Girwan Dhakal

205-239-6124 | girwandhakal@gmail.com | linkedin.com/in/gdhakal | github.com/girwandhakal

Education

The University of Alabama — Tuscaloosa, AL
B.S. & M.S. in Computer Science, Accelerated Master's Program — *Aug. 2022 – May 2027*

- Relevant coursework: Software Engineering, Algorithms, Database Management, Artificial Intelligence, Machine Learning.

Experience

Shipt — Birmingham, AL
Data Science Intern, AI Platform — *Jun. 2026 – Aug. 2026*

- Developed a lower-cost Gemini OCR pipeline on GCP and Snowflake to process previously excluded receipt data, expanding tax-recovery coverage to **99%** and identifying **$10K** in projected annual vendor savings.
- Tuned Shipt's product substitution-ranking model with Optuna, improving **Recall@K by 20%** and MRR from **0.550 to 0.559**.
- Wired internal RAG-as-a-service and LLM gateway tools to Prometheus and Grafana dashboards, improving visibility into service usage and supporting observability reviews.
- Shipped a LangGraph RAG assistant for infrastructure and support queries, supporting **40 internal users**; tested **10 tool calls** using Langfuse tracing and Promptfoo evaluations.
- Interviewed **10+** data scientists and engineers to uncover workflow pain points and shape product improvements around internal user needs.
- Enabled self-serve ad-hoc reporting for leadership by developing Snowflake Cortex Analyst agents that convert natural-language questions into SQL.

Alabama Credit Union — Tuscaloosa, AL
Data Analyst Co-op — *Aug. 2025 – May 2026; Aug. 2026 – Present*

- Created SQL export and backup workflows that preserved **25M+** member, account, and transaction records during a core banking migration.
- Maintained and optimized SQL Server ETL pipelines, reducing daily refresh time by **30%** by eliminating inefficient joins and redundant CTEs.
- Delivered **8 Power BI dashboards** across business functions and provided weekly training sessions to **100+ employees**, increasing adoption of data-driven decision-making.
- Diagnosed a threefold increase in daily ETL runtime, traced it to elevated server usage, and partnered with IT to resolve the infrastructure provisioning issue.

Alabama Life Research Institute — Tuscaloosa, AL
Machine Learning Researcher — *Jan. 2025 – Present*

- Built a child-speech analysis pipeline using a pretrained CRF model, hypothesis testing, and ensemble learning to identify patterns associated with late language development; submitted a **first-author paper** for peer review.
- Reduced word error rate on child-speech ASR transcripts by **18%** using transfer learning with Google FLAN-T5 for generative error correction.
- Unified speech, transcript, and gesture data from **13 datasets** into pipelines for multimodal model training and evaluation.
- Evaluated results across **30,000+ utterances** using mixed-effects models and statistical tests.
- Designed multimodal gesture-classification workflows with YOLO for person detection and ViTPose for keypoint extraction; benchmarked 3D CNN, graph-convolution, skeleton-sequence transformer, and vision-transformer models, advancing understanding of gesture-based communication.

RoomiCheck — Tuscaloosa, AL
Software Engineering Lead — *Dec. 2025 – Present*

- Led engineering for RoomiCheck, a student-led startup improving roommate matching through intelligent questionnaires and data analysis; won two entrepreneurship competitions, securing **$3,500 in total funding**.
- Piloted a FastAPI roommate-matching system on Vercel with Supabase among **100+ users**, gathering feedback that informed product changes and clarified product feasibility.

Technical Skills

Languages: Python, SQL, Java, JavaScript, TypeScript
Data & Machine Learning: PyTorch, scikit-learn, Optuna, Pandas, NumPy
AI & Observability: LangGraph, Snowflake Cortex Analyst, Langfuse, Promptfoo, Prometheus, Grafana
Platforms & Tools: SQL Server, Snowflake, GCP, Power BI, FastAPI, Vercel, Supabase, Git, Docker