

Marc Alphonsus, Senior ML engineer



Nationality:french

Address:France 75014, Paris, Avenue du Général Leclerc 105

Phone number:+33 06 62 6 2 62 48

Email address: alphonsus.marc@gmail.com

Web: <https://github.com/marcalph>

Profile

Experienced Senior Machine Learning Engineer with over 9 years of industry experience seeking a senior/lead position in a dynamic, multicultural environment focused on product development. Known for my passion for data and strong work ethics, I excel in creating innovative and implementing complex machine learning models and ensuring successful deployment. Passionate about technology, coffee connoisseur, and avid comic book enthusiast.

Work experience

09/2022 – present
france, UK

AI research engineer Helsing

Contributed to AI research efforts from conceptualization to model deployment on the edge, work is described within the confines of an NDA

- Designed and optimized models for video detection, video re-identification, user-defined single-object tracking, and tracking of small objects in both short and long-term scenarios
- Developed classification and re-identification models in an open-set context for long-range radar data
- Optimised fleet path-planning
- Conducted quantization techniques to enhance throughput of edge-deployed models

07/2021 – 08/2022
France

Lead ML engineer Jus mundi

Led a team of 6 data scientists and engineers tasked with constructing custom models and utilizing LLMs to deliver enhanced legal decisions to users

- Conceptualized an LLM-based pseudo-anonymization pipeline, spearheaded the identification of inter-document citations of legal entities (POI, WOA, etc.), developed a judiciary reasoning extraction system from court documents, a research effort based onto address long-range context issues in documents exceeding 200 pages
- Engineered a containerized ML training pipeline utilizing France's largest HPC cluster

05/2018 – 07/2021
France

Data team lead CGI

Served as AI Manager for the agency, contributing to multiple client projects as a consultant.

- Implemented NLP and Speech technology for IMA, integrating ASR/STT with in-house NLP models to improve assistance for car failures and accidents, while also working on vehicle classification and repair cost estimation.
- Led deep learning projects at MACIF, creating NLP models for text data processing and chatbot services for over 4.3M users monthly. Also worked on OCR for compliance purposes.

04/2014 – 05/2018
France

Data scientist MAIF

- Developed a multi-modal language model for email classification, summarization, and automatic answering, open sourced as <https://github.com/MAIF/melusine>
- Enhanced fraud detection models to identify under-insured assets using satellite data
- Implemented premium pricing strategies based on driving behavior and GNSS data

Work experience

09/2017 – 09/2019
France

TA machine learning University of Poitiers

Educated graduate students on various machine learning concepts

- Created syllabus outlining curriculum for machine learning course (including decision theory, regression, SVM, CART, random forests, MLP, and convolutional neural networks)
- Utilized programming languages Python and R to teach students key concepts in machine learning.

09/2013 – 03/2014
France

Statistician MAPA

- Conducted comprehensive analysis to interpret pricing drifts, price elasticity, and financial performance
- Utilized advanced statistical methods to identify and explain trends in pricing
- Generated detailed reports outlining statistical findings and recommendations for stakeholders

Education

09/2011 – 06/2013
France

Statistics | MSc Université de Poitiers

09/2008 – 06/2010

Mathematics | Bachelor Université de Paris

Skills

– LANGUAGES

Python



R



rust



SAS



– TOOLS

UNIX, git, redis, MongoDB,
SQL



k8s, nomad, ansible



pytorch, tensorflow, Apache
Spark, sklearn



Publications

12/2021

The Unreasonable Effectiveness of the Baseline: Discussing SVMs in Legal Text Classification

<https://arxiv.org/abs/2109.07234>

09/2021

LegalLMFiT: Efficient Short Legal Text Classification with LSTM Language Model Pre-Training

<https://arxiv.org/abs/2109.00993>

Achievements

- Earned first place at IMA 2019 Hackathon for showcasing skills in Natural Language Processing
- Secured second place in Vision at AflPC x LineUp7 2020 Hackathon
- Ranked in the top 4% in the 2021 MagNet time series competitions