



+33 634164788



sahishnutaatosh@gmail.com



Paris, France



[linkedin.com/in/sahishnuta/](https://www.linkedin.com/in/sahishnuta/)



<https://github.com/Sahishnuta>

COMPÉTENCES

COMPÉTENCES TECHNIQUES

- Python, PySpark
- Libraries: NumPy | Pandas | Scikit-learn | PyTorch | TensorFlow | OpenCV | NLTK | spaCy | Matplotlib | Dash | SHAP
- Statistics | Probability
- MySQL | PostgreSQL | MongoDB
- Machine Learning
- Azure Databricks, Azure Synapse, Azure Data Lake Storage, Azure Key Vault
- Snowflake
- Visualization – Tableau | Power BI
- Tools (GIT | Docker | Airflow)

COMPÉTENCES GÉNÉRALES

- Travail d'équipe | Intégrité | Résolution de problèmes | Prise de décision | Capacité à travailler dans les délais impartis.

LANGUE

- Anglais – Fluent
- French – Intermédiaire B2

CERTIFICATION

[PGP Certification in Data Science](#)

LOISIRS

- Cuisiner
- Écouter de la musique Bollywood

Sahishnuta TOSH

Professionnel des données en science des données et ML avec permis de travail 2027

À la recherche d'un poste dans le Data Science/ Ingénierie des données

EXPÉRIENCE PROFESSIONNELLE

ML Ingénieur, Omdena

À distance

Plateforme IA résolvant des défis réels de manière collaborative

Mai 2024 – Présent

- Développement d'un outil utilisant un LLM pour l'analyse en temps réel des données d'actualité de Kitwe
- Formation en NLP et LLM

Data Science Stagiaire, Amadeus IT group

Nice, France

Entreprise de voyage servant les compagnies aériennes

Janvier 2023 - Juillet 2023

- Améliorer les prévisions du trafic de passagers pour estimer les parts de marché et les recettes
- Collecte de données par web scraping, analyses et visualisations en utilisant Spark SQL
- Expérimentations avec des modèles ML Amélioration de la précision des prévisions et de l'explicabilité avec PySpark (SparkML)

Data Engineer, Infosys Private Limited

Bangalore, Inde

Cabinet de conseil spécialisé en IA et cloud

Janvier 2020 – Août 2021

- Développement d'un pipeline d'analyse des sinistres et segmentation client pour des offres personnalisées.
- Conception de l'architecture Delta Lake sur ADLS Gen2 avec des couches bronze, argent et or.
- Transformation des données via Databricks et Synapse, et intégration avec Power BI.
- Automatisation des flux de données avec les outils Azure, réduisant les coûts opérationnels.

Data Scientist, L&T InfoTech

Mumbai, Inde

Entreprise spécialisée dans le cloud

Mars 2018 - Septembre 2019

- Implémentation d'une tarification forfaitaire pour prédire les coûts hospitaliers
- Modèle de classification intégrant régression et clustering : Random Forest, KNN, SVM, boosting
- Modèle de séries temporelles pour les données des processus d'affaires
- Déploiement de modèles ML via Docker et Flask

Application Development Analyst, Accenture

Bangalore, Inde

Entreprise spécialisée en IT, conseil et transformation digitale

Mai 2015 - Mars 2018

- Close Loop Automation pour optimiser l'efficacité via des applications automatisées.

FORMATION

EPITA - Master of Science (BAC+5)

Paris, France

Systèmes d'intelligence artificielle

2021- 23

Silicon Institute of Technology - Bachelors

Bhubaneswar, Inde

Électronique et télécommunication

2010 - 14

PROJETS

- Création d'une application qui génère des images d'enfants à l'aide de Pix-to-Pix, FamGAN
- **Prédiction des prix des voitures d'occasion** : Prédiction des prix des voitures d'occasion et déploiement des modèles ML avec **FastAPI** et **Streamlit**, stockage des prédictions avec **PostgreSQL**, suivi et réentraînement avec **MLFlow**.
- **Détection d'anomalies** dans les données de capteurs de pompes à l'aide d'une approche heuristique, Isolation Forest et Prophet, avec des tests de stationnarité. Les prévisions sont faites avec **AR, Moving Average, ARIMA, et Auto-ARIMA**.