

# Taycir Yahmed

NLP Data Scientist / Machine Learning engineer  
Leveraging AI to help machines understand human languages

Paris, France  
(+33) 0783890129  
taycir.yahmed@gmail.com

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

## Skills

### Machine learning

pytorch, sklearn, keras,  
tensorflow, spark mllib/ml

### Programming

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

### Big data

Hadoop, spark, kafka,  
elasticsearch, hive.

### Web

html5, javascript, jquery,  
node.js, flask

### Other tools

linux, jupyter, git, docker,  
jenkins, travis, GPUs

## Certifications

Deep learning (Andrew Ng)  
Mobile developer (IBM)

## Languages

English: Fluent, Toeic : 975  
French: Fluent, Voltaire: 812  
Spanish: Intermediate  
German: Beginner  
Chinese: Beginner  
Arabic: Fluent

## Distinctions

- Entrepreneurship and  
innovation challenge, 2016  
- Humanitarian awards:  
Best communication, 2015  
- IBM mobile, cloud and big  
data hackathon, 2014

## Volunteering

Social media coordinator  
Open source contributor  
TED translator

## Education

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

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

### Engineer's degree, TELECOM PARISTECH, 2018

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

### Engineer's degree, INSTITUT MINES TELECOM - SUPCOM, 2016

- Top of the school. Telecommunication and computer science.
- Project: Customer satisfaction using NLP and machine learning.

## Experience

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

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: Internal document matching for specialized machine translation.

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

Mission: Summarization using a hybrid abstractive / extractive approach.

- Improved the ROUGE score using abstractive and extractive summarization.

### 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.

- Predicted post-collegiate earnings with -15% of error, using **regression**, handling missing values with **imputation** techniques and feature engineering.

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

Mission: Mobile app to determine the adequate profile for an emergency

- Design thinking and development, including **speech-to-text** and text mining.

### Seeds For The Future Program, Huawei | CHINA, SESSION 16

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