



06 16 46 98 21

stordeur.bru@gmail.com

[Portfolio](#)

Formation

Data Science & Business Analysis
- EDC Paris Business School

Mathématiques appliquées,
statistiques, parcours data science
- Université Aix-Marseille

Certifications

Data Upskilling
- Benjamin Dubreu

Azure AI Engineer Associate
- Microsoft

Dev/Data Security Champion
- Aroua Biri

Langues

Français

Anglais

Bruno
Data scientist

Expérience
3
ans

Mes motivations

Passionné par la data et le machine learning, je développe et déploie des modèles fiables et, tout en étant capable d'expliquer simplement les solutions aux profils non techniques.

Compétences clés

Techniques :

- **Langages** : Python, SQL, Pyspark
- **Manipulation des données** : Pandas, Numpy, Spark, Dataluku
- **Visualisation des données** : Matplotlib, Seaborn
- **Machine learning** : Keras, Tensorflow, Sklearn, MLflow
- **Tests** : Black, Pytest
- **API** : FastAPI
- **Prototype** : Flask, Streamlit
- **Cloud** : Azure (Certification AI Engineer Associate)
- **Orchestration** : Airflow,
- **Gestion du code**: Git, GithubAction
- **Gestion de l'environnement** : Docker, Poetry, DVC

Fonctionnelles :

- **Méthodologie** : Agile Jira,/Scrum
- **Machine learning** : NLP, Computer vision, Deep learning
- **Communication** : Vulgarisation d'un sujet technique
- **Veille technologique** : Synthétiser l'état de l'art d'un sujet

Expériences professionnelles

Cartes Bancaires – Data scientist (1 an)

Elaboration et déploiement de modèles destinés au scoring et à la détection de fraudes.

BNP Paribas – Data Analyst / Scientist (6 mois)

Analyse des incidents historiques afin d'identifier des zones à risques et de réduire les coûts

CHU de Brest – Data scientist (6 mois)

Elaboration de modèles sur des problématiques liées aux évènements d'hypertension intracrânienne

DigiTrips – Data scientist (1 an)

Elaboration de modèles de prédiction de réservation de vols ainsi que d'optimisation de coût par classification de la pertinence de recherche de vol

Cartes Bancaires

Data scientist

1 an

Projet :

Mise en place d'un système de scoring et de détection de fraude sur un grand projet européen (EPI)

Contexte :

L'European Payment Initiative (EPI) vise à créer un système de paiement paneuropéen unifié pour les consommateurs et les commerçants, en offrant des solutions innovantes comme les paiements instantanés et un portefeuille numérique. Dans ce contexte, la prévention et la détection de la fraude sont essentielles pour assurer la sécurité et la confiance dans ces nouvelles méthodes de paiement.

Missions :

- Définition de Data Catalog pour structurer les données (P2P et e/m-commerce)
- Assistance à l'intégrations des nouveaux data catalog dans les systèmes existants
- Création de règles de "cold start" pour pallier l'absence de données initiales
- Monitoring de l'activité et du scoring lors du lancement sur IBM Safer Payment
- Exploration et analyse des données pour une compréhension approfondies
- Développement de nouvelles règles de scoring en utilisant des modèles de machine learning et les retours de fraude des banques
- Veille technologique et présentation d'articles sur le sujet

Environnement technique :

Python, SQL, Pyspark, Github, Apache Zeppelin, Hue, IBM

BNP Paribas

Data Analyst

6 mois

Projet :

Analyse des données issues des incidents historiques afin de cibler les zones à risques et de réduire les coûts

Contexte :

Au sein de la division Risk ORM de BNP Paribas, la gestion des risques opérationnels est cruciale pour la surveillance et la réduction des risques à travers toutes les activités bancaires. L'analyse des incidents historiques aide à identifier les zones de vulnérabilité.

Missions :

- Analyse et explorations des données issues des incidents historiques.
- Analyse des corrélations entre les incidents historiques, leurs causes, type de risques et business unit défaillantes afin de cibler des zones à risque
- Analyse des contrôles mis en place
- Mise en parallèle des incidents historiques avec les contrôles mis en place afin de déterminer des périmètres à risque
- Exploitation et analyse des impacts de certaines business unit défaillantes
- Vulgarisation et restitution des résultats

Environnement technique :

Dataiku, Python, SQL

CHU Brest

Data scientist

6 mois

Projet :

Génération de données temporelles synthétiques et prédiction des événements d'hypertension intracrânienne.

Contexte :

Ce projet s'inscrit dans l'aide que la machine learning et deep learning peut apporter aux médecins du CHU pour prendre de meilleures décisions quant aux traitements des événements d'hypertension intracrânienne chez des patients.

Les trois principales missions sur lesquels j'ai pu évoluer ont été:

1. La génération de données synthétiques de qualité afin d'enrichir les datasets.
2. La prédiction de la pression artérielle et pression intracrânienne chez les patients.
3. La prédiction d'événements d'hypertension intracrânienne

Missions :

- Veille technologique et synthèse de l'état de l'art sur le sujet
- Utilisation d'un algorithme de deep learning (DGAN) pour générer des données synthétiques et définition de métriques (DTW) pour filtrer la qualité des données générées
- Extraction / nettoyage / préparation des données et feature engineering
- Prediction step by step à l'aide d'un modèle de deep learning
- Classification binaire à l'aide de modèles de machine learning

Environnement technique :

Python, SQL, Sklearn, Pandas, Numpy, Keras, TensorFlow, Linux, Bitbucket, Docker

Digitrips

Data scientist

1 an

Projet :

Optimisation du cache mémoire en classifiant les recherches de vols pertinentes ainsi que la prédiction de futures réservations de vols.

Contexte :

Les réservations de vols sur le site de Digitrips étaient systématiquement stockées dans le cache mémoire, peu importe leurs pertinences, ce qui engendre des coûts importants et de la latence pour les utilisateurs finaux. Lors de cette mission j'ai pu mener deux projets en parallèle.

1. Analyse des données issues des recherches et réservation de vols ainsi que le tri de leurs pertinences afin de libérer de l'espace du cache mémoire
2. Prédiction des futures réservations de vols.

Missions :

- Extraction des données depuis les bases de données SQL
- Analyse et explorations des données, features engineering et features selection
- Entraînement de modèles de machine learning et/ou deep learning
- Optimisation des modèles
- Time series forecasting pour la réservation de vol (ARIMA, SARIMA, LSTM)

Environnement technique :

Python, SQL, Sklearn, Pandas, Numpy, Keras, TensorFlow, Linux, Git