Vishrut Bezbarua, M.Sc.

(+33) 783235990 | vishrutbezbarua@gmail.com | vishrut-bezbarua

Professional Experience

DATA SCIENTIST (INTERN) **Feb 2024 — July 2024**
Observatoire Astronomique de Strasbourg *Strasbourg, France*

- Processed 3D data cubes (20GB total) by trimming, smoothing and normalization, reducing data size by 50% and volume by 87%.
- Developed custom algorithms in PYTHON to automatically extract structures of interest (filaments) from the processed data.
- Engineered data pipelines to measure the median evolution of 6 properties of these filaments in 4 cosmic times (redshifts).
- Conducted statistical analysis and implemented quality control protocols, ensuring exclusion of edge regions and outliers to maintain data integrity.

SCIENTIFIC RESEARCHER (INTERN) **April 2023 — July 2023**
Laboratoire Physique et mécanique des milieux Hétérogènes *Paris, France*

- Developed a Fortran program for linear stability analysis of Boussinesq equations in a 2-D domain for Rayleigh-Bénard convection.
- Implemented LU Decomposition, reducing execution time of large matrix multiplications by approximately 30%.
- Accurately determined the 'Critical Rayleigh Number' within a 5% error margin by solving large-scale eigenvalue problems.

DATA ANALYST (INTERN) **May 2021 — July 2021**
Indian Institute of Astrophysics *Bangalore, India*

- Extracted and cleaned a 25-entry dataset (CSV) from astronomical catalogs, ensuring data consistency via cross-referencing.
- Utilized MS Excel and Python for preprocessing, parameter calculations (e.g., mass, ratios, luminosity), and data visualizations.
- Created multi-range diagrams and performed statistical comparisons to identify distinct patterns and group characteristics.

Relevant Projects

Deep Learning Project on Breast Cancer Classification (*GitHub*) **Nov 2024**

- Implemented a neural network in PyTorch to classify breast cancer tumors using the Wisconsin dataset (569 samples).
- Applied train/test splitting, feature standardization, and tuned hyperparameters with ADAM and BINARY CROSS-ENTROPY.
- Achieved 99.78% training accuracy and 98.25% test accuracy, demonstrating strong generalization performance.

Customer Segmentation with KMeans Clustering (*GitHub*) **Dec 2024**

- Conducted exploratory data analysis on 525,461 transactions, uncovering 23% missing customer IDs and negative quantities.
- Cleaned data by removing invalid entries and engineered features like monetary value, purchase frequency, and recency.
- Applied KMEANS CLUSTERING with these features using the Elbow and Silhouette methods, defining 4 customer clusters/groups.
- Designed personalized retention strategies by interpreting cluster characteristics to drive business decisions.

Retail Orders Data Cleaning and SQL Data Analytics (*GitHub*) **Nov 2024**

- Engineered a data cleaning process using PANDAS to standardize a retail dataset, improving data analysis readiness.
- Executed SQL queries to derive insights on sales trends and product performance, aiding strategic business decisions.

Skills

Languages & Libraries	PYTHON (NUMPY, SCIKIT-LEARN, PANDAS, SCIPY, MATPLOTLIB, SEABORN, PYTORCH), C, C++, MySQL
Data Tools & Visualization	PowerBI, Tableau, Microsoft SQL Server
OS / Communication	Linux, Windows, MacOS / English (C2), Français (B1)

Education

- **Master 2 : Astrophysics and Data Science** | ***University of Strasbourg, France*** **Sep 2023 — July 2024**
  Coursework : *Database Managemnent with SQL, Machine Learning and AI, Bayesian Probability, Data Analysis (large-scale astronomical survey processing), Scientific Programming, Statistics, Parallel Computing in Astrophysics)*
- **Master 1 : Paris Physics Master** | ***Université Paris Cité, France*** **Sep 2022 — July 2023**
  Coursework : *Numerical Methods - Interpolation, Extrapolation, Root Finding, Numerical Stability, Ordinary Differential Equations, Partial Differential Equations, Optimization, Minimization, Linear Algebra, Matrix Inversion, Monte Carlo Algorithms*
- **Bachelor of Science : Physics** | ***Tezpur University, India*** **Sep 2019 — July 2022**
  Coursework : *Introduction to Python programming, Numerical Methods for Scientific Computing, Linear Algebra, Calculus*

Additional Training

Summer School : New Trends in Computing 2024, *IRMA, Université de Strasbourg* **Aug 2024**

- Topics : High-performance computing, heterogeneous machines, parallel in time numerical schemes

Achievements

Summer Research Fellowship by the *Indian Academy of Sciences* **2021**
MOBIL'ITI IRMIA++ Scholarship by the Université de Strasbourg **2023**