**Ranveer Singhania**

College Station, TX ● +1 9793265513 ● [rajpurohitharjun58@gmail.com](mailto:rajpurohithatjun58@gmail.com) ● raj-purohith-arjun.github.io/portfolio/ ●  [Linkedin](https://www.linkedin.com/in/raj-purohith-arjun-20a652200/)

**EXPERIENCE**

**Graduate Data Science Assistant** Dec.2024 - Present

*Texas A&M Institute of Data Science | College Station, TX*

* + Engineeredan **end-to-end modeling pipeline** for forecasting greenhouse gas emissions from BNI crop systems using 10,000+ IoT sensor timepoints, aligned with best practices across the Software Development Life Cycle (SDLC)
  + Achieved **30% higher accuracy** in GHG predictions, enabling analysis of climate-smart agriculture’s impact on soil health and carbon market eligibility through **business-driven analytics**
  + Automated **imputation and forecasting processes** by integrating BRITS and LSTM models, reducing RMSE and MAE errors by 15–20% over baseline models and improving **model stability**
  + Implemented robust **data preprocessing and feature engineering** across 17 environmental variables, reducing data preparation time by 40%
  + **Troubleshot performance bottlenecks** and optimized pipeline components using development tools such as Python, Git, and Jupyter; collaborated in an **inclusive, cross-functional team** to support real-time model deployment
  + Presented actionable insights via 5+ **interactive Tableau dashboards** and reports, influencing grant proposals and demonstrating the ROI of AI models in carbon offset initiatives

**Student Researcher, NLP & AI** Oct.2024 - Mar.2025

*FLAIR Lab, Texas A&M University | College Station, TX*

* + Improved entity recognition F1 by 25% on legal documents by training **BERT, T5, and GPT-2** models on OCR-simulated data using GPU infrastructure, reducing manual review time by 60%
  + Constructed a **multi-task BERT chatbot** for student services, improving intent and slot prediction accuracy by 20% over BiLSTM+CRF through T5-based data augmentation
  + Initiated lightweight inference-ready models via REST APIs using **ONNX and quantization techniques**, reducing latency by 18% and enabling scalable real-time predictions

**Machine Learning Engineer** Nov.2023 - Mar.2024

*Entropik | Chennai, India*

* + Engineered **facial emotion recognition** models using MobileNet and EfficientNet on **1,000+** hours of annotated video, improving accuracy by 21% via hard sample mining and robust noise-handling techniques
  + Built and deployed an **end-to-end inference pipeline** for real-time and batch processing using TensorFlow, OpenCV, and ONNX; ensured **application resiliency** and reduced inference latency under production loads
  + Integrated outputs into a **live dashboard** with role-based access controls, contributing to **secure and stable software deployment** that automated analytics for marketing teams and saved 120+ analyst hours/month
  + Collaborated across **Agile sprints** with **UX, DevOps, and product teams**, mapping emotion scores to campaign ROI metrics; this **business analysis** directly improved ad conversions by 17% and reduced feedback loops by 40%

**SKILLS**

Python, SQL, Git, Linux, Jupyter, Docker, Apache Spark, Hadoop, Airflow, AWS (S3, Lambda), GCP (BigQuery), REST APIs, ETL, CI/CD, SDLC, Agile, Secure Coding, Application Resiliency, Technical Troubleshooting, Development Toolsets, Scikit-learn, PyTorch, Pandas, NumPy, Data Structures & Algorithms, Data Analysis, Data Visualization, Natural Language Processing, Generative AI, LangChain, Neo4j, Business Analysis, RAG

**PROJECTS AND LEADERSHIP**

**Founding Engineer Project** | **AllyAI.net |** *SMARTLink – AI Graph Recommendation Engine*

* Designed a graph recommendation system using **Neo4j, OpenAI embeddings,** and **DuckDB** to connect users via career and academic link
* Launched a real-time interface using **Streamlit** and **LangChain**, achieved 33% higher match accuracy in pilot runs

**Lead Research Author | SRM University |** [*YOLOv8-Based ANPR*](https://link.springer.com/chapter/10.1007/978-3-031-82383-1_6) *– IRCCTSD Springer Nature*

* Modeled an ensemble ANPR system using **YOLOv8, and Deep Learning** Techniques to improve license plate detection and tracking in real-world low-light traffic environments. Achieved F1 Score of **0.97** through custom preprocessing, augmentation, and OCR pipelines

**Founder & Business Lead | GanfinityAI** *| ChatBot for SIMS Hospital, Chennai*

* Developed and deployed a **Rasa-based chatbot** at SIMS Hospital, automating 85% of **routine queries** and reducing front-desk load by 40%
* Scaled to **5,000+** users within a year, cut scheduling time by 50%, and secured $5,000 in funding by aligning features with clinical workflows in cross-functional teams

**Data Science Analyst Intern | *HighRadius*** *| AI-enabled Fintech B2B Cloud Financial Fraud Detection Application*

* Trained interpretable **fraud detection models** using **XGBoost and SHAP** on B2B transaction data, improving fraud precision by **30%** and reducing false positives. Optimized **SQL** pipelines on GCP, improving data retrieval by 40%, and collaborated with compliance teams to embed explainability into daily reporting workflows

**EDUCATION AND HONORS**

**Texas A&M University** | Texas A&M Institute of Data Science, College Station, TX Aug.2024 - May.2026

Master’s of Science in Data Science ( Computer Science Track) | GPA: 3.45

*Coursework: Deep Learning, Maths & Statistics for Data Science, Big Data Tools, Data Mining, Natural Language Processing, Software Engineering*

**SRM University** | SRM Institute of Science and Technology, India Oct.2020 - May.2024

Bachelor of Technology with Honors in Computer Science and Engineering ( AI & ML) | GPA: 3.92

*Achieved 1st rank in the department and was included in the Dean’s List in 2023 and 2024*