

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

Jun 2025 – Present

Right Skale

Pleasanton, CA

- Design, fine-tune, and deploy production-grade **ML models**, including **LLMs** and **retrieval-augmented generation (RAG)** pipelines for enterprise SaaS applications in the life sciences domain.
- Collaborate with CEO, product team, and cross-functional engineers to develop **AI-driven solutions**, integrating **LangChain** and **Bedrock LLMs** within AWS-based microservices.
- Conduct exploratory **data analysis** and preprocessing of structured and unstructured documents using **Python**, **SQL**, and **Power BI**; automate reporting, streamline audits, and support agile planning sessions.
- Drive experimentation with new models and frameworks to accelerate delivery of AI features; proactively pilot ideas without red tape, reporting findings directly to leadership.
- Support regulatory compliance and **internal audits** by documenting ML methodologies, quality control procedures, and performance validation protocols.
- Build CI/CD pipelines and observability infrastructure using **GitHub Actions**, **CloudWatch**, and **IAM**, ensuring **secure, scalable deployments**.

Data Scientist (Graduate Research Assistant)

Aug 2024 – Jan 2025

University of Arizona

Tucson, AZ

- Built scalable **ETL pipelines** in **Python (Pandas, SQL)** to collect and preprocess geospatial and sensor datasets for time-series modeling, improving forecast accuracy by **18%**.
- Visualized and interpreted model outputs using **Power BI** and **Jupyter**, guiding academic research while mentoring **15+ students** in data science workflows.
- Performed **critical data analysis**, **model validation**, and **statistical optimization** to deliver reproducible insights and recommendations for water resource management.
- Balanced academic research deliverables with exploratory analysis and communication of findings to interdisciplinary teams.

AI Engineer Internship

Apr 2022 – Jul 2022

Merkletree Technologies

Delhi, India

- Developed and tuned computer vision models (**TensorFlow Faster R-CNN**) for facial authentication; enhanced accuracy by **25%** using advanced data augmentation and preprocessing.
- Implemented **Dockerized CI/CD pipelines** and deployed models in **Kubernetes** environments optimized for **Linux**; reduced inference latency **35%**.
- Built **Flask APIs** for serving models, collaborated with mobile developers to enable seamless integration of **on-device inference** for **50k+ users**.
- Contributed to testing and monitoring frameworks that supported rapid model updates and production reliability.
- Demonstrated adaptability in responding to evolving system requirements and aligning outputs to meet stakeholder goals.

PROJECTS

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

Apr 2025

- Built a **GeoPandas + OSMnx** pipeline merging U.S. geospatial & clinical data to flag underserved areas.
- Developed **LangChain** LLM agents with **FAISS RAG** & **Streamlit** dashboard for real-time Q&A.

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

Mar 2025

- Built **ADF → Databricks (PySpark)** ETL for 1.2 TB/day clickstream data; enabled live KPIs.
- Modeled star-schema in **Synapse SQL** & automated reporting with **Power BI**.

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