

RITIK RANJAN

Mumbai, Maharashtra, India

📞 +91-9430036363 ✉ ritikranjan1212@gmail.com 🔗 linkedin.com/in/ritik-ranjan9900 🌐 github.com/Ritik9900 🏆 Kaggle.com/ritik9929

Education

Indian Institute of Technology, Madras

Bachelor of Science (BS) in Data Science & Programming, with Minors in Economics & Finance

May 2021 – April 2025

Chennai, India

Experience

Gmonks

October 2024 – Present

AI/ML Engineer

Mumbai, India

- Implemented a computer vision solution using a **YOLOv8** model, achieving a **0.339 mAP** for facial blemish detection and developing custom **OpenCV** algorithms for automated image retouching.
- Deployed the computer vision model as a scalable inference API on **AWS ECS Fargate**, a strategic decision for remote processing that reduced client application size by over **80 MB**.
- Engineered a secure **FastAPI** backend with **10+** endpoints to support the ML service, generating ephemeral signed URLs with **60-second** validity to authorize media asset access.
- Automated end-to-end infrastructure provisioning for both the ML service and backend API using **Terraform**, ensuring **100%** reproducible cloud environments and reducing setup time by **95%**.
- Orchestrated the full **MLOps** lifecycle for a system serving thousands of daily inference requests, from container versioning in **AWS ECR** to integrating the backend's security layer.

Projects

IndicTTS Deepfake Challenge | Kaggle

January 2025 – February 2025

- Placed Top **4.5% (9/201)** in Kaggle IndicTTS Deepfake Challenge by designing a multilingual deepfake speech detection system.
- Developed a data processing pipeline for **33K+** audio clips (**63 hrs**) across 16 Indian languages, enabling analysis with a ROC-AUC of **0.923**.
- Trained a ViT-based PyTorch model on Mel spectrograms with mixed-precision and multi-GPU setup, improving accuracy by **12%** and reducing epoch time by **40%**.

Denoising & 4x Super-Resolution of Low-Light Images | Kaggle

March 2025 – April 2025

- Secured a Top **6.2% position (13/210)** in the Kaggle Super-Resolution Challenge by enhancing low-light image quality.
- Optimized and fine-tuned PSNR-focused models to substantially enhance visual quality on a challenging low-light image dataset.
- Built a PyTorch pipeline with MRPNet for denoising and pixelShuffle Generator for **4x upscaling**, trained over **70 epochs** using mixed-precision and multi-GPU setup.

Big Data Engineering | GCP | Apache Spark

Jan 2025 – April 2025

- Engineered batch data pipelines on a Google Dataproc cluster, implementing complex **SCD Type II** data warehousing logic (PySpark & SparkSQL) and aggregating user click data into 4 distinct time-based buckets using hashing.
- Architected two real-time streaming systems: (1) a serverless, event-driven pipeline using **GCP Cloud Functions** and **Pub/Sub**, and (2) a high-throughput system using **Kafka** and **Spark Streaming** to process **1000+** records with a 5-second micro-batch interval and a 10-second processing window.
- Optimized a Spark **MLib** Decision Tree model by integrating the CrossValidator autotuner for automated hyperparameter selection, and successfully re-engineered a batch image classification model into a real-time inference pipeline using Spark Streaming.

Sentiment Prediction on Movie Reviews | Scikit-learn | NumPy | Pandas | SciPy

May 2023 – August 2023

- Ranked Top **8% (79/977)** in a Kaggle sentiment analysis competition with an F1 micro score of **83.20%**.
- Crafted a scalable classification pipeline using XGBoost and Scikit-learn on **50K+ labeled movie reviews**.
- Boosted model accuracy by **10%** and reduced training time by **25%** through hyperparameter tuning and vectorization optimization.

Technical Skills

Languages: Python | JAVA | Bash Scripts | HTML/CSS | JavaScript | SQL | Machine Learning

IDE & Developer Tools: VS Code | Eclipse | Google Colaboratory | Jupyter Notebook | PyCharm

Technologies/Frameworks: Linux | GitHub | JUnit | Postman | Celery | Redis | Pandas | JIRA | ClickUp | PyTorch | Tensorflow

Big Data Tools: PySpark | Spark ML | Spark Streaming | Google Dataproc | Kafka | Google Pub/Sub

Cloud Platforms: Google Cloud Platform (GCP) | Google Cloud Functions | Amazon Web Services (AWS)

Certifications

- Emotional Intelligence | NPTEL | IIT Kharagpur (January 2024 – March 2024)
- Data Mining | NPTEL | IIT Kharagpur (January 2024 – March 2024)
- Cloud Computing | NPTEL | IIT Kharagpur (July 2024 – October 2024)
- Supervised Machine Learning | Stanford Online | Deeplearning AI (September 2024)

Leadership / Extracurricular

Academic Societies

July 2022 – Present

Volunteer

Indian Institute of Technology, Madras

- Collaborated in interdisciplinary research discussions and organized community-building events as part of Sundarbans 271 House & Ramanujan Society at IIT-M.