# AYUSH RAI

Data Scientist | ML Engineer | Backend Developer

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

## Education

**Indian Institute of Information Technology Manipur** — **Imphal, Manipur**

Bachelor of Technology in **Electronics and Communication Engineering** — Nov 2022 – May 2026

## Experience

**Software Developer (Data & Backend)** — **Mar 2025 – Jul 2025**

Securify

- Designed a data-isolated, **multi-tenant system** to ensure **secure data handling** and access control for **6+** roles.
- Built automated **data processing pipelines** using **Celery**, **Flower**, and **Go**, improving the reliability of background **data management** and system tasks.
- **Analyzed and optimized database query performance** for **50+ REST APIs**, reducing query response times by **10%** and improving access to critical data.

**Google Cloud Boost Skill Program Participant** — **2024**

Google Arcade

- Leveraged **GCP services** including Compute Engine, Cloud Storage, and Cloud Run to deploy and manage scalable, secure, and reliable applications, demonstrating proficiency in cloud technologies and earning over 30 badges.

## Projects

**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**.

**TruthLens – Fake News Classifier** — **Ongoing**

- Designed a reproducible NLP pipeline for **fake news detection** using **TF–IDF feature extraction** and supervised models (**Naive Bayes, Logistic Regression, Random Forest, SVM**).
- Performed **EDA, text preprocessing, and vectorization**, evaluating models on a dataset of **23K+ news articles** with precision–recall and confusion matrix analysis.
- Aquiring **93% accuracy** and **F1-score 0.93**, with **Decision Tree and SVM** performing best; provided a baseline framework for text classification research.

## Technical Skills

| | |
|---|---|
| **Data Science & ML:** | **Python**, **SQL**, **Pandas**, **NumPy**, **scikit-learn** |
| **Frameworks:** | **PyTorch**, **TensorFlow**, **Django**, **Fast API** |
| **Cloud & Platforms:** | **GCP**, **AWS**, **Docker**, **GitHub** |
| **Other Skills:** | **Go**, **C++**, **JavaScript**, **React**, **Tableau** |

## Achievements

- Solved **300+ coding problems** across Codeforces, LeetCode, and CodeChef.
- Achieved **3-Star CodeChef rating (1733 peak)** with a **global contest rank of 164**.
- **Winner, CodeRush 2024** – secured **1st place** in Competitive Coding (Non-CS category).
- **Champion, 3 CTF competitions** hosted by the Data Security Club, outperforming **1000+ participants**.