

Ansh Kumar Dev

520-599-5589 | anshkumardev@gmail.com | linkedin.com/in/ansh-kumar-dev/ | github.com/Anshkumardev

EDUCATION

University of Arizona

Master of Science in Data Science

Tucson, AZ

Aug. 2023 – May 2025

Gautam Buddha University

B.Tech. in Artificial Intelligence

Greater Noida, India

Aug. 2019 – May 2023

EXPERIENCE

AI Engineer Intern

Right Skale

Jun 2025 – Present

Pleasanton, CA

- Design and deploy **agentic LLM pipelines** for processing **SEC 10-Q filings** using **LlamaParse**, **LlamaIndex**, and **LangChain agents**, enabling contextual reasoning over financial and unstructured tabular data.
- Collaborate with **finance SMEs** and **clinical-style QA teams** to ensure safe, factual, and empathetic agent responses with retrieval-grounded validation.
- Implement a production-grade **RAG system** using **OpenSearch Serverless**, **PostgreSQL**, and **Titan embeddings**, integrating **1M+ parsed JSON** chunks for **sub-second semantic search**.
- Build **prompt libraries** and tool schemas for multi-modal voice/chat flows in **AWS AI Foundry**, supporting dynamic agent behavior and error recovery routines.
- Deploy CI/CD pipelines with **GitHub Actions**, monitor via **CloudWatch Logs** and **X-Ray**, and enforce security using **IAM roles** and **SOX-compliant standards**.

Data Scientist (Graduate Research Assistant)

University of Arizona

Aug 2024 – Jan 2025

Tucson, AZ

- Built a geospatial **ETL pipeline** using **Python**, **SQL**, and **Pandas** to integrate and clean satellite and sensor data for hydrologic modeling.
- Engineered time-series features for a CV-RNN model, improving 12-month aquifer forecasting RMSE by **18%**.
- Led end-to-end data science workflow: **data wrangling**, **EDA**, and **ML model evaluation** using **Bayesian tuning** and **statistical testing**.
- Mentored 15+ students in best practices for reproducible research, **model design**, and **experimental interpretation**.

AI Engineer Intern

Merkletree Technologies

Apr 2022 – Jul 2022

Delhi, India

- Trained **YOLOv5** and **Faster R-CNN** models for biometric facial recognition, boosting face-auth accuracy **25%** via targeted augmentation.
- Optimized inference pipelines using **CUDA** on **embedded Linux**, reducing latency **35%** for real-time on-device verification in apps used by **50k+ users**.
- Built reproducible **CI/CD pipelines** using **Docker**, **Kubernetes**, **Flask APIs**, and **MLflow**, enabling seamless model deployment across environments.

PROJECTS

Healthcare Accessibility & Risk Mapping Platform | [Project Link](#)

Apr 2025

- Built a distributed **GeoPandas + OSMnx** pipeline unifying Geospatial, socioeconomic & clinical data across 3,000+ U.S. census tracts to identify underserved areas for policy action.
- Developed a **LangChain** multi-agent LLM workflow with **OpenAI API** and **FAISS RAG** and a **Streamlit** choropleth dashboard, enabling real-time Q&A and evidence-based insights.

E-commerce Analytics Lakehouse | [Project Link](#)

Mar 2025

- Engineered an **Azure Data Factory to Databricks (PySpark)** ETL ingesting 1.2 TB/day of order & clickstream data to power near-real-time KPI dashboards.
- Modeled star-schema views in **Synapse SQL** and built live **Power BI** reports, Reduce manual reporting by 90%

TECHNICAL SKILLS

Languages/ML: Python (Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, OpenCV), R, SQL, C++

Models: RoBERTa, SVM, Decision Trees, K-Means, PCA, Forecasting & Statistical Analysis

Cloud & MLOps: AWS (EKS, EC2, S3), Azure (ADF, Synapse, Databricks), GCP, Docker, Kubernetes, Spark, CI/CD, MLflow

Tools & APIs: Flask, REST APIs, TensorBoard, Git, Linux, Power BI, Streamlit