3/4	3/5	3/6	3/7	3/8	3/9	3/10	3/11	3/12	3/1	
4/9	4/10	4/11	4/12	4/1	4/2	4/3	4/4	4/5	4/6	4/7	4/8	4/9	4/10	4/11	4/12	4/1	
5/9	5/10	5/11	5/12	5/1	5/2	5/3	5/4	5/5	5/6	5/7	5/8	5/9	5/10	5/11	5/12	5/1	
6/9	6/10	6/11	6/12	6/1	6/2	6/3	6/4	6/5	6/6	6/7	6/8	6/9	6/10	6/11	6/12	6/1	
7/9	7/10	7/11	7/12	7/1	7/2	7/3	7/4	7/5	7/6	7/7	7/8	7/9	7/10	7/11	7/12	7/1	
8/9	8/10	8/11	8/12	8/1	8/2	8/3	8/4	8/5	8/6	8/7	8/8	8/9	8/10	8/11	8/12	8/1	
9/9	9/10	9/11	9/12	9/1	9/2	9/3	9/4	9/5	9/6	9/7	9/8	9/9	9/10	9/11	9/12	9/1	
10/9	10/10	10/11	10/12	10/1	10/2	10/3	10/4	10/5	10/6	10/7	10/8	10/9	10/10	10/11	10/12	10/1	
11/9	11/10	11/11	11/12	11/1	11/2	11/3	11/4	11/5	11/6	11/7	11/8	11/9	11/10	11/11	11/12	11/1	
12/9	12/10	12/11	12/12	12/1	12/2	12/3	12/4	12/5	12/6	12/7	12/8	12/9	12/10	12/11	12/12	12/1	
13/9	13/10	13/11	13/12	13/1	13/2	13/3	13/4	13/5	13/6	13/7	13/8	13/9	13/10	13/11	13/12	13/1	
14/9	14/10	14/11	14/12	14/1	14/2	14/3	14/4	14/5	14/6	14/7	14/8	14/9	14/10	14/11	14/12	14/1	
15/9	15/10	15/11	15/12	15/1	15/2	15/3	15/4	15/5	15/6	15/7	15/8	15/9	15/10	15/11	15/12	15/1	
16/9	16/10	16/11	16/12	16/1	16/2	16/3	16/4	16/5	16/6	16/7	16/8	16/9	16/10	16/11	16/12	16/1	
17/9	17/10	17/11	17/12	17/1	17/2	17/3	17/4	17/5	17/6	17/7	17/8	17/9	17/10	17/11	17/12	17/1	
18/9	18/10	18/11	18/12	18/1	18/2	18/3	18/4	18/5	18/6	18/7	18/8	18/9	18/10	18/11	18/12	18/1	
19/9	19/10	19/11	19/12	19/1	19/2	19/3	19/4	19/5	19/6	19/7	19/8	19/9	19/10	19/11	19/12	19/1	
20/9	20/10	20/11	20/12	20/1	20/2	20/3	20/4	20/5	20/6	20/7	20/8	20/9	20/10	20/11	20/12	20/1	
21/9	21/10	21/11	21/12	21/1	21/2	21/3	21/4	21/5	21/6	21/7	21/8	21/9	21/10	21/11	21/12	21/1	
22/9	22/10	22/11	22/12	22/1	22/2	22/3	22/4	22/5	22/6	22/7	22/8	22/9	22/10	22/11	22/12	22/1	
23/9	23/10	23/11	23/12	23/1	23/2	23/3	23/4	23/5	23/6	23/7	23/8	23/9	23/10	23/11	23/12	23/1	
24/9	24/10	24/11	24/12	24/1	24/2	24/3	24/4	24/5	24/6	24/7	24/8	24/9	24/10	24/11	24/12	24/1	
25/9	25/10	25/11	25/12	25/1	25/2	25/3	25/4	25/5	25/6	25/7	25/8	25/9	25/10	25/11	25/12	25/1	
26/9	26/10	26/11	26/12	26/1	26/2	26/3	26/4	26/5	26/6	26/7	26/8	26/9	26/10	26/11	26/12	26/1	
27/9	27/10	27/11	27/12	27/1	27/2	27/3	27/4	27/5	27/6	27/7	27/8	27/9	27/10	27/11	27/12	27/1	
28/9	28/10	28/11	28/12	28/1	28/2	28/3	28/4	28/5	28/6	28/7	28/8	28/9	28/10	28/11	28/12	28/1	
29/9	29/10	29/11	29/12	29/1	29/2	29/3	29/4	29/5	29/6	29/7	29/8	29/9	29/10	29/11	29/12	29/1	
30/9	30/10	30/11	30/12	30/1	30/2	30/3	30/4	30/5	30/6	30/7	30/8	30/9	30/10	30/11	30/12	30/1	
1/10	1/11	1/12	1/1	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10	1/11	1/12	1/1	1/2	

Calendar: <https://tinyurl.com/agenda-m2ds>

The year in 4 parts (parts 1-2 \approx semester 1, parts 3-4 \approx semester 2))

- **Part 1:** Courses 11/09-20/10, Exams 23-27/10
- **Part 2:** Courses 06/11-15/12, Exams 08-12/01, Data Camp 18-22/12
- **Part 3:** Courses 15/01-22/03, Exams 25-29/03
- **Part 4:** Internship from 01/04, Defense in August/September/October

Vacations

- Fall : 28/10-05/11
- Winter : 23/12-07/01

M2DS Courses

Syllabus: <https://tinyurl.com/syllabus-m2ds>

Part P1 & P2

Course name	Professors	Part	ECTS
Introduction to Operation Research	Eric SOUTIL	P1	3
Natural Language Processing and Sentiment Analysis	Chloé CLAVEL	P1	3
Deep Learning I	Geoffroy PEETERS	P1	3
An Introduction to Machine Learning Theory	Stephan CLEMENCON - Hicham JANATI	P1	3
Practical introduction to machine learning	Rémi FLAMARY - Ekhine IRUROZKI	P1	3
Big Data Framework	Duc PHAM-HI	P1-2	6
Statistical Learning Theory	Jaouad MOURTADA	P1-2	3
High-dimensional statistics	Alexandre TSYBAKOV	P1-2	3
Hidden Markov models and Sequential Monte Carlo methods	Nicolas CHOPIN	P1-2	3
Nonparametric estimation and testing	Cristina BUTUCEA	P1-2	3
Masterclasses and Hilckathon	Eric MOULINES et Emmanuel GOBET	P1-2	3
Optimization for Data science	Alexandre GRAMFORT - Pierre Ablin	P1-2	6
Convex Analysis and Optimization Theory	Pascal BIANCHI - Olivier FERCOQ - Walid HACHEM	P1-2	6
Advanced AI for text and graphs	Michalis VAZIRGIANNIS	P1-2	6
Monte Carlo Methods: from MCMC to Data-based Generative model	Randal DOUC - Emmanuel GOBET - Alain DURMUS	P1-2	6
Partially observed Markov chains in signal and image	Wojciech PIECZYNSKI	P2	3
Computer Vision	Alasdair NEWSON	P2	3
Data camp (mandatory course)	Alexandre GRAMFORT - Thomas MOREAU	P2	3
An Introduction to Reinforcement learning	Erwan LE PENNEC	P2	3
High dimensional matrix estimation	Karim LOUNICI	P2	3
Law and ethics of artificial intelligence	Winston MAXWELL	P2	3

M2DS Courses

Syllabus: <https://tinyurl.com/syllabus-m2ds>

Part P2 & P3

Course name	Professors	Part	ECTS
Generalisation properties of algorithms in ML	Aymeric DIEULEVEUT	P2-3	6
Online learning and aggregation	Alexandre TSYBAKOV	P3	3
Optimal Transport: Theory, Computations, Statistics, and ML Applications	Marco CUTURI	P3	3
Cooperative Optimization for Data Science	Andrea SIMONETTO	P3	3
Operation research for Data Science	Zacharie ALES	P3	3
Big Data and Insurance Project	Denis OBLIN	P3	3
Cloud data infrastructure	Nicolas TRAVERS	P3	3
Audio and music information retrieval	Geoffroy PEETERS	P3	6
Tail events analysis: Robustness, outliers and models for extreme	Pavlo MOZHAROVSKY	P3	3
Stochastic approximation and reinforcement learning	Pascal BIANCHI - Walid HACHEM	P3	3
Deep Learning II	Yohan PETETIN/ Newson	P3	3
Representation Learning for Computer Vision and Médical Imaging	Pietro GORI - Loïc LE FOLGOC	P3	3
Recent Developments in Responsible AI	Charlotte LACLAU - Florence D'ALCHE-BUC	P3	3
DATA stream processing	Maurras TOGBE - Jérémie SUBLIME - Mariam BARRY	P3	3
Structured Data: Learning and Prediction	Florence D'ALCHÉ-BUC	P3	3
Missing Data and causality	Mariane CLAUSEL	P3	3
ML Research Seminar	Éric Moulines - Rémi Flamary	P3	6
Capstone Project	Marylou GABRIE - Anna KORBA	P3	6

Courses registration

Registration link

<https://m2ds.flamary.com/>

Limit date for submission: **05/09/2022 23:59**

Requirements

- Select at least 42 ECTS, no more than 48 ECTS.
- No conflicts possible (only one course per 1/2 day).
- For course outside of the master: pick one from M2DS (without numerus clausus) and contact us afterward.
- Use your unique email (you can update your selection with it).

Suggestions

- Spread the courses between P1/P2/P3 (P3 has ML seminar or capstone).
- Limit the number of courses with numerus clausus.
- Be curious, have fun, you have access to materials even if not selected.
- Some suggested "tracks" in the next slides.

Contact : stephanie.clevenot@polytechnique.edu

Track : Research in theoretical ML

Example track

P1-P2

- Deep Learning I
- An Introduction to Machine Learning Theory
- Monte Carlo Methods: from MCMC to Data-based Generative model
- Nonparametric estimation and testing
- High-dimensional statistics
- Generalisation properties of algorithms in ML

P3-P4

- Structured Data: Learning and Prediction
- Online learning and aggregation
- Missing Data and causality
- Optimal Transport: Theory, Computations, Statistics, and ML Applications
- ML Research Seminar
- Internship in a research lab.

Track : Optimization for ML and data science

Example track

P1-P2

- Deep Learning I
- Convex Analysis and Optimization Theory
- Optimization for Data science/Optimisation pour les datasciences
- An Introduction to Reinforcement learning
- Introduction to Operation Research
- Data Camp

P3-P4

- Operation research and Big data
- Cooperative Optimization for Data Science
- Missing Data and causality
- Optimal Transport: Theory, Computations, Statistics, and ML Applications
- ML Research Seminar or CAPSTONE Project

Track : ML and applications

Example track

P1-P2

- Deep Learning I
- Practical introduction to machine learning
- Natural Language Processing and Sentiment Analysis
- Partially observed Markov chains in signal and image
- Data Camp

P3-P4

- Advanced AI for text and graphs
- DATA stream processing
- Recent Developments in Responsible AI
- Audio and music information retrieval
- Missing data and causality
- Capstone Project

Track : Data science for industry

Example track

P1-P2

- Practical introduction to machine learning
- Deep Learning I
- Natural Language Processing and Sentiment Analysis
- Big Data Framework
- Law and ethics of artificial intelligence
- Data Camp

P3-P4

- Operation research and Big data
- DATA stream processing
- Missing Data and causality
- Deep learning II
- Capstone Project

Internship

Schedule and contact

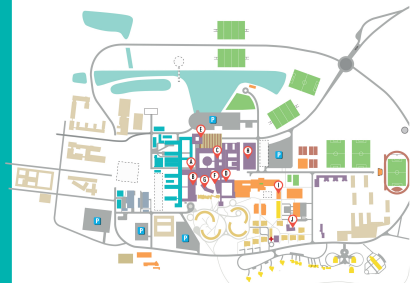
- Internship can start from April 1 2023, minimum 16 weeks.
- Start looking for internships early (December at least).
- Subject: need to have contributions in the field of data science and must be validated by an "enseignant référent".
- Contact : ny-kanto.andriahenintsoa@polytechnique.edu

Internship subject validation (enseignants référents)

- **X/Polytechnique** Rémi Flamary, Éric Moulines
- **ENSAE** Anna Korba
- **ENSTA** Zacharie Ales
- **Télécom Paris** Hicham Janati, Olivier Fercoq
- **Télécom SudParis** Randal Douc
- **ISEP** Jérémie Sublime
- **ECE** Duc Pham Hi

IP Paris Campus

Campus de l'École polytechnique - IP Paris



- A Bâtiment S7 - Informatique
- B Bâtiment 8 - Centre Poly-média - Langues - HSS - PC
- C Bâtiment 13 - Ensemble Central (DRH)
- D Bâtiment 21 - Pôle Social - Moyens Généraux - Aumôneries
- E Bâtiment 28 - Boncourt - Présidence - Direction générale

- F Bâtiment 69 - Bachelor - Executive Master
- G Bâtiment 81 - Finances - Achats
- H Bâtiment 86 - Dashi - X Navigation Center
- I Bâtiment 409 - Bienvenue (CASEP - CSR)
- J Bâtiment 411 - EXED

Accueil

- 2 Accueil Evénement
- 3 Amphithéâtres
- 4 Bibliothèque - Mus'X
- 5 Centre aquatique et sportif
- 6 Cour des Conférences
- 7 Cour Vaneau
- 8 Direction du Concours
- 9 Présidence & Direction générale
- 10 Direction Graduate Degree - 1^{er} étage
- 11 Direction Bachelor - RDC
- 12 Direction Master / Doctorat - SOE
- 13 Escalier Education
- 14 Bureaux logements

- Crous truck
- Arrêt Bus 91 06



- 14 Grand hall
- 15 Laboratoires / Départements
- 16 Dashi - X Navigation Center
- 17 LIX, Inria, MSR - Bât. Alain Turing
- 18 Laboratoire d'Optique Appliquée
- 19 Centre de secours
- 20 Restauration - Conciergerie
- 21 Magasin de produits chimiques



Other information

Access to campus

- From Paris : RER B (to Massy Palaiseau or to Lozere) or RER C (to Massy Palaiseau)
- From Massy Palaiseau: TransEssonne 91.06 ou 91.10
<http://www.albatrans.net/les-lignes-les-horaires/>
<https://me-deplacer.iledefrance-mobilites.fr/>
- IP Paris Campus smartphone App.

More information : <https://www.ip-paris.fr/acces-et-mobilite>

Restaurants and student life

- Several restaurants on the Campus
- Restaurant Magnan (polytechnique) requires to activate the Student ID card.

More information :

<https://www.ip-paris.fr/campus/vie-etudiante/vie-pratique>