

VISHRUT BEZBARUA

☎ (+33) 783235990 | ✉ vishrutbezbarua@gmail.com | in [vishrut-bezbarua](#)

EXPÉRIENCE PROFESSIONNELLE

DATA SCIENTIST (STAGIAIRE)

Février 2024 — Juillet 2024

Observatoire Astronomique de Strasbourg

Strasbourg, France

- Traitement de 20 Go de données 3D par découpage, lissage et normalisation, réduisant la taille de 50 % et le volume de 87 %.
- Développement d'algorithmes personnalisés en PYTHON pour extraire automatiquement des structures d'intérêt.
- Création de pipelines de données pour mesurer l'évolution de six propriétés de ces structures sur quatre intervalles de temps.
- Réalisation d'analyses statistiques et mise en œuvre de protocoles de contrôle de la qualité, en veillant à l'exclusion des régions marginales et des valeurs aberrantes afin de préserver l'intégrité des données.

CHERCHEUR SCIENTIFIQUE (STAGIAIRE)

Avril 2023 — Juillet 2023

Laboratoire Physique et mécanique des milieux Hétérogènes

Paris, France

- Simulation de la convection 2D de Rayleigh-Bénard en FORTRAN 90, avec résolution de systèmes 180×180 sur 600 itérations.
- Étude des nombres de Rayleigh jusqu'à 10 000 pour modéliser les phénomènes thermiques.
- Implémentation d'une décomposition LU, simplifiant les calculs de 50% pour traiter 100 000 points de données.
- Optimisation des calculs de 30 %, améliorant les performances des simulations.

DATA ANALYST (STAGIAIRE)

Mai 2021 — Juillet 2021

Indian Institute of Astrophysics

Bangalore, Inde

- Extraction, prétraitement et ingénierie des caractéristiques des données stellaires avec *pandas*.
- Visualisation des paramètres interdépendants à l'aide de *matplotlib*.
- Comparaison de différents types d'étoiles pour vérifier un lien évolutif.

PROJETS PERTINENTS

Classification du cancer du sein à l'aide de réseaux neuronaux ([lien](#))

Nov 2024

- Développement d'un réseau neuronal pour classer le cancer du sein à partir de 30 caractéristiques sur 569 instances.
- Prétraitement des données avec *Gaussian Scaling*, atteignant 99,78 % de précision en entraînement et 98,25 % en test.
- Optimisation avec une perte croisée binaire, réduisant la perte à 0,0107 pour des prédictions fiables.

Segmentation client avec le clustering KMeans ([lien](#))

Déc 2024

- Réalisation de l'analyse exploratoire des données sur 525 461 transactions, avec une réduction du bruit de 23%.
- Application de l'algorithme de clustering KMeans sur trois variables clés, identifiant avec précision quatre groupes de clients.
- Développement de stratégies de fidélisation en analysant le comportement des acheteurs fidèles et à forte valeur.

Analyse de données de bout en bout avec Python et SQL ([lien](#))

Déc 2024

- Standardisation d'un ensemble de données de détail grâce à un processus de nettoyage des données piloté par Python.
- Exécution de requêtes SQL pour analyser les ventes et la performance des produits, appuyant les décisions stratégiques.

COMPÉTENCES

Programmation & Bibliothèques

PYTHON (NUMPY, SCIKIT-LEARN, PANDAS, MATPLOTLIB, PYTORCH), C, FORTRAN 90, MYSQL

Outils de Données & Visualisation

POWERBI, TABLEAU, Microsoft SQL Server

Systèmes d'Exploitation / Communication

Linux, Windows, MacOS / Anglais (C2), Français (B1)

EDUCATION

Master 2 : Astrophysique et Science des Données | Université de Strasbourg, France

Sep 2023 — Juillet 2024

Cours : Gestion de Bases de Données avec SQL, Machine Learning et IA, Probabilité Bayésienne, Analyse de Données (traitement d'enquêtes astronomiques à grande échelle), Programmation Scientifique, Statistiques, Calcul Parallèle en Astrophysique)

Master 1 : Paris Physics Master | Université Paris Cité, France

Sep 2022 — Juillet 2023

Cours : Méthodes Numériques - Interpolation, Extrapolation, Recherche de Racines, Stabilité Numérique, Équations Différentielles Ordinaires, Équations aux Dérivées Partielles, Optimisation, Minimisation, Algèbre Linéaire, Inversion de Matrices

Licence : Physique | Tezpur University, Inde

Sep 2019 — Juillet 2022

Cours : Introduction à la Programmation en Python, Méthodes Numériques pour le Calcul Scientifique, Algèbre Linéaire, Calcul

FORMATIONS ACADÉMIQUES ET ATELIERS

École d'Été : New Trends in Computing 2024, IRMA, Université de Strasbourg

Août 2024

- Sujets : calcul haute performance, machines hétérogènes, méthodes numériques parallèles.

DISTINCTIONS

Bourse MOBIL'ITI IRMIA++ par l'Université de Strasbourg

2023

Summer Research Fellowship par Indian Academy of Sciences

2021