



LEJEUNE Arnaud

Data Scientist

www.alejeune.com

(438) 505-5723 - arnaudlejeune44@gmail.com

Description

Ingénieur Français, passionné des problématiques autour de la data et du machine learning, ainsi qu'un joueur de jeux video depuis l'enfance. J'ai décidé d'allier ma passion pour le jeux video et mes compétences professionnelle pour évoluer dans cette industrie passionnante.

Vit actuellement à Montréal (Canada), en PVT jusqu'en Août 21' et en attente de résidence permanente.

Compétences

Outils :

- | | |
|----------------|-----------------|
| / Pandas | / TensorFlow |
| / Scikit-Learn | / Keras |
| / Spark | / AWS / GCP |
| / MongoDB | / Ubuntu / UNIX |

Langages :

- | | |
|----------|--------|
| / Python | / Java |
| / Scala | / WEB |

Langues :

- | |
|----------------------------------|
| / Français : langue maternelle |
| / Anglais : Bilingue (TOEIC 965) |

Certifications :

- | |
|---|
| / Data Engineering on GCP (en cours de passage) |
|---|

Divers

/ Président du B.D.E d'IMT Atlantique campus de Nantes 2016 – 2017.

/ Gardien de handball en championnat universitaire et 10 ans de tennis

Expériences

Data Scientist – Eidos-Montreal

Août 2019 –

Au sein de l'équipe "Analytics and Data Science", participation à divers projets d'analyse et de prédiction afin d'assurer le suivi de projets passés (Deus Ex : Mankind Divided & Shadow of the Tomb Raider) et l'effort de production de projets en conception / production (non annoncés)

/ Mise au point d'un outil de NLP visant à catégoriser et synthétiser les avis utilisateurs et critiques professionnelles. Transfer learning et fine-tuning d'un modèle BERT sur 500k reviews scrappées (metacritic, IGN, steam ...). Supervision de la labélisation d'un dataset de reviews en interne et entraînement de modèles de classification (catégorisation et sentiment).

/ Exploration et nettoyage d'une base de données live de ventes digitales au niveau du groupe Square Enix. Mise au point de modèles de prédiction de ventes sur nos titres live. Utilisation de méthodes classiques de Time-Series et modélisation RNN / LSTM.

/ En charge de l'organisation et des processus de versioning de notre base de code. Mise en place d'un GitLab commun et setup complet d'une VM commune pour la Data Science.

/ Développement de scripts permettant l'automatisation de tâches de nettoyage et preprocessing de données internes. Utilisation d'Airflow pour paramétriser les jobs et interaction avec Google Big Query.

Consultant Data – Wavestone

Sept 2017 – Août 2018

/ Etude des problématiques liées au Big Data, aux bases de données Graphes et aux bases NoSQL. Outre une cartographie des différents cas d'usage, j'ai réalisé un PoC sur Neo4J (en python) et un benchmark des différentes solutions Graphe.

/ Participation à diverses missions de conseil pour des grands comptes français (SNCF, ENEDIS, Ministère des Armées) sur des sujets autour du Big Data, du Machine Learning et de l'architecture de leur SI.

Formation

Mastère Spécialisé Big Data – Télécom ParisTech

2018 – 2019

- | | |
|---------------------------------------|---|
| / Machine Learning (Sup. et non-sup.) | / Statistiques avancées |
| / Data Engineering (Spark & Hadoop) | / Systèmes NoSQL (MongoDB, Neo4J, ...) |
| / Deep Learning (CNN, RNN, ...) | / Reinforcement Learning (QL, SARSA, ...) |

Ecole d'ingénieur généraliste – IMT Atlantique

2015 - 2018

Majeure : Organisation et Management des Technologies de l'Information

- | | | |
|---------------------------|--------------------------|----------------------------|
| / Big Data | / Business Intelligence | / Architecture applicative |
| / Architecture Logicielle | / Conduite du changement | / Gestion de projet |

Tronc Commun :

- | | |
|---------------------|---------------------------------------|
| / Statistiques | / Programmation (Impérative, POO ...) |
| / Sciences sociales | / Gestion des bases de données |

Classes préparatoires MPSI/MP – Lycée Montaigne

2013 - 2015