

Research internship AI / machine learning
Performance improvement of a “deep search engine”
November 2022

You are completing a master in artificial intelligence and machine learning, and you are looking for a research internship (a few months long).

*You would like to work on **real data** in **directly operational conditions**, for a small **fast developing start-up company**?*

*You are **autonomous, creative, and in the same time very consistent**?*

*You want to help us **revolutionize access to culture for all** by creating **the first « deep search engine »**, i.e. a search engine that indexes objects according to their meaning, and not only according to their content - challenges that even Google gave up?*

We have many beautiful topics to work on ;-)



About ideactiv

[ideactiv](#) is **the first search engine for cultural activities** : it enables everyone to instantly find, from either a web or a mobile UI, concerts, visits, exhibitions, conferences, etc, matching their interests, anywhere in France, and soon in other countries.

Its mission is to **ease access to culture everywhere, to everyone**, on a European platform, dedicated to culture and tourism, freely accessible to all cultural and touristic actors.

Thanks to its unique AI capabilities, it is **the only app constantly offering 150 000 activities in France**, from thousands of organizers and announcers.

The mobile app¹ was launched 18 months ago and has **more than 10 000 returning users**. The website is only 6 months old and has approximately the same amount of users.

The content of the internship

The internship will create operational capabilities on part or all of the following topics:

1. **detection of « online » activities**, based on the title, description, address of the activities. This is not a trivial task, there are lots of traps. You will develop one or many solutions, test them and put them in production.
2. **automatic interpretation of dates**. The goal is to be able to « understand » (in a structured format) dates formulated for humans. This is a great challenge, given the extreme variability of the dates in natural language. It is also a topic of universal interest (far beyond culture). You will search what solutions already exist, and develop your own ML-based solution.
3. **consolidation of similar addresses**. The goal is to determine whether similar addresses are linked to the same actual place, and then to create, from the various addresses, the best consolidated address that will be shown to the user. This is a difficult topic. You will search what solutions already exist, and develop your own ML based solution.

A few examples of these topics are described in the attached powerpoint presentation.

It is also possible to work on the app and website usage data (analysis of the usage logs in order to determine the most used functions, the usage patterns, the sessions durations, and generally speaking the users behavior and interests).

On all these topics, ideactiv will provide data in sufficient quantities to train and test your algos. You will also be able to compare the performance of your solutions with those of the current ideactiv algos.

The internship will be mostly remote based, with regular online and in person discussions with ideactiv founder.

The internship can be followed by a thesis and/or a job offering in order to go deeper on these topics (and others) which are the heart of ideactiv activity.

Interested? Let's talk about it! hello@ideactiv.com

¹ Free app "ideactiv" on AppStore and GooglePlay. Website www.ideactiv.com