

# Taycir Yahmed

R&D Software Engineer in NLP and Machine Learning

Paris, France  
+33 (0) 783890129  
taycir.yahmed@gmail.com

LinkedIn:// [tayciryahmed](#)  
Github:// [tayciryahmed](#)  
Website:// [tayciryahmed.com](#)

## Programming

python (expert), C++ (prior),  
SQL (proficient), java (prior)

## Machine Learning

pytorch, sklearn,  
tensorflow, spark ml

## Big Data

hadoop, spark, kafka,  
elasticsearch.

## Other tools

linux, git, docker, CI.

## Languages

English: Trilingual  
French: Trilingual  
Arabic: Trilingual  
German: Beginner  
Chinese: Beginner  
Spanish: Intermediate

## Education

### Master of research, PARIS-SACLAY UNIVERSITY, 2018

- Deep/Machine learning, image processing, NLP, big data.
- Projects: Fake news detection, recommender system, chatbots.

### Master of Science, TELECOM PARISTECH, 2017

- Data science, optimization, statistics, applied mathematics. GPA 4.
- Projects: Face recognition, sounds identification, credit scoring.

### Master of Engineering, INSTITUT MINES TELECOM - SUPCOM, 2016

- Top of the school. Information and communication technologies.
- Project: Customer satisfaction using NLP and machine learning.

## Experience

### NLP & Machine Learning Eng, BNP Paribas | FRANCE, JAN 18 - NOW

Mission: Multilingual text classification automated pipeline.

- Identified more than .6M clients' missing information to ensure the company's compliance, using fine-tuning of multilingual Language Models (BERT).
- Built an annotation tool with active learning to automate classification tasks.

Mission: R&D to advance speech-to-text in low resource setups.

- Helped call centers controllers identify and prioritize problematic situations, using speech-to-text and text analysis, on low resource languages.

Mission: Data augmentation for neural machine translation.

- Increased the BLEU score using bilingual dictionaries, clause detection and segmentation with **attention weights**, back-translation and synthetic contexts.
- Handled rare words for better translations using BPE, normalization with **named-entity** recognition and external knowledge integration.
- Improved context **disambiguation** using pre-trained word embeddings and linguistic features: part-of-speech tags, dependency parsing labels and lemmas.

Mission: Document matching to build multilingual parallel datasets.

- Matched multilingual internal documents to create a parallel in-domain corpus, using OCR, cross-lingual **bag-of-words** decoding and sentence alignment.
- Trained **specialized** neural machine translation models, using fine tuning, up-sampling and domain tags as additional features, on GPUs.

### Big data science intern, Renault | FRANCE, JUL - SEP 17

Mission: Production-ready predictive maintenance solution for the factories.

- Predicted +87% of future defects, by implementing an end-to-end solution, with batch machine learning, **continuous integration** and log monitoring.
- Reduced manufacturing costs by providing business teams **real-time** visualization dashboards, based on **stream processing** in a big data cluster.

### Analytics/data science project, CB Group | FRANCE, FEB - JUN 17

Mission: Data analytics to support deeper market understanding.

- Delivered insights about the company's **market share** in 732 economic sectors.
- Modeled the geographic density: hypothesis testing and variable importance.
- Developed a **dashboard** to visualize the findings with Google design patterns.

### Data science intern, IBM | TUNISIA, JUN - AUG 16

Mission: Pilot project to test IBM Data Science Experience cloud platform.

### Seeds For The Future Program, Huawei | CHINA, 2016

Technical training at Huawei labs and language courses at Beijing University.

### Mobile & analytics intern, IBM Labs | FRANCE, JUN - AUG 15

Mission: Built an app for emergencies, using speech-to-text and text mining.