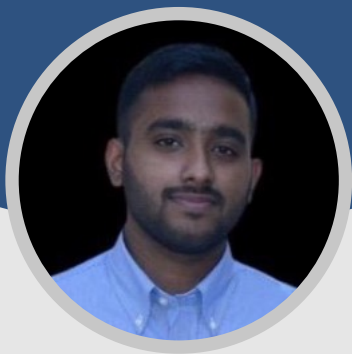


Navirash THARMASEELAN



Informations personnelles



Navirash THARMASEELAN



navirash.tharmaseelan@outlook.fr



+33 6 51 62 86 44



linkedin.com/in/navirash-tharmaseelan

Langues

Anglais (TOEIC 855) ●●●●●●

Français ●●●●●●

Tamil ●●●●●●

Certification

Microsoft Certified : Azure Data Engineer Associate
mars 2023

Microsoft Certified : Azure Data Scientist Associate
mars 2023

Microsoft Certified : Azure Data Fundamentals
mars 2023

AWS Cloud Quest : Cloud Practitioner
mars 2023

Microsoft Certified : Azure Fundamentals
mai 2022

Centres d'intérêt

■ Football, culture Japonaise, nouvelle technologies, automobile

Soft-skills

■ Communication, Sens de l'initiative, Sens de l'organisation

Profil

Prochainement diplômé en tant qu'ingénieur généraliste en informatique et numérique spécialisé en data science et data engineering. Je suis à la recherche d'un CDI en tant que data scientist, data engineer, MLOPS engineer ou DevOps à compter de septembre 2023. Il me tarde de rejoindre vos équipes pour me confronter à de beaux défis et consolider mes compétences.

Expérience professionnelle

Data scientist / Data engineer

de sept. 2022 à août 2023

Société Générale, Val de Fontenay

- Étude de cas sur les larges langages modèles (LLM) dans le secteur bancaire.
- Intégration du modèle Vicuna, modèle basé sur LLaMA entraîné avec GPT-4 dans le SI Société Générale sur AWS en utilisant Docker.
- Analyse conversationnelle chatbot, mail : clustering (BERTopic, DBSCAN, AutoCluster, KMeans, Agglomérative Clustering, Spectral Clustering).
- Mise en place d'un batch en Spring Boot Java pour supprimer les documents datant de plus de 3 mois dans une base de données MongoDB.
- Développement d'une API (Flask) pour un projet d'analyse et de classification sémantique d'échanges de mails.
- Développement d'une interface (ReactJS, Java Spring Boot)
- Création d'un Github app pour le projet d'analyse et de classification sémantique d'échanges de mails pour pouvoir générer un token de façon automatique et cloner le projet sur des instances EC2 (cloud AWS).
- Mise en place des credentials disponibles dans le projet d'analyse et de classification sémantique d'échanges de mails dans le Vault et pouvoir y accéder à partir des instances EC2.
- Mise en place des tests unitaires, intégration et stress test (Java, Python).
- Déploiement du projet d'analyse et de classification sémantique d'échanges de mails sur AWS via des scripts packer, terraform.

Machine & Deep Learning engineer

de nov. 2021 à avr. 2022

Société Générale, Val de Fontenay

- Réaliser le prototypage, collecter et préparer les données.
- Prise en main de Mélusine : bibliothèque Python permettant la classification d'e-mails.
- Nettoyage des données (Lemmatisation, suppression des stopwords...).
- Entraînement et déploiement de modèles d'intelligence artificielle (CNN, RNN, SVM et Random Forest) en Python dans le cadre de la classification de messages entre un client et son conseiller (création de 20 modèles de Random Forest)
- Élaboration de nouvelles stratégies pour améliorer les modèles d'intelligence artificielle (augmentation du F1-score de 0.89 à 0.99)
- Réalisation d'une synthèse des différents travaux (Data visualisation avec Python)
- Intégration des modèles dans un processus applicatif.

Formation

Ingénieur généraliste en informatique et numérique

de sept. 2019 à août 2023

EFREI Paris, Villejuif

Matières enseignées : Big Data Framework, Scala, Spark, Deep Learning, Machine Learning, MLOPS, VueJS, NodeJS, ReactJS, Python, Java, Docker, Jenkins, cloud Azure, Statistiques, Linux, SQL, No-SQL, Data management.

CPGE - MPSI

de sept. 2018 à août 2019

Lycée Marcelin-Berthelot, Saint Maur

Projets académiques

SafeChild

de sept. 2022 à janv. 2023

EFREI Paris

- Service permettant de détecter l'harcèlement dans les réseaux sociaux.
- Rôle : Data Engineer & Chef de projet
- Outils utilisés : Python (Flask, Scikit-learn), ReactJS, Tensorflow, Docker, Grafana, Jenkins, JIRA, GitHub.
- Compétences : Autonomie, management d'équipe, connaissance techniques.

Finance Use Case

EFREI Paris

- Construire un projet de ML pour la classification du risque de crédit immobilier.
- Rôle : Data Scientist
- Outils utilisés : VS code, JIRA, Github, Airflow, MLFlow, Docker.
- Compétences : Sphinx (documentation), cookie cutter, Airflow, MLFlow, ML algorithmes, Docker, Shap, Github actions.