

Marc ALPHONSUS

Data Science Lead, +8 ans d'XP

Data scientist senior avec un large portefeuille d'expérience; du scoring à la détection de fraude, en passant par la reconnaissance d'image et la modélisation du langage dans des contextes big data. Je recherche un poste de lead tech laissant la part belle à la technique. Habitué au travail en équipe, versé à la philosophie devops, j'apporte une attention toute particulière à l'industrialisation et l'automatisation des modèles.

Mes domaines d'expertise tournent autour du Deep Learning en particulier le traitement du langage naturel et l'apprentissage par renforcement. J'adore la combinaison de créativité et de théorie nécessaire dans le job.

IDF, 31 ans
Permis A, B
nationalité Française

alphonsus.marc@gmail.com
[linkedin](#)  / [github](#) 
+33 6 62 62 62 48

Experience

CGI Lead Data Science 2018 - 2020

Manager le lab en offrant une vitrine des savoirs-faire Data Science. Construire et valoriser la stratégie IA des clients Grand Ouest.

MACIF :

- Macif NLP: Création des briques de traitement de langage naturel pour le groupe (routing/classification, NER, analyse du sentiment)- amélioration de la performance de 85% sur 3 classes à **92% sur 56 classes** > `tensorflow`, `huggingface`, `spacy`, `docker`
- Hub IA: Exposition d'une plateforme de services d'IA sécurisée et monitorée utilisables par tous les salariés du groupe sous forme d'un catalogue d'API > `WSGI`, `API Gateway`, `docker`
- AutoML Workflow : Conception et mise à disposition d'un pipeline de réentraînement automatique des modèles en production alimenté grâce à une boucle de feedback des utilisateurs (active learning, multitask-learning) > `python`, `gitlab CI`, `docker`
- RAD/LAD: Mise en place d'un service automatique d'analyse de documents (classification automatique et OCR) **95% sur 80 classes** > `pytorch`, `docker`
- MeetUP: Définition d'une stratégie data à l'échelle du département et évangelisation des concepts clés de la data science auprès des différentes directions du groupe `OKR`, `Gslides`, `Jupyter`
- Chatbot EPI: refonte l'UX de l'espace personnel internaute - la mise en place de l'IA a permis de tenir les objectifs clés de l'entreprise (délai de traitement d'une demande <1 jour) avec une meilleure qualification du besoin grâce à un moteur de (extraction du produit, du bénéficiaire, etc) > `tensorflow`, `spacy`, `rasa`

IMA (InterMutuelles Assistance) :

- Assistance Augmentée: PoC d'une IA d'assistance téléphonique - Combinaison d'un modèle d'ASR/STT et de modèles de traitements du langage (NER, Classification, Similarité) pour assister et accélérer l'assistance lors des pannes et accidents de la route;

Process mis en place : Géolocalisation du client, compréhension de la panne à partir de la description du client, délégation d'un dépanneur/fourniture de service (taxi, hôtel...) et envoi d'un SMS récapitulatif
> [pytorch](#), [Go](#), [docker](#)

CGI - Lab IA Grand Ouest :

- Management et mentorat de collaborateurs, conférences et vulgarisation
- Avant-vente, cadrage/gestion de projet
- Veille technologique et learning days

MAIF Data scientist 2014 - 2018

Apporter aux départements porteurs de projets, l'expertise nécessaire en apprentissage statistique et concevoir des outils au service du sociétaire.

- Mail: Routage automatisé des e-mails et résumé automatique, devenu [Mélusine](#) ; design de l'architecture des réseaux neuronaux et des algorithmes de résumés extractifs > [tensorflow](#), [scikit-learn](#), [docker](#)
- MAIF & Go: Tarification d'une offre automobile connectée basée sur des comportements de conduite remontés via capteurs GPS (IoT) > [R](#), [caret](#)
- Valeur Sociétaire: Modélisation de la rentabilité des clients et prospects à horizon long (LTV) > [pySpark](#), [Hadoop Map/Reduce](#)
- Téléphonie: Prévion de la charge d'appels pour dimensionnement RH des plateaux téléphoniques **taux d'erreur à 5%, proposition de publication** > [R](#), [caret](#)
- Détection d'anomalies: Identification de risques sous-assurés (fraude & devoir de conseil) > [scikit-learn](#)

IRIAF Chargé de TD Machine Learning 2017 - 2019

Initier des étudiants à l'apprentissage machine, au programme: théorie de la décision, régression logistique, SVM, CART, forêts aléatoires et réseaux de neurones.

MAPA Statisticien 2013 - 2014

Concevoir et réaliser des études visant à expliquer et corriger les dérives tarifaires du portefeuille, l'élasticité au prix des sociétaires et les résultats financiers.

Technical Experience & Skills

Projets choisis

TextNets: Mes projets autour du NLP, répliques d'articles, analyse de similarité textuelle, résumé automatique, traduction, search...
DeepCount: dénombrement sur image fortement peuplées, estimation par density map

Languages: [Python](#), [R](#), [SAS](#), [Go](#)

Langues: [English](#), [Français](#), [Deutsch](#)

Frameworks: [tensorflow](#), [pytorch](#), [scikit-learn](#), [Apache Spark](#), [opengym](#)

Cloud: [AWS](#), [Google Cloud Platform](#)

Basics: [UNIX](#), [algorithms](#), [git](#), [docker](#), [redis](#), [MongoDB](#), [SQL](#)

Formal Training & MOOCs

2019	Spinning up Reinforcement Learning ; OpenAI 6.S094 ; MIT
2018	CS224n, CS229, CS231n ; Stanford University
2017	Deep Learning Specialization ; Deeplearning.ai x Coursera
2014	Machine Learning ; Stanford University x Coursera Deep Learning ; Google x Udacity
2013	Master, Statistiques Appliquées ; Université de Poitiers
2010	Licence, Mathématiques Informatique et Applications ; Université Paris V