

✉ Donn  que par email

✉ loick.bourdois@outlook.com

in [Loïck Bourdois](#)

globe lbourdois.github.io

EXPERIENCE PROFESSIONNELLE ET DE RECHERCHE

CATIE, Data Scientist

Jan. 2021 – pr sent

Membre de l' quipe Algorithmes et Donn es

Bordeaux, France

Missions :

Accompagnement des entreprises dans leur d veloppement technologique sur des sujets de science des donn es.

En pratique, participation   une vingtaine de projets diff rents pouvant  tre class s de trois fa ons diff rentes :

- Etudes de faisabilit /de pr -industrialisation et encadrements.
 - Clustering de donn es non annot es
 - Exp rimentation de m thodes de l' tat de l'art
 - Encadrement d' tudiants et acculturation en IA
- Aide   la conception et au prototypage de solutions d'apprentissage profond en image/texte, d'apprentissage machine ou de s ries temporelles.
 - Contr le qualit  (ex : d tection de d fauts dans des images)
 - Optimisation de process (ex : recherche de documents similaires, analyse satisfaction client, etc.)
 - Maintenance en temps r el (ex : d tection d'anomalies en embarqu )
- Recherche & D veloppement du CATIE notamment en NLP.
 - Entra nement   partir de z ro de mani re efficace d'un mod le encodeur-d codeur en fran ais de 147M de param tres (le Flash Attention T5) : <https://github.com/catie-aq/flashT5>
 - Am lioration de l' tat de l'art en fran ais en Question Answering : <https://hf.co/CATIE-AQ/QAmembert> et en Reconnaissance d'Entit s Nomm es : <https://hf.co/CATIE-AQ/NERmembert-base-3entities>
 - Jeux de donn es de prompt en fran ais : <https://hf.co/datasets/CATIE-AQ/DFP>

INSERM U1219 « Bordeaux population health », Data Scientist

Nov. 2019 – D c. 2020

Membre de l' quipe IETO (Injury Epidemiology Transport Occupation)

Bordeaux, France

Travail effectu  :

- Entra nement sous PyTorch de *transformers* (BERT/GPT2) pour classifier automatiquement des anamn ses du service des urgences du CHU de Bordeaux. Publication : <https://ai.jmir.org/2023/1/e40843/>
- D sidentification des anamn ses du service des urgences du CHU de Bordeaux (mix de techniques de type *transformers*, Bi-LSTM et de r gles). Publication : <https://journals.flvc.org/FLAIRS/article/view/128480>
- Communications dans des congr s (Dataquitaine 2020 , PFIA 2021)

EDUCATION

Universit  de Bordeaux, Master MAS parcours MSS mention bien

2017 – 2019

Master Math matiques appliqu es et statistique, parcours Mod lisation Statistique et Stochastique

Bordeaux, France

Principaux cours :

- *Machine Learning* et *Deep Learning*, *Data mining*, S ries temporelles, Analyse de donn es en grande dimension
- Algorithmes stochastiques, Optimisation convexe, Cha nes de Markov, Martingales, Analyse de survie et de fiabilit 
- Ethique, Management, Th orie des organisations

COMPETENCES

Python Machine Learning/Deep Learning : PyTorch, Scikit-learn, WandB | Data Science : Pandas, Numpy
NLP : Hugging Face (transformers, sentence-transformers, datasets) | Demo : Matplotlib, Seaborn, Streamlit

Langues Anglais

DIVERS

Blog Mon blog personnel d di    l'apprentissage profond : <https://lbourdois.github.io/blog/>
Cit  dans le rapport « Les nouveaux d veloppements de l'intelligence artificielle » de l' Office parlementaire d' valuation des choix scientifiques et technologiques
<https://www.senat.fr/notice-rapport/2024/r24-170-notice.html>

Traduction Traduction du cours d'apprentissage profond de *Yann Le Cun* et *Alfredo Canziani* de l'Universit  de New York :
<https://lbourdois.github.io/cours-dl-nyu/>

Programme Membre du programme [Hugging Face Fellows](https://huggingface.co/hugging-fellows) : <https://huggingface.co/hugging-fellows>

Formation TP de *deep learning* (25H/an)   destination d' tudiants de l'ENSAM et de l'IUT de Bordeaux
Cours d'initiation au *machine learning* (8h/an)   destination d' tudiants de l' cole YNOV

Loïck BOURDOIS

📧 Given only by email

✉ loick.bourdois@outlook.com

🌐 [Loïck Bourdois](#)

🌐 lbourdois.github.io

👤 PROFESSIONAL AND RESEARCH EXPERIENCE

CATIE, *Data Scientist NLP*

Member of the Algorithms and Data team

Jan. 2021 – present

Bordeaux, France

Missions:

Support companies in their technological development on data science topics.

In practice, participation in about 20 different projects that can be classified in three different ways:

- Feasibility/pre-industrialization studies and education.

- Clustering of non-annotated data
- Experimentation of state-of-the-art methods
- Student mentoring and AI acculturation

- Assist in the design and prototyping of image/text deep learning, machine learning or time series solutions.

- Quality control (e.g. image defect detection)
- Process optimization (e.g. search engine for similar documents, customer satisfaction analysis, etc.)
- Real-time maintenance (e.g. on-board fault detection)

- Research and development at CATIE, especially in NLP.

- Efficient training from scratch of a 147M parameters encoder-decoder model in French (the Flash Attention T5):

<https://github.com/catie-aq/flashT5>

- Improving the state-of-the-art in French in Question Answering: <https://hf.co/CATIE-AQ/QAmembert> and in Named Entity Recognition: <https://hf.co/CATIE-AQ/NERmembert-base-3entities>

- Datasets of French prompts: <https://hf.co/datasets/CATIE-AQ/DFP>

INSERM U1219 « Bordeaux population health », *Data Scientist*

Member of the IETO (Injury Epidemiology Transport Occupation) team

Nov. 2019 – Dec. 2020

Bordeaux, France

Work performed:

- Train NLP transformers models (BERT/GPT2) in PyTorch to automatically classify emergency records from the emergency department of the Bordeaux University Hospital. Publication: <https://ai.jmir.org/2023/1/e40843/>
- De-identification emergency records in the emergency department of the Bordeaux University Hospital (mix of transformers, Bi-LSTM and rules techniques). Publication: <https://journals.flvc.org/FLAIRS/article/view/128480>
- Communications in French conferences (Dataquaine 2020, PFIA 2021)

🎓 EDUCATION

University of Bordeaux, *Master MAS specialization MSS with mention*

Master Applied Mathematics and Statistics specialization Statistical and Stochastic Modeling

2017 – 2019

Bordeaux, France

Main courses:

- Machine Learning & Deep Learning, Data mining, Time Series, Big Data
- Stochastic algorithms, Convex optimization, Markov chains, Martingales, Survival and reliability analysis
- Ethics, Management, Organization theory

⚙️ SKILLS

Python Machine Learning/Deep Learning: PyTorch, Scikit-learn, WandB | Data Science: Pandas, Numpy
NLP: Hugging Face (transformers, sentence-transformers, datasets) | Demo: Matplotlib, Seaborn, Streamlit

Languages French, English

MISCELLANEOUS

- Blog** Personal blog dedicated to deep learning (in French): <https://lbourdois.github.io/blog/>
Quoted in the report “Les nouveaux développements de l’intelligence artificielle” by the “Office parlementaire d’évaluation des choix scientifiques et technologiques” (a working group of French deputies and senators): <https://www.senat.fr/notice-rapport/2024/r24-170-notice.html> (in French)
- Translation** French translator of the *Deep learning course* by Yann LeCun and Alfredo Canziani from New York University <https://lbourdois.github.io/cours-dl-nyu/> (in French)
- Fellowship Course** Member of the [Hugging Face Fellows](https://huggingface.co/hugging-fellows) program: <https://huggingface.co/hugging-fellows>
Deep learning practical work (25h/year) for students at ENSAM and IUT of Bordeaux
Introductory course in machine learning (8h/year) for YNOV students