Ansh Kumar Dev

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

Education

University of Arizona — Tucson, AZ
Master of Science in Data Science — *Aug. 2023 – May 2025*

Gautam Buddha University — Greater Noida, India
B.Tech. in Artificial Intelligence — *Aug. 2019 – May 2023*

Experience

Machine Learning Engineer (Graduate Research Assistant) — Aug 2024 – Jan 2025
University of Arizona — *Tucson, AZ*

- Developed and deployed scalable ML pipelines using **Python, Pandas, and SQL**, reducing aquifer prediction error by **18%**.
- Led optimization of time-series features and model architecture (**CV–RNN**) for geospatial forecasting.
- Applied **Bayesian tuning** and statistical significance testing to evaluate model generalizability and robustness.
- Managed **ML workflows and retraining protocols** using reproducible methods to ensure operational reliability.
- Mentored peers in **ML model evaluation**, **data preprocessing**, and **CI/CD** best practices.

Machine Learning Research Intern — Jun 2022 – Nov 2022
Gautam Buddha University — *Greater Noida, India*

- Built anomaly detection models for IoT systems using **time-series analysis, C++**, and **Python**, improving predictive accuracy by **40%**.
- Streamlined feature engineering and pipeline design for real-time ML deployment in sensor environments.
- Published work on scalable **data streaming** architectures and textbfstatistical reliability frameworks.
- Optimized ingestion and processing pipelines to enhance system throughput by **20%**.

AI Engineer Intern — Apr 2022 – Jul 2022
Merkletree Technologies — *Delhi, India*

- Productionized computer vision model using **TensorFlow**, **Docker**, and **Kubernetes**, cutting latency **35%**.
- Implemented containerized CI/CD workflows for ML deployment via **MLflow** and **Flask REST APIs**.
- Worked on real-time model inference and integration into mobile platforms for 50k+ user base.
- Contributed to end-to-end **MLOps** systems, including model monitoring and versioning.

MLOps Intern — May 2021 – Aug 2021
LinuxWorld Informatics Pvt Ltd — *Jaipur, India*

- Built fully reproducible ML pipelines on **AWS**, with CI/CD and automated retraining mechanisms.
- Improved F1 score by **15%** through **Bayesian hyperparameter tuning** and domain-driven feature extraction.
- Developed ML APIs using **Flask** for real-time mobile app integration and analytics.
- Assisted in infrastructure setup and pipeline validation for scalable ML deployment workflows.

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.

Multi-Label Emotion Classification with RoBERTa | *Project Link* — Jan 2025

- Fine-tuned **RoBERTa-base** on 28 K emotion-tagged texts, boosting micro-F1 to 90% (+12%) for high-precision multi-label detection.
- Tracked experiments in **MLflow** and applied Bayesian hyperparameter search with mixed-precision GPU, cutting compute time by 40%.

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