

Fedi HAMDI

[in linkedin.com/in/fedi-hamdi](https://www.linkedin.com/in/fedi-hamdi) | [0622039329](tel:0622039329) | fedihamdi.netlify.app | fedihamdi@yahoo.com | [GitHub](#)

Skills

- Python | Spark | R | SQL | Flask | Django | ETL | PostgreSQL | NoSQL | KQL | Databricks | Docker | Linux | PySpark | Jenkins | Hadoop | Git
- Azure | AWS | Cloud Computing | CI/CD | Pytest | Integration Testing | GitHub Actions | Grafana | Alerting | Unit Testing | Splunk | OOP
- Microservices | MLOps | Airflow | MLflow | DVC | API | Machine Learning | Tableau | PowerBI | Time management | Budget management
- 3 ans d'expériences | Autonomie | Curiosité | Dynamisme | Teamwork | Problem solving | Anglais, Japonais, Allemand

Experience

ML Engineer / Data Scientist

Engie

La Défense 12/2022 – 01/2024

- Amélioration de la production d'énergie de 5 % en un an grâce à la détection d'anomalies et à des modèles de maintenance prédictive.
- Développement et mise en production d'algorithmes avancés visant à optimiser la production d'énergies dans le monde.
- Déploiement, orchestration et monitoring de modèles de machine learning visant à améliorer les rendements des éoliennes (en France, Italie, USA, etc).
- Analyse des anomalies de performance sur des assets en vue d'optimiser la production d'énergie renouvelable et d'améliorer le processus de maintenance prédictive.
- Mise en place de modèles d'analyse de la qualité des données dans le but de garantir des résultats fiables (Data Quality and UK Data Governance Frameworks).
- Déploiement des connecteurs vers des clusters ADX (Azure Data Explorer) en vue de simplifier l'accès aux données, entraînant une réduction des coûts de 15% soit 50 milles euros € par an à l'échelle de mon équipe (estimé).
- Développement des dashboards et de systèmes d'alerte sous Grafana assurant une surveillance en temps réel des actifs d'Engie USA, contribuant à une réduction d'effort de 60%.
- Gestion des lancements de packages de machine learning internes, implémentation de tests unitaires et d'intégration.
- Environnement : Python, Spark, R, Git, CI/CD, Splunk, Jenkins, Airflow, MLflow, Pytest, Docker, Kubernetes, ttest, SonarQube, AureML, Azure Storage, Databricks, Artifactory, Grafana, ADX, DVC.

Data Engineer

IPSEN Pharma

Boulogne-Billancourt 06/2022 - 11/2022

- Conception et mis en production de deux projets stratégiques en relation avec le FDA américain : une bibliothèque de codes intégrant les concepts cliniques clés et un centre de données épidémiologiques mondial.
- Dirigé le développement d'une bibliothèque de codes encapsulant les concepts cliniques majeurs (symptômes, diagnostic, causes et traitements).
- Conception et réalisation intégrale du déploiement d'un centre de données épidémiologiques mondial, contribuant ainsi à l'optimisation de la gestion des données médicales au sein de l'entreprise.
- Mise en place d'APIs stratégiques assurant une communication transparente entre les applications, optimisant ainsi la documentation médicale, la facturation des services de santé en relation avec les assurances et la collecte de données épidémiologiques.
- Réduction des coûts de mise en production de 20%.
- Environnement : R, Python, Rstudio, AWS, EC2, ECR, S3, Conda, Shell, Git, Docker.

MLOps Engineer

JustAI

Montreuil 01/2022 - 05/2022

- POC d'aide à la décision basé sur l'analyse des émotions lors d'un entretien.
- Implémentation du scraping en temps réel des données LinkedIn afin d'enrichir les informations contextuelles.
- Mise en production des modèles de deep learning qui détectent les émotions et les attributs faciaux en contexte d'entretien.
- Environnement : PyTorch, Torchvision, OpenCV, Pandas, Numpy, CLIP-OpenAI, DeepLabV3+, VGG16, KNN, Gradio, HuggingFace.

Data Scientist

Business and Decision

01/2020 - 12/2020

- Évaluation des risques du crowdfunding axée sur l'Intelligence Artificielle : use-case : Fintech Kiva.
- Conduite d'analyses exploratoires et de pré-traitements dans le but d'évaluer les risques du crowdfunding.
- Développement de modèles de Deep Learning et de Machine Learning visant à l'évaluation des risques avec une précision de 90%.
- Environnement : Python, Statsmodels, Tensorflow: CNN, LSTM, Embeddings, Scikit-Learn: RF, LR, SVM, SMOTE, Pandas, NLTK.

Education

Master

HETIC

Paris 11/2021 - 12/2023

- Master Big Data & Intelligence Artificielle – Titre RNCP niveau 7

Diplôme d'ingénieur

ESSAI

09/2016 - 12/2020

- Statistique et Analyse de l'information (Option : Actuariat)

Projets

- **BreatheEasy** : Application préventive des maladies liées à l'environnement, identifiant les zones à risques pour une santé optimale. ([Code](#))
- **Fendi** : Package Python pour créer des apps web propulsées par l'IA Générative personnalisées avec CV et profil LinkedIn. ([Doc](#)) ([Demo](#))