



STAGE DE 06 MOIS EN DATA

Pourquoi moi ?

Ce qui me distingue, c'est mon aptitude à intégrer les théorèmes de statistiques et de probabilités dans des contextes quotidiens. Avez-vous, par exemple, envisagé l'application de la loi Beta pour estimer la probabilité de succès d'un projet ?



+33 7 82868996



gasyd.gagnon@groupe-esigelec.org



Mobilité nationale



<https://www.linkedin.com/in/gasyd-gagnon/>

Skills

- Data science
- Data visualisation
- Data engineering
- Esprit d'équipe
- Esprit d'écoute
- Développement front-end et back-end

Education

DIPLÔME D'INGENIEUR ESIGELEC

2021-2024

PROJETS RÉALISÉS

GOCOD

Octobre 2023

Outil de création de projets assistée par IA, permettant de générer un projet en quelques minutes grâce à un questionnaire simplifié, accompagné de recommandations personnalisées.

Compétences : FastAPI (python), Reactjs, MongoDB, Neo4J, Git, CI/CD, Déploiement avec terraform, Tests unitaires avec Pytest, Docker

ANALYSE DE DONNÉES EN TEMPS RÉEL

Novembre 2023

Analyse de sentiments dans des commentaires **Reddit** en temps réel avec spark streaming et apache kafka.

Compétences : Apache **Kafka**, **Spark**, **Python**, **Spak SQL**, **Textblob**, **Spark Streaming**, **NLP**

Chatbot avec RAG (Retrieval-Augmented Generation)

Septembre 2023

Utilisation de LLaMA 2.0, FAISS et LangChain pour la réponse aux questions sur des documents.

Compétences : Maîtrise des mesures de recherche par similarité vectorielle (distance euclidienne, similarité cosinus), **NLP**, **Vectorisation de textes**, **RAG**

Experience

STAGIAIRE DATA SCIENTIST CHEZ HANALYTICS

Juin - Octobre (2023)

Data Science :

Prédiction temporelle avec **ARIMA**, **SARIMA** et **Prophet**

Développement et publication d'une bibliothèque python sur pypi

Clustering avec **K-Means** : segmentation client

Prédiction de churn avec des modèles de régression logistique

Data engineering :

Création et orchestration de pipeline de données

Création et sécurisation de VM Google avec déploiement d'applications

Data analyzing :

Modélisation de données avec **DBT**

Requêtes SQL (CTE, Windows functions, subqueries)

Datavisualisation avec **Looker studio**, **Power BI** et **Metabase**