Alok Kumar

alok85820018@gmail.com   +91 6376863079   [GitHub](#)   [LinkedIn](#)   [Leetcode](#)

Experience

Cognizant

Program analyst Trainee - AIA Data Integration   **Kolkata - July 2025 - Present**

- **Database Management** & **Query Optimization**: Managed relational databases using **SQL**, **MySQL**, and **Snowflake**; enhanced performance through **indexing** and **execution plan analysis**.
- **Cloud Integration** & **Data Warehousing**: Implemented scalable solutions with **AWS RDS**, **S3**, and **Redshift**; designed efficient **warehouse architectures** for historical data analysis.
- **ETL and Data Pipelines**: Built and automated **ETL workflows** to enable seamless **data transformation** and integration from multiple sources into cloud-based systems.

Cognizant

Database intern - AIA   **Chennai - March 2025 – July 2025**

- **Database Management** & **Query Optimization**: Collaborated on building and refining **relational databases** using **SQL**, **MySQL**, and **Snowflake**, focusing on **data accuracy** and optimizing **query structures** for better performance.
- **Cloud Integration** & **Data Warehousing**: Contributed to setting up **cloud-hosted environments** with **AWS tools** like **RDS**, **S3**, and **Redshift**, and assisted in modeling **data warehouse schemas** for business intelligence.
- **ETL and Data Pipelines**: Helped design and monitor **ETL processes** to automate **data ingestion** and transformation, ensuring reliable integration of structured data into centralized repositories.

Databits Technologia

Machine Learning and Cloud Engineer Intern   **Kolkata – Aug. 2024 – Oct. 2024**

- Built and optimized **machine learning models** using **Python**, **Scikit-learn**, and **TensorFlow** for **predictive analytics** and **classification**, ensuring high accuracy across diverse datasets.
- Deployed models via **REST APIs** on **AWS** (**EC2**, **S3**, **Lambda**), enabling **scalable**, **real-time inference** and automating **cloud-based workflows**.
- Engineered end-to-end **training pipelines**, including **data preprocessing**, **feature engineering**, and integrated **monitoring tools** to detect **model drift** and maintain performance.

Project Contribution

1.Disease Prediction – Liver Cirrhosis Detection ([LINK](#))

- Developed **supervised ML models** in **Python** using **Scikit-learn** and **Pandas** to predict **liver cirrhosis**, achieving **92% accuracy** through robust **data preprocessing** and **model evaluation** (e.g., **ROC-AUC, cross-validation**).
- Deployed the final model via a **Flask web app**; utilized **Git**, **GitHub**, and **Jupyter Notebook** for collaborative, version-controlled development. [**GitHub Link**]

2. Shopping Karo – Frontend Web Application ([LINK](#))

- Built a responsive **e-commerce frontend** using **HTML**, **CSS**, **JavaScript**, and **React.js**, featuring an intuitive **UI/UX**, **product listing**, and **shopping cart functionality**. Implemented **state management** with **React Hooks** and **Context API**; used **Axios** for API integration and ensured **mobile-first** design compatibility.
- Applied **component-based architecture**, optimized loading performance, and hosted the project on **GitHub** with clean, modular code. [**GitHub Link**]

Skills

Languages: Python, C++, SQL, JavaScript, HTML, CSS

Technical skills: Data Structures & Algorithms (DSA), Exploratory Data Analysis (EDA), Data Visualization, Data Preprocessing & Integration, Process Automation, KPI Tracking, Predictive Analytics, Supervised & Unsupervised Learning, Generative AI, Natural Language Processing (NLP), Transformer Models, Flask, Statistics, Linear & Logistic Regression, Normalization.

- **Machine Learning**: SHAP, LIME, KNN, K-Means, Decision Trees
- **Deep Learning**: Neural Networks, LLMs, GNN, CNN
- **Time Series Forecasting**: LSTM, ARIMA
- **Cloud and Big Data**: AWS, Azure, Snowflake, Teradata, PySpark, Google Cloud Platform (GCP)

Tools and Libraries: Power BI, Power BI Service, Power Query, Excel (VBA, Macros), Azure ML, SQL, PyTorch, TensorFlow, Keras, OpenAI API, Scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, YFinance, Google Analytics, MLflow, Git, GitHub, DVC, Microsoft Office.

Education

Haldia Institute of technology   **Haldia - 2021-2025**

*B Tech in Computer Science (**CGPA: 8.31)***

Happy Valley School   **Bhagalpur 2018-2020**

Senior Secondary *School PCM (**CGPA: 7.2)***

Achievement & Certificate

Member of Data Science Club at HIT

National Means-cum-Merit Scholarship Holder

Python Programming Masterclass - Udemy

Machine Learning A-Z - Udemy