

TRISTAN MUSCAT

DATA SCIENTIST

 Paris, France  Tristan Muscat  06 11 62 52 59  tristan.muscat@protonmail.com

 <https://datartichaut.com>

EXPÉRIENCE PROFESSIONNELLE

Orange Cyberdefense

Lead Data Scientist - Sep. 2023 - Oct. 2024, Paris

Environnement technique : GCP (Kubernetes, Helm, Terraform), Python (scikit-learn, MLflow, Airflow, FastAPI, Celery)

Sujet : Création d'un modèle de détection de phishing pour sécuriser la navigation internet des utilisateurs.

- Utilisation de MLflow model registry pour versionning et suivi des modèles ainsi que déployer des serveurs d'inférence dans le cluster kubernetes (GKE).
- Utilisation d'airflow pour automatiser les tâches liées aux données.
- Création d'un dashboard streamlit pour le suivi du modèle, reporting Evidently, et explicabilité avec Shap.
- Création d'un pipeline de modélisation allant du parsing jusqu'aux predictions.
- Création de pipeline CI via Gitlab CI et la mise en place d'Argo CD pour le déploiement.
- Gestion de l'infrastructure dans un cluster Kubernetes (GKE) déployé avec Terraform, et déploiement des composants avec Helm.

Kereon Intelligence

Data Scientist - Oct. 2020 - Oct. 2022, Paris

Missions

- **Hackathon IMA Innovation, 2022, deuxième place :** "Amélioration d'un modèle de prévision en fonction de données météorologiques".
- **Projet interne, 2022 :** projet d'extraction de texte de scan de livres, réalisé en environnement Microsoft Azure.
- **Ocean Hackathon, 2021 :** organisation d'un événement data et innovation en partenariat avec le Village By CA.
- **MACIF, 2021 (SQL, Python) :** réalisation de l'architecture data nécessaire à la création d'un tableau de bord de pilotage à chaud en temps réel.
- **Excelia, 2021 :** formation "Data for Business". Supports de cours, études de cas, évaluations.
- **Territoria Mutuelle, 2021 :** réalisation d'un diagnostic de data gouvernance.
- **Cerfrance, 2021 (Python, Hasura, GraphQL) :** automatisation de traitements de données pour produire un reporting comptable.
- **France Active, 2020 (Python) :** création d'un modèle de scoring de risque dans le cadre d'ocroi de crédits.

MENTORAT

OPENCLASSROOMS, 2020 - 2024

Mentor et Evaluator chez Openclassrooms, parcours Data Scientist et Data Analyst

COMPÉTENCES

Data Science

Programmation : Python, R.
Libraries : scikit-learn, H2O.
Softwares : Dataiku.

Data Engineering

SGBD : SQL, MongoDB, Redis.
Backend : Celery, FastAPI.
Outils : MLflow, Airflow.

Data Visualisation

Python : Dash, Streamlit.
Softwares : Tableau, PowerBI, Dataiku.

Data Management

RGPD.
Qualité des données : Great Expectation.

Cloud

Azure Certifié.
GCP : un an d'expérience.
AWS trois ans de mentorat.

Build

Docker.

Déploiement

Kubernetes / Helm / Terraform.

CI/CD

Gitlab CI.
Argo CD.

LANGUES

- **Français** (langue maternelle)
- **Anglais** (courant)

EXPÉRIENCE PRO ...

- SUITE

AddixWare

Data Scientist - Nov. 2019 - Nov. 2020, Paris

Missions

- **LVMH, 2020 (Python)** : modélisation d'une grandeur physique (point éclair) de cosmétiques pour répondre à un cahier des charges.
- **Covid-19, 2020 (Python)** : étude des facteurs aggravants du covid-19 : travail sur une base de données sur la propagation du virus en Chine et au Mexique pour analyser les causes et obtenir un diagnostic.
- **Sanofi Pasteur, 2019 (Python)** : accompagnement en data science sur la création d'un processus industriel breveté.

Magellan Consulting

Data Scientist - Mai 2019 - Août 2019, Paris

Missions

- **Crédit Agricole, 2019** : Travaux de cadrage de refonte de la gestion des connaissances à l'échelle Groupe.
- **Interne, 2019 (Python, TensorFlow, Android Studio)** : Travaux sur le thème des services cognitifs en machine learning: computer vision (Convolutional Neural Network).

Voirin Consultants

Data Scientist - Avril 2018 - Janvier 2019, Paris

Missions

- **SIAAP, 2018** : Rédaction d'un document sous forme de vade-mecum, sur l'utilisation de la data dans l'entreprise, de l'implémentation d'une gouvernance à la mise en production d'une solution.
- **Caen-la-mer, 2018** : Élaboration d'un schéma directeur des systèmes d'informations pour une communauté urbaine.

Optimind Winter

Data Scientist - Mars 2016 - Avril 2018, Paris

Missions

- **Carte Blanche, 2018** : Chiffrage de l'impact financier sur le domaine de la santé du projet RàC zéro.
- **Prévère, 2017, RPA (UiPath)** : Développement d'un programme (UiPath) permettant la lecture automatique et le traitement de fichiers PDF.
- **La Poste, 2017 (R, RShiny)** : Modélisation et identification des drivers d'absentéisme.
- **EURIA, 2017, (R)** : Encadrement de deux bureaux d'étude Data Science en provisionnement et segmentation client.
- **Interne, 2017, (R)** : Modélisation de prédiction de migration de rating pour optimiser le SCR de marché en assurance. Utilisation de facteurs externes, analyse de sentiments de revues financières (NLP).
- **Barclay's, 2017, (R)** : Modélisation d'écoulements de flux de passifs (dépôts à vue et épargne en banque).
- **Predica, 2016, (R)** : Modélisation de comportement client concernant les versements, arbitrages et rachats en assurance vie.

FORMATION

Certifications Microsoft Azure

2023, Paris

AZ-900 - DP-900 - AI-900

Certifications Dataiku

2023, Paris

Dataiku Core Designer

Université Dauphine, Paris

2014-2016, Paris

M1 MIDO, M2 Actuariat

E.I.S.T.I.

2013-2016, Paris

Diplôme d'Ingénieur

Université de Cergy Pontoise

2013-2014, Paris

Licence Informatique

Lycée Saint Joseph, Avignon

2011-2013, Avignon

Classe Préparatoire (CPGE-PC)

Lycée Saint Louis, Orange

2012, Orange

Baccalauréat Scientifique

CENTRES D'INTERETS

Membre de Kactu'Z Actuariat, 2017 -

Organisation d'événements pour la communauté actuarielle.

Trésorier LADD, 2018 -

Les Actuaire De Dauphine