

Jonathan OHEIX

Freelance Data Scientist

31bis rue Jules Charpentier

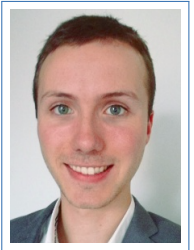
37000 Tours

☎ 06 05 04 74 48

✉ jonathan.oheix@outlook.fr

in jonathanoheix

🌐 jonathanoheix



Expérience professionnelle

- 2017 - janv. 2019 **Data Scientist (1 an et 3 mois)**, *Influans*, Paris 2e.
— Optimisation de moteurs de recommandation via du Machine Learning (Random Forest, ALS)
— Deep Learning et analyse d'image à l'aide de réseaux de neurones convolutifs (CNN)
— Implémentation d'algorithmes d'optimisation : algorithmes génétiques et recuit simulé
Environnement technique : Python, Spark, Scala, Keras
- 2017 **Consultant Data Scientist (11 mois)**, *Louis Vuitton*, Paris 1er.
— Utilisation de Machine Learning (Random Forest) pour prédire la valeur des clients
— Association de produits grâce à des algorithmes de Pattern Mining
— Analyse de données d'objets connectés et présentation des résultats lors de conférences internes
Environnement technique : Python, Spark
- 2016 **Consultant Data Scientist (3 mois)**, *Groupe BPCE*, Paris 13e.
— Segmentation de la clientèle professionnelle à l'aide d'arbres de régression
Environnement technique : SAS
- 2016 **Stagiaire Consultant Data Scientist (6 mois)**, *SoLocal Group*, Boulogne-Billancourt.
— Optimisation d'un moteur de recherche : Learning to Rank, Query Expansion avec Word2Vec
Environnement technique : Python, Elasticsearch, Programmation Web

Projets Data Science

- 2019 **Reconnaissance faciale en temps réel avec du Deep Learning.**
— Création et entraînement d'un réseau de neurones convolutifs (CNN)
— Détection et classification en temps réel des visages via une interface web
Environnement technique : Python, Keras, OpenCV, Flask
- 2018 **Analyse de sentiment des avis clients de Booking.com via traitement du langage (NLP).**
— Nettoyage des données textuelles brutes : étiquetage morpho-syntaxique, lemmatisation
— Feature engineering : calcul de scores de sentiments, transformation du texte avec Doc2Vec
— Classification des avis (bon/mauvais) à l'aide de Machine Learning (Random Forest)
Environnement technique : Python, NLTK, Gensim, Scikit-Learn, Pandas

Formation

- 2013–2016 **Ensai (École nationale de la statistique et de l'analyse de l'information)**, *Rennes*.
Spécialisation : Statistique et ingénierie des données
- 2011–2013 **DUT Statistique et Informatique Décisionnelle**, *IUT de Vannes*.

Compétences

Machine Learning	Natural Language Processing, Deep Learning, Random Forest, Gradient Boosting, Séries temporelles, SVM, Régression pénalisée, Learning to Rank, Pattern Mining, GLM, Analyses factorielles
Programmation	Python, Git, Scala, R, SQL, Java, Programmation web
Big data	Spark, Hadoop, Elasticsearch, Pig, Hive, H2O
Cloud computing	Google Cloud, AWS
Anglais	Courant : TOEIC 945/990 points, travail en Irlande pendant 1 mois

Centres d'intérêt

- Data Science Participation à des challenges : Kaggle, Le Meilleur Data Scientist de France 2018
Rédaction d'articles pour la page Towards Data Science