Mihir D. Chauhan

mihirchauhan951@gmail.com | (720)-813-5491 |LinkedIn | Github | Portfolio

EDUCATION

University of Colorado at Boulder **Boulder, Colorado**
M.S. in Data Science, **GPA:** 3.96/4.00 *August 2024 - May 2026*
- o **Related Coursework:** Data Mining, Statistical Methods and Applications I & II, Machine learning , Natural Language Processing, Data scale systems, Big data, Machine Learning.

University of Mumbai **Mumbai, India**
B.E in Computer Engineering, **GPA:** 9.29/10.00 *August 2020 - May 2024*
- o **Related Coursework:** Python programming, Quantitative analysis, Cloud Computing, Object-Oriented Programming, Data Structures & Algorithms, Database Management Systems, Computer Networks.

SKILLS

Programming: Python, SQL, R, C++, Java, JavaScript
Scientific Python: NumPy, Pandas, SciPy, Matplotlib, Plotly, Seaborn
ML/AI: Scikit-learn, TensorFlow, PyTorch,, model evaluation, error analysis, LLM, LangGraph, LangChain, RAG.
Cloud & DevOps: Azure, Google Cloud, AWS (ECS Fargate, RDS PostgreSQL, ElastiCache Redis, S3, CloudFront, Secrets Manager, CloudWatch, VPC, ALB), Docker, Kubernetes.
Data/DE: Feature engineering, data cleaning, Postgres, BigQuery, Supabase, REST APIs
Visualization & Dashboards: Power BI, Tableau, Streamlit
Dev Practices: Git, FastAPI, testing and documentation, stakeholder presentations

EXPERIENCE

Research Intern **Aurora, Colorado**
University of Colorado Anschutz Medical Campus *Jan 2026 – Present*
- Supporting **annotation workflows** and **LLM/model evaluation** using **rubrics**, **QA checks**, and **error analysis** to surface **failure modes** and **clinical relevance** issues.
- Led **prompt engineering** by refining **templates**, **structured outputs**, and **guardrails**; ran systematic **prompt experiments** to improve **reliability** and reduce **errors**.

AI Research Engineer (Capstone Project) **Boulder, Colorado**
Honda 99P Labs *Sept 2025 – December 2025*
- Designed a **multi-agent CEWI debate engine with hybrid retrieval** (**vector database**) and a curated evidence pack for reproducible arguments.
- **Implemented a FastAPI and Uvicorn backend** with instrumentation inspired by **LangFuse and Graphiti-style context engineering**; applied advanced RAG patterns for high-signal citations.

Junior Data Scientist Intern **Boulder, Colorado**
Kobeyo *May 2025 – August 2025*
- Implemented an **MLOps pipeline** for a custom **BERT model**, creating a **self-improving system** with a **human-in-the-loop feedback** mechanism that continuously fine-tuned the model and **boosted classification accuracy** in a production environment.
- Developed and **deployed a full-stack data ingestion API** using Python, which orchestrated a **resilient web scraping engine (Playwright)** and **integrated with a Supabase (PostgreSQL) backend** to provide real-time, structured data.
- Leveraged **LLMs (OpenAI) and advanced prompt engineering** to build a module for **extracting structured job skill data from unstructured text**, enhancing the granularity of the company's core dataset.

PROJECTS

CodeSense: AI-Powered Code Review Automation Platform **Boulder, Colorado**
Independent Project *Oct 2025 – December 2025*
- **Architected a Cloud-Native AI Code Reviewer** by integrating **GitHub Webhooks** and **GPT models** to automate security scanning (Bandit) and generate real-time patch diffs for code vulnerabilities.
- **Engineered a scalable FastAPI backend** featuring **15+ REST endpoints**, JWT authentication, and **async Redis RQ workers** to handle multi-tenant repository management and high-volume webhook processing.
- **Deployed a robust AWS infrastructure via Terraform**, utilizing **ECS (Fargate), RDS, and ElastiCache** within a secure VPC, while managing containerization and CI/CD through **Docker** and **GitHub Actions**.

AI Interview Coach **Boulder, Colorado**
Independent Project *March 2025 – May 2025*
- Built an **AI Interview Coach** using **Python, FastAPI, Streamlit, and LangChain** to generate role-specific interview questions and answer guidance from job descriptions.
- Implemented a **RAG-based workflow** to ground responses in curated prep material and provide structured feedback (rubrics + follow-ups) for mock interview practice.

Yoga Posture Correction and Detection | https://journals.stmjournals.com/joosdt/article=2024/view=161704/ **Mumbai, India**
Team Lead *Aug 2023 – May 2024*
- **Published an article** and led a team to develop **a yoga pose estimation model** with **96.5% accuracy** using **neural networks** and **TensorFlow,** deploying it with **TensorFlow.js** for real-time feedback.
- Processed **3,600+ images** for data cleaning and structuring, enhancing **model accuracy** and **training effectiveness by 30%,** while creating an **interactive platform,** fostering **40% higher user participation**.

Certifications / Publications

- Completed several Google Cloud certificates and Badges:
  https://www.cloudskillsboost.google/public_profiles/f4a0a0fd-5ea0-44d9-8734-b180faa9f1b9