Simran Jadhav

simranjadhavv022003@gmail.com | https://www.linkedin.com/in/simran-jadhav20/ | https://github.com/Simrann020

EDUCATION

University Of Colorado, Boulder — Aug 2024 – May 2026 | Boulder, USA
Master's in Data Science (Current GPA: 3.8/4)

Mumbai University, Rajiv Gandhi Institute of Technology — Jul 2020 – May 2024 | Mumbai, India
Bachelor's of Engineering in Computer Engineering (GPA: 3.7/4)

SKILLS

Programming & Databases: Python, R, SQL, C++, Java, MySQL, PostgreSQL, MongoDB.
ML & Big Data: pandas, NumPy, scikit-learn, TensorFlow, PyTorch, XGBoost, LightGBM, Spark, Hadoop, Dask, xarray.
Cloud & MLOps: AWS, Azure, GCP; Docker, Airflow, MLflow; Flask, FastAPI, REST APIs; CI/CD (GitHub Actions).
Data Engineering: ETL/ELT, time-series feature engineering, schema/QA validation, Parquet, orchestration.
Visualization: Tableau, Power BI, Matplotlib, Seaborn; dashboard design.
Professional: Leadership, Problem Solving, Communication, Cross-functional collaboration.

EXPERIENCE

Nexus Weather and Climate - Data Science & Engineering Intern — May 2025 – Present Boulder, USA

- Engineered production ETL and ELT for NWP, station, and wave and precip feeds; produced partitioned Parquet datasets; scheduled Airflow on Docker runs for daily refresh and backfills across multi-million record climate tables.
- Implemented a data quality and inference gate with schema, unit, range, and shape checks plus dedupe and backfills; replaced a legacy SVM with a tuned Random Forest; standardized metrics MAE, MSE, CRPS with automated evaluation reports and plots.
- Designed a CNN–LSTM forecasting module with probabilistic calibration and custom loss functions such as CRPS style and quantile weighted objectives to improve calibration and tail behavior; automated training and evaluation with YAML configs; exported single file artifacts and a REST service; shipped dashboards for runtime health and drift.
- Established CI and CD with GitHub Actions for nightly training and evaluation, artifact versioning and checksum tracking, and failure notifications; cut p95 inference latency under 350 ms by streamlining feature pipelines and caching.

Geomicrobial Physiology Lab at CU Boulder - Graduate Research Assistant, Isoverse — Dec 2025 – Present | Boulder, USA

- Reverse engineered legacy vendor binaries that follow MFC CArchive Serialize and mirrored exact field order with inheritance; produced deterministic R readers with schema stable outputs for three classes and completed serialize analysis for eight additional classes.
- Introduced a per class CSV schema with validation gates including unit checks, range checks, array count guards, and clear error messages; added strict and lenient modes ,so older files keep working while corruption is surfaced.
- Built a minimal test suite with tiny binary fixtures, golden outputs in RDS files, and regression checks for type stability across versions; added crafted corrupted samples to confirm fail fast behavior.
- In progress establishing an Isoverse server stack with Nginx in front of ShinyProxy and Keycloak using OpenID Connect; enforcing authentication at the edge, forwarding trusted identity headers to ShinyProxy, and blocking any direct path to app containers through network policy.
- In progress hardening the platform with TLS, HSTS, secure cookies, idle and absolute session lifetimes, reliable logout that invalidates tokens, and timeouts aligned with long-lived Shiny sessions; documenting request flow and required headers in a runbook.

PROJECTS

Safer Ride - Risk Aware Bike Routing | *AWS* — Sep 2025 – Nov 2025

- Built AWS ingest and spatial store with S3, Glue, Athena, and RDS PostGIS; partitioned crash and hazard data,and rendered hotspot overlays on Amplify.
- Lambda FastAPI computes a route risk score from crash density, severity, and hazard proximity.

AI Powered Code Debugger | *Python, Flask, GPT 4o, Docker, REST API* — Mar 2025

- Flask app with GPT 4o analyzes Python files, and logic faults, suggests safe refactors, and stages changes through linting and tests.
- Delivered a Docker image with documented REST API, token auth, structured logs, and run history so teams reproduce fixes

PCOS Analysis and Detection | *Python, Scikit-learn* — Jan 2025 – Mar 2025

- Built pipeline with cleaning, encoding, and class balance handling; trained cross validated models and tuned thresholds to favor recall
- Added a validated input form with calibrated probabilities, top drivers, and concise summaries that support quick review by

Data Integration for Denver International Airport | *ServiceNow, Azure DevOps* — Nov 2024

- Automated extracts from ServiceNow and Azure DevOps, reconciled schemas, and wrote quality rules for stale tickets, ownership
- Delivered an incident dashboard that shows live queues, bottlenecks, and trends by team, improving triage speed and visibility across.

Yoga Posture Detection and Correction | *Python, TensorFlow JS* — Aug 2023 – May 2024

- Trained a feed forward pose model with curated datasets and augmentations, achieving 96.5 percent accuracy under varied angle and
- Built a TensorFlow JS coach that tracks joints in the browser, gives posture feedback, and logs sessions for progress review.

Amazon Prime Data Analysis using Tableau | *Tableau, SQL* — May 2022

- Designed a SQL backed Tableau model with dimensions for title, genre, region, and release window, refreshed on a regular cadence.
- Built views for top performers, rising titles, and long tail segments, with filters, bookmarks, and a summary page for weekly reviews.

PUBLICATION

- Published a technological paper on Yoga Posture Detection and Correction. Link