Vedangi Bengali

vedangibengali@tamu.edu | Homepage | LinkedIn | GitHub | Google Scholar

Computer Science Ph.D. candidate with research in building scalable algorithms on large graphs & hypergraphs. Experience translating complex problems into reliable agentic and retrieval systems, by using deterministic methods to make agents more trustworthy. Prior industry experience at HSBC and Walmart Global Tech.

Education

Texas A&M University — College Station, TX
Ph.D. Candidate, Computer Science — GPA 4.0/4.0 — Advisor: Dr. Nate Veldt — *Expected May 2027*
M.S. Computer Science (Thesis Option), GPA 4.0/4.0 — *Aug 2022*

College of Engineering Pune — Pune, India
B.Tech., Electronics and Telecommunication — CGPA 8.8/10 — *May 2017*

Selected Graduate Coursework: Algorithms for Big Data & Graph Mining, Distributed Systems & Cloud Computing, Software Engineering, Analysis of Algorithms, ML, AI, Multimodal LLM Agents, Security Engineering.

Technical Skills

Languages: Python, C/C++, Java, Julia, SQL (PostgreSQL, SQLite), JavaScript
AI & ML: PyTorch, scikit-learn, LightGBM, Keras, RAG pipelines, dense embeddings, BM25, reciprocal rank fusion, LLM tool-calling and agent loops, offline evaluation harnesses
Tools: AWS, Git, OpenMP, FastAPI, Django, React, REST APIs, Spark, Snowflake

Engineering and AI Projects

RepoCompass — code-understanding agent for Python repositories | *Python, FastAPI, React, SQLite, networkx*
- Built and deployed a 14-tool agentic system that indexes public GitHub repositories into a queryable call graph, and answers code-related questions; indexed 12,500 symbols and 8,000 call edges from the Django codebase.
- Implemented a hybrid retrieval system combining graph traversal for structural queries with regex, BM25, and dense embeddings for conceptual questions, fused by Reciprocal Rank Fusion and selected through a query router.
- Validated retrieval, graph construction, tool execution, and citation grounding through 360 automated offline tests and hand-labeled queries. Exposed the system through a FastAPI backend and interactive React interface.

ResponseGuard-OT — industrial cyber-response assurance | *Python, React, AWS, NetworkX, PuLP* 2026
- Designed an operator-facing agentic workflow that translates industrial incident scenarios into graph constraints, verifies proposed response plans, and repairs unsafe actions through constrained optimization.
- Integrated chemical-process and power-grid simulators with an interactive incident-replay interface enabling users to compare original and repaired response plans while retaining approval over recommended actions.

AIDE — retrieval-augmented economic intelligence assistant | *Snowflake Cortex, Streamlit, SQL* 2026
- Built an end-to-end RAG assistant with document ingestion, metadata-aware chunking and retrieval, grounded citations and a user-facing Streamlit application.
- Implemented query planning and reflection and evaluated retrieval precision and answer faithfulness to identify unsupported or weakly grounded responses.

PHOENIX — forecasting and risk-tiering service | *Snowflake ML, Snowpark, Streamlit, SQL* 2026
- Built a data-to-dashboard pipeline producing 30-day COVID forecasts, country-level risk tiers, and LLM-generated briefings across 15 countries.
- Combined structured forecasts with LLM-generated explanations in an interactive application while preserving country-level source data and metrics; several country models reached 2–3% MAPE.

ML projects

Laser-Track Geometry Prediction | *Python, gradient-boosted trees, conformal prediction*

∗ Developed a leakage-controlled ML pipeline for predicting melt-track geometry from thermal imaging; reduced MAE 2.8× over constant baseline on a fully held-out process condition, and produced calibrated prediction intervals.

Graph Clustering for Regulatory Gene Discovery | *Julia* 2025

∗ Applied multi-resolution LambdaCC clustering to RNA-seq co-expression networks to shortlist candidate transcriptional regulators; stress-tested stability across similarity measures, graph sparsity, and resolution parameters before handing candidates to wet-lab collaborators.

Research and Industry Experience

Graduate Research Assistant — Texas A&M University Aug 2022 – Present
Algorithms and machine learning for large-scale graph and hypergraph data *College Station, TX*

· Designed and implemented scalable graph and hypergraph algorithms in Julia and C++ (OpenMP), processing real-world networks containing millions of nodes and edges, and outperforming exact ILP baselines by orders of magnitude while maintaining competitive solution quality.
· Built leakage-controlled, forward-in-time ML training and evaluation pipelines for response prediction over hypergraph communication datasets under severe class imbalance, with per-feature ablations attributing gains to specific signals rather than model capacity.
· Implemented and benchmarked linear, MLP, and transformer-style set encoders in PyTorch behind a shared negative-sampling and evaluation harness for higher-order link prediction.
· Maintained reproducible benchmark suites comparing new methods against exact solvers and prior work across public datasets; released open-source implementations accompanying published papers.

Data Engineer Intern — Walmart Global Tech Jun 2021 – Aug 2021
Cloud migration capacity planning *Dallas, TX*

· Built forecasting pipelines over historical CPU-usage telemetry (SARIMA, Prophet, LSTM) to project weekly compute demand for a Teradata-to-BigQuery migration; output fed capacity-planning decisions for the migration.

Software Engineer — HSBC Software Development Jul 2017 – Aug 2018
Core banking platform *Pune, India*

· Designed and launched Autosolv, a Django search and Q&A service surfacing fixes for recurring AS400 build errors, cutting repeat build failures by 15% and reducing manual triage for the build team.

Publications

- **V. Bengali**, N. Veldt. "Faster Approximation Algorithms for Parameterized Graph Clustering and Edge Labeling." **CIKM 2023**.
- **V. Bengali**, N. Veldt. "The Densest SWAMP Problem: Subhypergraphs with Arbitrary Monotonic Partial Edge Rewards." **ECML-PKDD 2025**. Extended journal version under review.
- F. Adriaens, **V. Bengali** et al. "Improved Hardness and Approximations for Cardinality-Based Minimum s–t Cuts in Hypergraphs." Under final journal revision, 2025.
- "A Survey on LLMs for Story Generation." **Findings of EMNLP 2025**. Extended journal version accepted 2026.
- **V. Bengali**, N. Veldt. "Measuring and predicting reciprocity in hypergraph interaction networks." **Submission in process** 2026.

Teaching and Service

- Teaching assistant in Machine Learning, Discrete Structures, Programming Languages, and Database Systems, supporting 200+ students; built course autograders and ML assignments, and mentored students one-on-one through project debugging.
- Reviewer: AAAI 2025, ECML-PKDD 2025, ALENEX 2026, ACM TIST 2026.