

Bibek Dhungana

Nashville, TN bibek.dhungana@vanderbilt.edu (832) 310-6869
bibekdhungana.com linkedin.com/in/dhunganabibek github.com/dhunganabibek

Summary

Machine learning Ph.D. student at Vanderbilt University with 3+ years of production software and data engineering experience. I work on applied deep learning, neural retrieval, and recommendation, and I build the systems around a model as well as the model itself: data pipelines, event-driven services, and deployment. Production impact includes 11M+ records reconciled via automated ETL, 2M+ weekly API interactions served, and 99.9% pipeline uptime. 6x AWS Certified.

Education

Vanderbilt University, Nashville, TN

2025 – Present

M.S. & Ph.D. in Computer Science

GPA: 4.0 / 4.0

M.S. expected 2027; Ph.D. expected 2030. Advised by Dr. David Hyde (SOL Lab).

Relevant Coursework: Machine Learning, Deep Learning, Artificial Intelligence, Numerical Analysis, Big Data, Advanced Machine Learning, AI Systems, Representation Learning, Quantum Computing

Texas Tech University, Lubbock, TX

2019 – 2023

B.S. in Computer Science, Minor in Mathematics

GPA: 3.9 / 4.0

Relevant Coursework: Data Structures & Algorithms, Operating Systems, Computer Architecture, Database Management Systems, Computer Networks, Software Engineering, Theory of Automata, Discrete Mathematics

Research Experience

Vanderbilt University, SOL Lab, Nashville, TN

Aug 2025 – Present

Graduate Research & Teaching Assistant · Advised by Dr. David Hyde

- Built [research.team](#), an LLM-powered grant discovery and team-formation platform, grounded in a mixed-methods HCI study (N=36 interviews, N=306 surveys); first-authored paper under review at ACM CSCW 2027.
- Shipped a dense retrieval engine over 80,000+ federal grants and 500 researcher profiles that beat BM25, TF-IDF, and keyword baselines (MRR 0.54 → 0.74; nDCG@5 +47%).
- Cast grant and collaborator matching as one recommendation problem: nearest-neighbor candidate generation over embeddings, similarity ranking, and offline evaluation on a held-out relevance set.
- Teaching Assistant for Algorithms (CS 3250), supporting 60+ students through weekly office hours, exam review sessions, and grading on core data structures and complexity concepts.

Mobility Automation Lab, Texas Tech University, Lubbock, TX

Jan 2021 – Jan 2022

Research Assistant

- Built an edge-to-cloud streaming pipeline on AWS EC2/S3 ingesting real-time NVIDIA Jetson sensor streams, enabling live position tracking and proactive fault detection via WebSocket dashboards.
- Implemented RGB-D SLAM using ROS and probabilistic sensor fusion, improving real-time mapping accuracy for autonomous robots in dynamic environments.

Industry Experience

St. Jude Children's Research Hospital (ALSAC), Memphis, TN

Jul 2024 – Jul 2025

Software Engineer · AWS Glue, Apache Kafka, MuleSoft, SQL Server, .NET, OpenShift, Docker, Kubernetes

- Engineered AWS Glue ETL jobs to reconcile 11M+ InContact records with NICE APIs, eliminating call-recording gaps and automatically generating NACHA-compliant audit reports.
- Designed a Kafka + MuleSoft event-driven streaming pipeline propagating donation events across downstream services, deployed on OpenShift through Jenkins CI/CD and sustaining 99.9% uptime.
- Enforced data quality with multi-layer validation across .NET APIs and SQL Server data contracts, rejecting malformed donor records at ingest rather than surfacing them in downstream reporting.
- Replaced a legacy .NET donation system with a React + .NET platform, cutting call-center handling time by 20% and achieving full PCI compliance across all transaction flows.

AGCO (Precision Planting), Bloomington, IL

Jun 2023 – Jun 2024

Software Engineer · Python, Node.js, AWS Aurora, DynamoDB, S3, Express.js, GitHub Actions

- Built a unified data catalog consolidating 5,000+ assets (PDFs, videos, decision trees) from Brandfolder, Wistia, and Stonly into AWS Aurora, with nightly sync pipelines and automated metadata extraction and cross-source taxonomy mapping.
- Migrated AGCO's monolithic marketplace to a distributed architecture on AWS Aurora, removing shared-database coupling and enabling independent per-service deployments.
- Automated nightly 900-page reporting via a Puppeteer + S3 batch pipeline, eliminating manual report generation and saving hours of analyst time per week across Precision Planting sites.

- Delivered 15 production REST APIs on Aurora Serverless with standardized data contracts and JWT auth, consolidating data access for downstream consuming services.

Heavy Construction Systems Specialists (HCSS), Houston, TX

Jun 2022 – May 2023

Software Engineer Intern · .NET, React, Azure, PostgreSQL, Docker, Azure DevOps

- Launched a unified API developer portal spanning all HCSS products, scaling to 2M+ weekly interactions with sub-second response times.
- Automated API publishing and validation for 10 engineering teams by shipping a custom Azure Marketplace pipeline extension, cutting per-release manual overhead 3x.
- Reduced API release lead time from 2 days to 6 hours by designing multi-stage CI/CD pipelines across QA, staging, and production environments in Azure DevOps.

NSF I-Corps Program, Lubbock, TX

May 2021 – Aug 2021

Software Engineering Intern · Spring Boot, React, MySQL, AWS

- Designed normalized MySQL schemas and ETL pipelines on AWS to power real-time product analytics, surfacing KPIs that informed prioritization decisions.
- Delivered full-stack features with React, Spring Boot, and OAuth2, and conducted 20+ customer discovery interviews to define MVP scope.

Projects

- **ResearchTeam:** live demo of the grant discovery and researcher-matching platform described above.
- **Quantum Adaptive Self-Attention (QASA)**, *independent reimplement and extension of published work*: Hybrid quantum-classical Transformer that replaces the final encoder block's FFN with a parameterized quantum circuit (data re-uploading + entanglement) in PyTorch and PennyLane, trained end-to-end via the parameter-shift rule. Reproduced the reported behavior and extended it with a hardware-noise study: came within 0.02 R^2 of a classical Transformer baseline (0.88 vs. 0.90) with ~6% fewer parameters, and retained R^2 0.86 under IBM Quantum hardware-noise evaluation.
- **StreamCab: Real-Time Taxi Fare Intelligence:** Containerized Kafka + Spark Structured Streaming pipeline (Docker Compose) ingesting NYC TLC taxi trips and computing 10-minute zone aggregates in PostgreSQL. Trained a memory-aware XGBoost fare predictor (continued boosting over batched parquet) achieving 3.42% MAPE – a 78% error reduction over the zone-hour-mean baseline (MAE \$0.49 vs. \$1.99), surfaced in a live dashboard.
- **Fully Local Retrieval-Augmented Generation Pipeline:** Privacy-first RAG system (LangChain, ChromaDB, Hugging Face MiniLM embeddings, Ollama) ingesting academic PDFs into a semantic vector store and synthesizing grounded outputs via a locally-served LLM (Llama 3 / gpt-oss 20B), with no external API calls or cloud costs.

Technical Skills

Machine Learning & MLOps: PyTorch, Transformers, Hugging Face, scikit-learn, XGBoost, MLflow, LLMs, Generative AI, RAG, Recommender Systems, LangGraph, Feature Engineering, Experiment Tracking, Model Deployment, NLP, Computer Vision

Data Engineering & Big Data: Apache Spark, Apache Kafka, Spark Structured Streaming, Apache Airflow, AWS Glue, ETL/ELT, Data Modeling, Data Quality, Workflow Orchestration, Hadoop, Distributed Systems, NumPy, Pandas

Cloud & DevOps: AWS (SageMaker, Glue, Redshift, Lambda, EC2, S3, Aurora), Kubernetes, Docker, Terraform, CI/CD, GitHub Actions, Jenkins, OpenShift, Azure

Databases & Warehousing: PostgreSQL, MySQL, SQL Server, Amazon Redshift, MongoDB, DynamoDB, Redis, Data Warehousing, Vector Databases (S3 Vectors, Qdrant, ChromaDB)

Languages & Frameworks: Python, TypeScript, JavaScript, C#, Java, SQL, Rust, React, Node.js, .NET Core, FastAPI, Spring Boot, Tailwind CSS

Quantum Computing: PennyLane, Qiskit, Parameterized Quantum Circuits, Parameter-Shift Rule, Hybrid Quantum-Classical Models

Certifications

AWS Certified (6x): Machine Learning Engineer, Data Engineer, Solutions Architect – Associate, Developer – Associate, Cloud Practitioner, AI Practitioner · **HashiCorp:** Terraform Associate