Anuj Patel

+1 682 217-2520 | E-Mail: **anuj.patel.29dec@gmail.com** | GitHub | LinkedIn | Portfolio

Machine Learning and Data Science graduate student with experience in large-scale data analysis, predictive modeling, and statistical experimentation. Proficient in Python (Pandas, NumPy, Scikit-learn) and SQL for data preprocessing, exploratory data analysis (EDA), feature engineering, and machine learning model development. Experienced in building end-to-end data pipelines, evaluating models using statistical metrics, and translating complex datasets into actionable insights through visualization and reporting to support data-driven decision making.

EXPERIENCE

University of Texas at Arlington Feb 2026 - Present

Research Assistant (Physics-Informed Neural Networks)

- Implemented a **Physics-Informed Neural Network (PINN)** using **Python, TensorFlow/PyTorch** to model soil water retention curves across **1,000+ soil samples**, performing **feature engineering on particle size distribution (PSD) and soil property datasets**
- Integrated **physics-based constraints into the loss function**, reducing prediction error by **15–20%** and improving model generalization over baseline models
- Conducted **statistical analysis, hypothesis testing, and error analysis on 1,400+ samples**, improving model evaluation robustness and validating generalization across diverse soil distributions

RinaySoft, Pune, India June 2024 - June 2025

Software Engineer (Applied AI/ML)

- Engineered an **LLM-powered email automation system using LangChain, OpenAI API, and FastAPI**, automating classification, summarization, and response generation, reducing manual processing effort by **95%**
- Orchestrated **multi-agent LLM workflows using LangChain** for email routing and prioritization, reducing manual triaging effort by **~60%** and improving task handling efficiency across pipelines
- Constructed a semantic search pipeline using HuggingFace embeddings and **ChromaDB**, indexing **10K+** documents to enable efficient **similarity search** and improve information **retrieval accuracy**

KEY PROJECTS

CHURNSENSE | GitHub

- Architected an **end-to-end ML pipeline (data preprocessing/modeling/evaluation/deployment simulation)** on **10K+ records**, improving model scalability and enabling efficient monitoring of churn predictions
- Prepared a churn prediction model on **10K+** records, boosting early detection by **25%** via **RFM + behavioral features**.
- Enhanced ML models using k-fold **CV, ROC-AUC**, and recall, ensuring high-impact business outcomes.
- Built interactive dashboards for churn analysis, enabling data-driven customer retention strategies, improving decision-making efficiency

INCREMENTAL MODEL |GitHub

- Applied **clustering + PCA** on **5K+** records, boosting pattern detection accuracy by **25%**.
- Developed **LLM-powered, time-decayed topic modeling pipeline** with **<5 min latency** for real-time insights.
- Improved trend detection speed by **35%**, enabling faster, data-driven decision-making.

AI/ML LAB | GitHub

- Initiated a **modular AI platform** integrating embeddings, **RAG**, and **multi-agent** systems
- Reduced insight generation latency to **<5 minutes** using optimized pipelines
- Structured a serverless architecture using AWS Lambda and Step Functions to automate data ingestion and model retraining pipelines, decreasing manual intervention by **75%** and infrastructure costs by **15%.**

SKILLS

- **Languages:** **Python, SQL, C++, R**
- **ML/AI:** **Machine Learning, Deep Learning, NLP, LLMs, RAG, Multi-Agent Systems, Generative AI**
- **Frameworks:** **Scikit-learn, TensorFlow, PyTorch**
- **Data:** **Pandas, NumPy, Feature Engineering, EDA, Statistical Modeling**
- **Systems & MLOps:** **Model Deployment, Pipeline Design, Model Evaluation, Monitoring**
- **Cloud & Tools:** **AWS (EC2, S3, Lambda), Docker, CI/CD (GitHub Actions), MySQL**
- **LLM Tools:** **LangChain, FAISS, Pinecone, OpenAI API**

PUBLICATION & CERTIFICATIONS

- **AWS Certified Generative AI Developer - Professional | BADGE** Feb 2026
- **NVIDIA Certified Professional: Agentic AI | BADGE** Jan 2026
- **A Study on the Potential of AI in the Healthcare Sector - Research Paper** Feb 2024

EDUCATION

University of Texas at Arlington, Arlington, Texas, U.S Aug 2025 - Present
Master's of Science in Computer Science (Machine Learning & Data Science Focus)

Prestige (PIEMR), RGPV, Indore, MP, India Aug 2021 - May 2025
Bachelor of Technology (Computer Science & Engineering)