# Vageesha Datta Ganapaneni

✉ vageeshadattag@gmail.com | 🔗 linkedin | ⓖ Github | 📞 +1(469) 805-1906 | 🏠 Dublin, California

## EDUCATION

- **Master of Science in Computer Science**
  The University of Texas at Dallas
  *Relevant Coursework: Machine Learning, Artificial Intelligence, Computer Vision, Natural Language Processing, Statistics in AI and ML, Web Programming Languages*

## TECHNICAL SKILLS

- **Languages:** Python, JavaScript, TypeScript, Rust, SQL, Bash, Linux
- **Frameworks & Libraries:** PyTorch, TensorFlow, LangChain, LangGraph, Flask, FastAPI, React, Node.js, D3.js
- **Tools & Platforms:** Docker, Kubernetes, AWS, Git, FAISS, Redis, vLLM, REST APIs

## EXPERIENCE

- **Handshake** — Remote
  AI Fellow — *Oct 2025 – Present*
  - Applied prompt engineering, conducted 1K+ machine learning code reviews of coding AI agents to detect rogue behaviors and integrate corrected experimentation, improving model compliance accuracy by **2%**.
  - Enhancing the scientific integrity evaluation framework by designing custom metrics, contributing to the development of internal benchmarks for responsible AI governance within Handshake's MOVE research program.

- **HUD** — San Francisco, CA
  Research Engineer – Agentic AI Evals — *Oct 2025 – Nov 2025*
  - Designed a distributed evaluation framework by parallelizing rollouts, vLLM inference, and fine-tuning across H100 GPU clusters, improving **evaluation throughput** by **15%** and accelerating agentic model benchmarking.
  - Evaluated AI agents across multiple **codebases**, assessing their ability to interpret **pull requests (PRs)**, implement code changes, and design test cases verifying modified behavior through multi-turn reasoning.
  - Refactored the evaluation harness into a modular, configuration-based system supporting plug-and-play model integration, reducing setup effort by **25%** and enabling faster experimentation for future client evaluations.
  - Developed interactive **React** dashboards connected to **Python** and **Docker**-based analytics, cutting manual result inspection time by nearly **20%** and improving visibility across research and engineering teams.

- **Allen Institute for AI (Ai2)** — Dallas, TX
  AI Researcher — *Jan 2025 – May 2025*
  - Applied **Optimization by Prompting (OPRO)** to enhance few-shot reasoning, refining prompt design to achieve a **12%** gain in hypothesis generation accuracy across active reasoning benchmarks.
  - Developed belief-tracking and uncertainty quantification components using **entropy** and **KL divergence**, allowing early detection of reasoning drift and improving interpretability in complex inference tasks.
  - Built an interactive **D3.js** platform visualizing confidence evolution across reasoning chains, streamlining analytical reviews and increasing transparency for research teams.

- **IBM** — India
  AI Software Engineer — *Jan 2020 – May 2022*
  - Deployed AI-driven automation pipelines across hybrid cloud environments, improving operational throughput by **20%** and minimizing manual touchpoints in core business processes.
  - Processed and structured over **10 million** data records to enable accurate training of predictive models that strengthened process reliability and decision support.
  - Introduced reusable **libraries** and **REST APIs** that standardized automation routines, reducing integration timelines by **30%** and facilitating organization-wide adoption of AI tooling.
  - Containerized machine learning solutions with **Docker** and orchestrated deployments using **Kubernetes**, bringing down release time from hours to minutes while ensuring consistent performance across environments.

## PROJECTS

- **Vision-Based Web Automation Agent**: [Source Code]
  - Built an autonomous web automation agent using GPT-4o Vision that executes multi-step tasks through an Observe-Think-Act loop, leveraging IoU-based filtering to deduplicate bounding boxes for accurate interactive element detection.
  - Implemented intelligent action execution with visual grounding and text-matching fallbacks, achieving 98.7% accuracy in solving web automation tasks with robust loop detection and verification mechanisms.

- **MoodBoard AI:Context-Aware Pin Recommendations**: [Source Code]
  - Built a context-aware recommendation system that generated personalized "mood boards" by combining graph embeddings of user–pin interactions with multimodal inputs, enabling richer context modeling for recommendations.
  - Incorporated LLM-based embeddings of image and text metadata into the retrieval pipeline, improving recommendation quality and boosting recall by 17% in offline evaluations.