Michaël SOK

Data Scientist Confirmé

8 rue nicolo 75016 Paris **3** 06 79 76 08 99 ⋈ mdb.sok@gmail.com

Expérience Professionnelle

De décembre 2018 Quantmetry, Data Scientist Consultant, Paris, France.

- à aujourd'hui Banque Génération de données structurées via des GANs (Cramèr-GAN)
 - o Banque Intelligibilité d'un processus de filtrage de messages SWIFT (SHAP)
 - o Média Segmentation de cookies internets à partir de la base CRM
 - Package Graphiques pour des modèles d'intelligibilité additifs (SHAP, LIME, etc.)
 - o Formations Introduction à l'Intelligibilité en ML et Introduction à la Data Science et au Big Data)

De juin 2018 Quantmetry, Data Scientist Stagiaire, Paris, France.

- à novembre 2018 Télécommunication Refactoring et logging de code d'un algorithme Dikjstra
 - o Compagnie Aérienne Segmentation client pour le service Marketing (K-Means)
 - o Banque Intelligibilité de modèle dans le cadre d'un challenge (Shap et Anchor)

De février 2017 InBox, Stage de 6 mois, Malakoff, France.

- à juillet 2017 Retail Création d'un pipeline de preprocessing et d'un modèle de scoring
 - Dashboarding Front-end d'une application avec la librairie R shiny
 - Veille Benchmarking des méthodes pour pallier au problème de déséquilibre de classes

Formation

De 2014 Ingénieur Statisticien Économiste, ENSAE Paristech (92).

à 2018 Spécialisation Data Science - Apprentissage Statistique

Compétences

Programmation Python, R, SQL

Packages python numpy, pandas, scikit-learn, tensorflow, shap, plotly, dash, ...

Cloud Provider Google Cloud Platform

Anglais Fluent - Toeic 2015 - 900 / 990

Certifications

GCP Professional Data Engineer on Google Cloud Platform

Coursera Deep Learning Specialization - deeplearning.ai

Articles

Blog Quantmetry

18 mai 2020 Les exemples adverses ne sont pas des erreurs, ce sont des features!

23 avril 2020 Repenser sa stratégie de prévision et optimiser son activité

21 février 2020 Techniques d'intelligibilité en Deep Learning appliquées à l'image

24 janvier 2020 Tabular data generation using Generative Adversarial Networks

9 septembre 2019 Intelligibilité d'un modèle lié un dispositif clé de la conformité

17 juillet 2019 Paris Data Ladies au Trianon! (NLP, GANs et Graphes)

Présentations

Webinar

20 mai 2020 Rendre les modèles intelligibles pour une IA de confiance

23 avril 2020 Repenser sa stratégie de prévision et optimiser son activité

Meet ups

11 décembre 2019 Paris Machine Learning - Génération de données tabulaires par des GANs

17 octobre 2019 Paris Generative Models - Génération de données tabulaires par des GANs