



Hi!ckathon #5

Student guide



Agenda



1. About the Hi!ckathon
2. Understanding the subject
3. Work methods
4. Deliverables and deadlines
5. Coaching

What is the Hi!ckathon ?

The event

Competitive Data / AI event of Hi! PARIS, the interdisciplinary research center for Artificial Intelligence founded by Institut Polytechnique de Paris and HEC Paris.

The theme

AI & Sustainability

Beyond producing an AI model, the competition will ask you to realistically project your solution in a market / real-world context.

When?

From Friday **November 29th, 6:30pm** to Monday **December 2nd, 9pm.**

- Friday to Sunday Evening at **Télécom Paris**
- Monday Evening (Closing Ceremony) at **ENSAE Paris.**

Timeline of the event



All activities will be held on site

Friday 29th

Télécom Paris

Opening Ceremony

- Welcome
- Corporate donors round table
- Release of the topic
- Instructions and team number release

6:30pm



Topic release

8pm

Saturday 30th

Télécom Paris

Start of the technical advice

9am

10 am

>>1 Business Coaching per group<<

Activities

Sports & Games (chess, VR, Switch)

12pm

Movie Time

9pm

Activities

Sports & Games (chess, VR, Switch)

10 am

12 pm

Sunday 1st

Télécom Paris



Return of the deliverables
(4 PM)

Code assessment
(4PM – 5PM)

4pm

End of the Hackathon

6pm

Monday 2nd

ENSAE

Closing Ceremony

- Live pitch + Jury session 2
 - Award ceremony
 - Wrap-up
 - Networking cocktail
- 6:30pm

9am

Jury session 1

Meals over the weekend: 2 breakfasts and 1 afternoon tea

What are the deliverables ?



Build an AI model and make predictions

Develop a state-of-the-art AI model that addresses the sustainability use case



Create a scientific / business document

Justify the scientific and business approach of your AI solution



Produce a video presentation

Film a short video to present your project and defend its merit








Defend your project in front of a live audience*

Some groups will be asked to present their project in front of a live audience during the closing ceremony.

The groups who will have to present live during the closing ceremony will be notified Sunday evening. More information on this can be found in **4. Deliverables and deadlines.*

Hi!ckathon awards

Award	Description	Given by
Business opportunity	Rewards the team that best projected itself in a market / real-world context through an impactful pitch .	
Scientific approach	Rewards the team which best demonstrates the ability to justify its scientific & business approach and to confront the state of the art in data science	
Technical excellence*	Rewards the model with the highest score . The score is a synthetic measure of the model's accuracy and code quality.	
Interdisciplinarity	Rewards the teams able to combine several disciplines in the production of the scientific & business approach and presentation .	
Public crush	Rewards the public's favorite team during the live pitch session.	

*Winners of the **Technical excellence award** will be asked to present their project during the event's closing ceremony. The winning group will be contacted by a member of the organizational team right after the 4PM Sunday submission deadline.

Guidelines for participants



Going to unauthorized floors/areas is prohibited

- Only the ground-floor and first floor (level 0 and 1) are accessible during the competition.
- Areas bared by masking tape cannot be reached by participants.



Please keep your room/designated area clean

- Participants must keep their room/designated area clean during the duration of the competition.
- Teams who don't comply with this rule by the end of the weekend will be penalized or disqualified (the same rule applies for accessing restricted areas).



Need to sleep or take a shower ?

- Sleeping rooms are available to rest during the night (only between 11PM-7:45AM).
- Showers are available at Télécom Paris during the weekend.



Access to the venue during the weekend

- The main entrance of Télécom Paris will be closed for the whole duration of the final sprint, starting Friday at 8PM.
- To go in or out, you have to pass by the security desk.



Any trouble during the competition ?

- Hi!ckathon #5 team members are available anytime during the weekend, from 8AM to 10PM.
- Don't hesitate to reach out to any of us regarding any issues/concerns.

Guidelines for participants

Where can I eat near the venue ?

- **PANIFIQUE** (TAKEAWAY): 13 Av. de la Vauve, 91120 Palaiseau
Opening hours: from 6:30 am to 8 p.m / Monday to Friday
- **FRANPRIX** (In-store shopping): 13 Av. de la Vauve, 91120 Palaiseau
Opening hours: from 9 am to 9 pm/ Monday to Sunday
- **Living Room Palaiseau** (Dine-in / Takeaway): 10 Boulevard Gaspard Monge, 91120 Palaiseau
Opening hours: from 12 pm to 2:30 pm and from 7 pm to 10:30 pm / Monday to Saturday
- **Le Julie's** (Dine-in/ Takeaway): 40 Cr. Pierre Vasseur, 91120 Palaiseau
Opening hours: Monday to Friday (7 am to 11:30 pm) , on Saturday (5 pm to 11:30 pm)
- **FOOD ASIA** (Dine-in/ Takeaway): 2 Boulevard Gaspard Monge, 91120 Palaiseau
Opening hours: from 11:30 am to 9:30 pm/ Monday to Saturday
- **All Together** (Dine-in): 27, boulevard Thomas Gobert - 91120 Palaiseau
Opening hours: from 11:30 am to 2:30 pm then from 6 pm to 10:30 pm/ Monday to Friday

Most meals won't be provided during the weekend, except for breakfast at 8AM Saturday and Sunday morning, a snack Saturday afternoon as well as a final cocktail Monday evening.

Agenda



1. About the Hi!ckathon
2. Understanding the subject
3. Work methods
4. Deliverables and deadlines
5. Coaching

AI & Sustainability – Water shortage prediction

Dataset

A data set of +3 million lines, more than 100 columns, about weather, hydrology, abstraction, socio-economic data, piezometry, etc.

Story Telling

- A) You are a consultant whose aim is to support local players in their water abstraction strategy.
- B) You are an association whose aim is to alert local authorities to the use of water in a given area.
- C) You are a start-up whose aim is to develop technological solutions to improve sustainable water management.
- D) Open problem, choose your situation.

Objectives and Deliverables

ML task

Predict groundwater levels on Summer

Scientific & Business Approach

You need to justify your predictions, demonstrating how your approach is innovative, viable and socially profitable.

Presentation “Pitch” goal

In 3 minutes, you must explain your project, demonstrating its viability, showing how your proposal is innovative and original, and what the added value to society is.

What we want

Sustainability Imperative
Holistic Approach
Collaboration for Impact
State-of-the-art Machine Learning

Agenda



1. About the Hi!ckathon
2. Understanding the subject
3. Work methods
4. Deliverables and deadlines
5. Coaching

What resources do I have ?

For the final sprint



- A welcome document
- A student guide
- **Datasets** to train/evaluate your AI solution
- A variable dictionary
- An **HFactory** guide
- A guide notebook to load the datasets in HFactory

Where can I find these resources ?

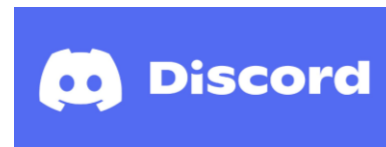
- **Final Sprint**: Go to the « *Hi!ckathon #5 Final Sprint* » activity on HFactory + **Google Drive**
- **Preparation phase**: Go to the « *Hi!ckathon #5 Journey* » activity + **Google Drive** + **Github repo**

From the preparation phase



- Three intermediate-level **Data Science courses** (notebooks + video recordings)
- A **business pitch course** (slides + video recordings)
- Introductory Data Science and Python courses

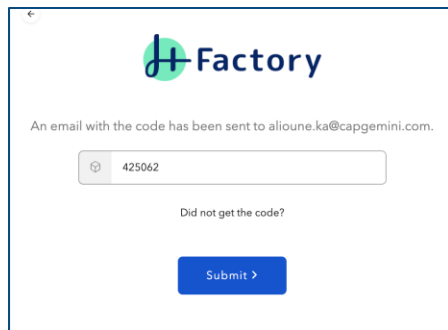
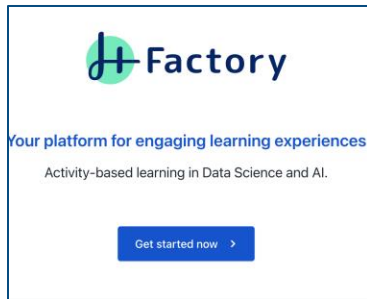
What tools should I use ?



- Find the **challenge topic** and **datasets**
 - Submit **deliverables**
 - Work on the project with your group
 - Access **CPU/GPU instances** and popular IDEs (JupyterLab, VSCode, ...)
 - You received during the opening ceremony an email to register to HFactory
- Get the latest **news** and **announcements** of the Hi!ckathon
 - **Contact the Hi!ckathon organizational team** for assistance
 - Discuss with fellow team members in private rooms
 - Access the discord here:
<https://discord.gg/DpF2Yg3A>

Step-by-step guide on the Hfactory platform

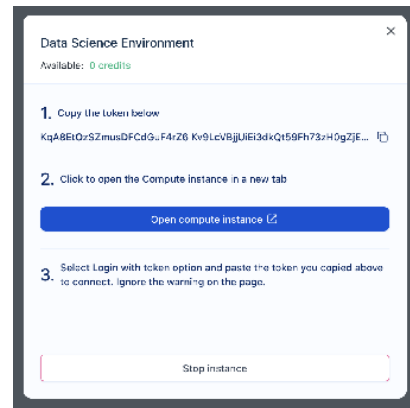
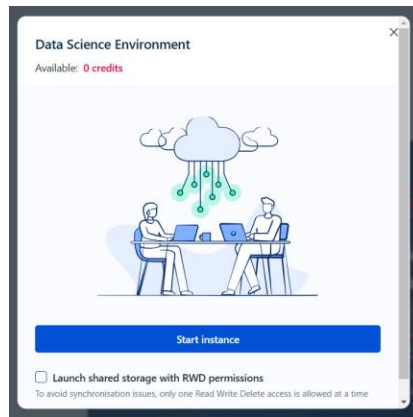
- 1 Register to the Hicckathon #5 activity on HFactory (you should have received an invitation by email)



- 2 Click on the **compute** icon in the top menu bar to access the Data Science computing environment

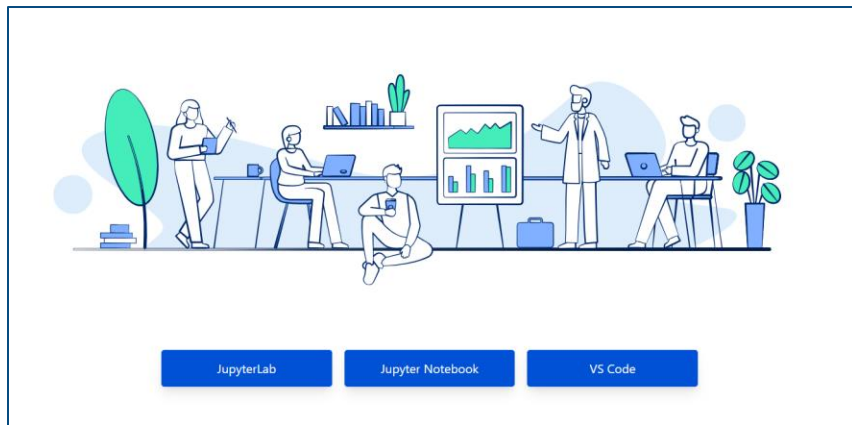


- 3 Start a CPU/GPU computing instance. Then, copy the token then click on "Open compute instance" to login with it.

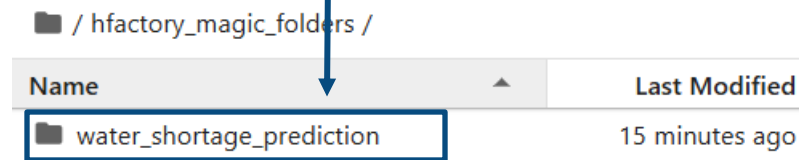
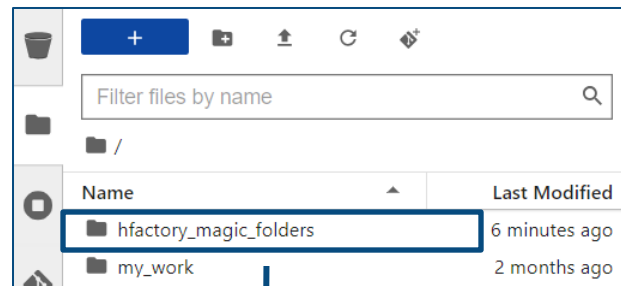


Step-by-step guide on the Hfactory platform

- 4 Now, you can select an IDE to use with the computing environment (JupyterLab, Jupyter Notebook, VS Code)



- 5 Once you've selected an IDE, you can **open the computing environment**. In this environment, you can access the dataset and work in your own or group environment.



Need help with HFactory ?

Guidelines about HFactory

- HFactory gives you access to **CPU/GPU instances** with JupyterLab, Jupyter notebook and VSCode
- Group workspaces/directories are available to work collaboratively during the sprint
- At the beginning of the competition, CPU instances will only be available. 30 GPU credits per group will be attributed Saturday evening.
- All of the **Hi!ckathon #5 Datasets** are in the “**Studio**” portion of the “Hi!ckathon #5 – Final Sprint” activity (they can also be downloaded from a course page).
- **Submissions** must also be made in the same Hi!ckathon #5 activity.

A detailed guide called ***HFactory_guide.pdf*** is available in the HFactory activity if you need help using the platform. You can also find the pdf [here](#).

Agenda



1. About the Hi!ckathon
2. Understanding the subject
3. Work methods
4. Deliverables and deadlines
5. Coaching

Deliverable 1: Data Science model & prediction



What should I submit ?

- Submit the **code** of your Data Science model (in .ipynb or .py format). *Your code will be evaluated by the Hi!ckathon team for code quality.*
- Submit your **predictions** on the test set to the leaderboard (in .csv format). *Note that you can only make up to ten submissions to the leaderboard.*

How is code quality evaluated ?

- **Executability:** You will be asked to run your code live in front of Data Science coaches during code assessment. To speed-up the assessment, we will only ask you to run the code on 10 000 rows. Groups that don't show up or can't run their code live will be disqualified.
- **Code structure:** Good use of functions/classes to structure your code. Adding comments, docstrings and/or markdown cells to document your code.

Deliverable 1: Data Science model & prediction



What metric is used for evaluation ?

A weighted F1 score is used to score submissions to the leaderboard.

- **F1 score:** Harmonic mean of the precision and recall metrics
- **Precision:** Proportion of true positives among all positive predictions
- **Recall:** proportion of true positives among all actual positives

		Predicted	
		0	1
Actual	0	TN	FP
	1	FN	TP

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The weighted F1 score is the weighted average of the F1 scores calculated for each class. The weights correspond to the support, which is the number of true instances for each class.

Click [here](#) to learn more about the weighted f1 score (see weighted parameter)

Deliverable 2: Scientific / Business approach



Requirements / Target

- Submit a **short** PDF document detailing the **scientific** and **business** approach you developed for your project/solution.
- Make your document **clear and concise**. It shouldn't exceed a 4 minutes reading time (maximum 4 pages). Adding more pages will lead to the **disqualification** of your group.
- Don't hesitate to add **visualizations** to make your approach engaging and easily understandable to the jury.
- The following slide showcases a template of the **mandatory sections** you need to include in your document.

You can include additional subsections or questions beyond those provided in the template. However, ensure that your document remains concise.

Deliverable 2: Scientific / Business approach



Mandatory sections

I. Overview

- Problem statement (What is the issue ? Why is it critical ?)
- Presentation of your project / solution
- Presentation of your team

II. Business approach

- What is the value of the solution ?
- How does it integrate into real-world operations?
- How is it different from other solutions ? What makes it novel ?

III. Scientific approach

- What is the technical foundation of your solution ? (pre-processing, model)
- How did you make your approach environmentally frugal ? (can also be answered in the business approach)
- How does it scale ?

IV. Results and Future Potential

- What results have you achieved ? (score, explainability)
- What is the future potential of your solution ?

Deliverable 3: Presentation video



Requirements / Target

- Create a **short, facecam video** presenting your solution.
- The video must **not exceed 3 minutes**. Videos longer than 3 minutes will result in **disqualification**.
- **No editing or cuts** are allowed, but you can use props if you want to.
- This video will be used to select the qualified groups for the final live pitch session held during the closing ceremony.

Guidelines for the pitch

- Use the **business pitch course** given during the pre-hi!ckathon training phase as guidelines/help
- You can find the slides of the course **here**, and its recording **here**
- You can also ask business coaches advice to build your presentation (see 5.Coaching section).

Deliverable 3: Presentation video



Here are some excerpts of the business pitch course :

Sharing your Impact-Making Idea

A good presentation is not an option:

- 7% words
- 38% tone of voice
- 55% body language



TU/e
INNOVATION
SPACE

© Copyright Vittorio Pignatelli

Explain your Solution

- Max 14-15 words
- Simple & understandable way (visualise/prototype)
- Provide an explanation but not technical details
- If needed you can use more than one slide

TU/e
INNOVATION
SPACE

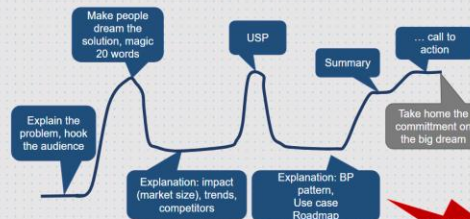
Why You???

- What is special in your solution? Why shall we choose you and not the others?
- What do you do better than your competitors (USP)

TU/e
INNOVATION
SPACE

Resonate

Manage the tone of your presentation ... like a storyteller



Look at http://www.ted.com/talks/Eliezer_Yudkowsky_a_powerful_stroke_of_insight.html
Look also at http://www.ted.com/talks/Eliezer_Yudkowsky_a_powerful_stroke_of_insight.html
Advised reading: «Resonance» by Nancy Duarte

Remember, 2-3 minutes!

Deliverable 4: Defend your project in front of a live audience



Requirements / Target

- Present your project to a live audience/jury during the Hicathon's closing ceremony
- Same as the video, **your speech shouldn't exceed 3 minutes.**
- You can deliver the same presentation as the video or modify it if you want to.
- **This deliverable only concerns 7-10 groups** who will be notified Sunday evening of their participation.
- The live presentation is a **component of the jury's evaluation to win awards.** *However, not being selected to perform the live pitch does not disqualify you from winning an award.*
- Please prepare a **single slide** to present your project for the live presentation.

Presence at the closing ceremony is mandatory for every participant, even for those who weren't selected for the live presentation.

Deadlines and disqualification

Important deadlines



- The **code**, **scientific/business approach** and **presentation video** should be submitted before **Sunday at 4PM** via the *submission form* on HFactory
- Don't forget to attend the **code assessment** from **4PM to 5PM** on Sunday. You will be asked to run your code live (data preprocessing + inference)

Disqualification



- Groups must **submit all required deliverables** and adhere to **all specified requirements** to avoid disqualification. This includes mandatory participation and/or attendance in both the **code assessment** and **live presentation** sessions.
- Any form of **plagiarism** will also to disqualification (for all parties involved). *For students whose participation is mandatory, your pedagogic director will also be notified of any plagiarism.*

Where do I make submissions ?

All of the submissions should be made on the “Hickathon #5 – Final Sprint” **HFactory** activity.

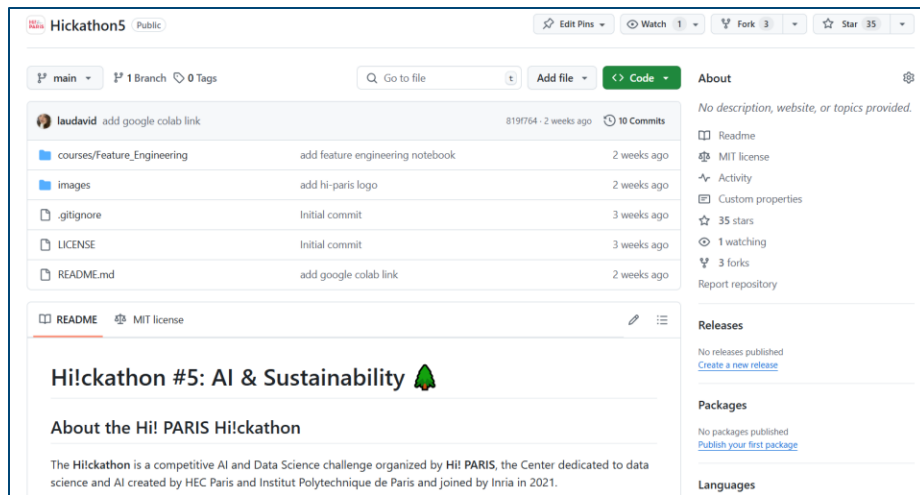
- Submit your **test predictions** to the **Leaderboard** section
- Submit all remaining files (**python code, scientific/business paper, video pitch**) to the **Submissions** section

The screenshot shows the HFactory submission interface. On the left, a sidebar contains three buttons: 'Predictions', 'Leaderboard', and 'Submissions'. Two blue boxes with arrows point to these buttons: one labeled 'Predictions' pointing to the 'Leaderboard' button, and another labeled 'Code, paper, video' pointing to the 'Submissions' button. The main content area is titled 'Submissions' and includes a 'Submission' section with a deadline of 'Sun. Dec 1st, 2024 at 4:00pm' and a time remaining of '2 day(s) / 4 hour(s) / 19 min(s)'. Below this, a dashed box contains the instructions: 'What should you submit?' followed by a list of three items: 'The Python code (in .ipynb or .py format) you used to train your model and make predictions to the leaderboard', 'The Scientific/Business Paper in PDF format (maximum 4 pages)', and 'The video business pitch (maximum 3 minutes)'. It also states 'Your model predictions should be submitted to the leaderboard page, not this one.' and provides 'Guidelines for submissions' stating that only one of each deliverable per group should be submitted, and that group numbers should be included in file names. At the bottom, there is a section for adding submission files with a text input field and a button to click to upload.

Share your project on github

The Hi! PARIS center adopts an **open-source distribution logic**, according to which your AI solution can be shared to the Hi!ckathon #5's public github repository.

If you want to share your work on a personal github, we ask that you mention in your README.md the Hi!ckathon #5 github repository.



<https://github.com/hi-paris/Hickathon5>

Agenda



1. About the Hi!ckathon
2. Understanding the subject
3. Work methods
4. Deliverables and deadlines
5. Coaching

Coaching during the weekend

Business coaching



- Q&A sessions with business coaches from Hi! PARIS's corporate donors will be organized **Saturday** and **Sunday**.
- Each team will be assigned to a Q&A session at a **specific time slot**. Not all team members need to join the session.
- Sessions aren't specific to a single team. Multiple teams (but not all) can attend the same session

Technical coaching



- The organizational team is made up of **Hi! PARIS Research Engineers**, as well as **Data Science Consultants**. You can ask us any technical question.
- We are available anytime during the weekend (8AM-10PM), except during sleeping hours.
- We can provide **guidance in understanding the dataset** or in **following a Data Science methodology**. We cannot not intervene directly on your code or give ready-made solutions.

Thank you for following this guide ! 🧡

If you have any questions, don't hesitate to reach out to a member of the Hilckathon #5 team during the weekend.

And more importantly, **good luck** 😊

Hi!

HEC
PARIS



INSTITUT
POLYTECHNIQUE
DE PARIS

PARiS

PARIS ARTIFICIAL INTELLIGENCE FOR SOCIETY