AYUSH RAI

Portfolio | Ayush-31r | ayush-rai-v1 | rai4ayush@gmail.com | +91 63883 61301

Education

Indian Institute of Information Technology Manipur **Nov 2022 – May 2026**
Bachelor of Technology in **Electronics and Communication Engineering** **CGPA:** 7.87/10

Experience

Data Science Intern, Securify **Mar 2025 – Jul 2025**

- Designed a data-isolated, **Multi-tenant data architecture** to securely aggregate and analyze service-usage data across **6+** organizational roles.
- Built automated **Data ingestion and transformation pipelines** using **Celery**, **Flower**, and **Go** to collect third-party integration and user-activity data into schema-specific analytical databases.
- Performed **exploratory data analysis (EDA)** and implemented anomaly-detection logic on API usage and background job metrics to identify abnormal spikes, failures, and cost-risk patterns.

Projects

InsightOps – Support Ticket Risk Analysis System **Dec 2025**

- Developed an end-to-end NLP system to identify **high-risk customer support tickets**, enabling faster prioritization of critical issues using **20K+ historical tickets**.
- Trained and evaluated supervised models (**Logistic Regression**) achieving **0.89 ROC–AUC** and **68% F1-score** on a held-out validation set.
- Implemented a **retrieval-augmented explanation layer (RAG)** to surface similar past cases and policy context, reducing simulated manual triage time by **32%**.

Parkinson's Disease Detector **Oct 2025**

- Built a **CNN-based model in PyTorch** to detect Parkinson's Disease from patient speech using the **MDVR-KCL dataset (2K+ samples)**.
- Processed raw audio into **spectrograms and MFCC features** with normalization and augmentation to enhance robustness on noisy biomedical data.
- Achieved **90.2% accuracy** and an **F1-score of 0.90**, deploying an interactive **Streamlit demo** showcasing real-time voice-based disease detection.

FlowCast – Urban Mobility Forecasting **Aug 2025**

- Predicted **average inter-ward travel time and fare** in Bangalore using **Uber Movement data** and ward-level **GeoJSON** boundaries with spatial–temporal feature engineering.
- Engineered features such as **Haversine distance**, hour of day, and rush-hour indicators; visualized congestion trends via **Plotly choropleths**.
- Trained and deployed a **Random Forest model** achieving **MAE ≈ 485 s** and $\mathbf{R^2 = 0.67}$, with distance contributing **76%** to variance; hosted interactive predictions via **Streamlit**.

Technical Skills

Data Science & Analytics:	**Python**, **SQL**, **Pandas**, **NumPy**, **EDA**, **Feature Engineering**, **Statistics**, **Model Evaluation**
Machine Learning:	**scikit-learn**, **PyTorch**, **TensorFlow**, **Classification**, **Regression**, **Natural Language Processing**
Data Visualization:	**Matplotlib**, **Seaborn**, **Plotly**, **Tableau**
GenAI & NLP:	**TF-IDF**, **Embeddings**, **RAG**, **Vector Databases**, **LangChain**
Tools & Cloud:	**Git**, **Docker**, **GCP**, **AWS**

Achievements

- Achieved **3-Star CodeChef rating (1733 peak)** with a **global contest rank of 164**.
- **Winner, CodeRush 2024** – secured **1st place** in Competitive Coding.
- **Champion, 3 CTF competitions** hosted by the Data Security Club, outperforming **1000+ participants**.
- More **than 50 Badges** collected by leveraging GCP services in **the Google Cloud Skill Boost Program**.