Varuni H.K

LinkedIn: https://www.linkedin.com/in/varunihk/ Email: varunihk7@gmail.com

Github: https://github.com/varuni7 Mobile: +91-8296176533

Portfolio website : https://portfolio-varuni-varuni7.vercel.app/

Languages & Interests

Languages: Python, Go, Haskell

Technical Interests AI for code optimization , Agentic AI, Multimodal Information Retrieval , Indexing Techniques (currently exploring learned indexes) HNSW data structure , NLP , Graph Theory GNNs and Distributed Graphs
Soft Skills: Research, Strategic planning, Design Thinking, Leadership, Team Building, Public Speaking

Publications

Polaris : Multi Agentic System for Conversational Enterprise Analytics - accepted at AAAI - 26 (Workshop on Agentic AI Benchmarks and Applications for Enterprise Tasks, Singapore) lead author

GAT - Grapheme Aware Tokenization - accepted at MIT Undergraduate Research Technology Conference (MIT URTC), Massachusetts Institute of Technology, USA , lead author

Efficient Indexing for Google Zanzibar Based Authorization Systems Using Graph Dynamic-Transitive Closures - accepted at the 28th International Conference on Advanced Communications Technology, Pyeongchang, South Korea, lead author

Exploring caching and indexing mechanisms for Zanzibar based authorization systems - presented at the 11th international conference on Applied Systems Innovation, Kyoto, Japan , lead author

Leveraging Permissioned Blockchain for securing controlled drug prescriptions in India - presented at the 6th international conference on Blockchain Computing and Applications, Dubai , second author

SubGINs - Approximate and Exact Subgraph Matching with GINs - submitting to ICML'26

Graph-RAG for Go code optimisation and refactoring leveraging abstract syntax trees accepted at 21st International Conference on Artificial Intelligence Applications and Innovations

Active Research Member

- **RISE MICCAI Mentorship Program** Oct 2025 – Oct 2026

  Research Mentee
  - Working under **Dr. Seyyedeh Qazale Mirsharif** on *Personalized Lung Cancer Treatment Prediction* using heterogeneous graph models

- **ML Commons**

  Working Group Member *Sep 2025 - Presemt*
  - **MLPerf Storage Working Group**: Contributing to the development of MLPerf Storage benchmarks to characterize performance of storage systems supporting machine learning workloads
  - **Medical AI Working Group**: Developing the MedPerf platform and GaNDLF to establish standardized benchmarks for Medical AI

Education

- **PES University** Bengaluru, India

  Bachelor of Technology - Computer Science; GPA: 8.88 (CNR Rao Scholarship Awardee) September 2021 - June 2025
- **Christ Junior College** Bengaluru, India

  PCMC; Karnataka Board Exam - 98.5% *June 2019 - June 2021*

Experience

- **Software Engineer** July 2025 - Present

  Couchbase Index
  - Working on the indexing layer and encryption at rest
  - Worked on alter index pipeline to allow for efficient replica repair.

- **Couchbase AI Intern** Jan 2025 - June 2025

  Interactive Data Experience
  - Building **conversational** layer with analysis, reasoning and visualization over Couchbase operational database using **Multi Agentic workflows** with **Langraph**
  - Researched Reinforcement Learning based tool calling for agentic workflows

- **JP Morgan Cloud & AI Engineering Intern** June 2024 - August 2024

  Cloud Migration, ML Ops , Infrastructure as Code and Environment as Code
  - **Go based cloud automation tool** to minimize Terraform code ensuring optimised resource management across parameters like availability, security and cost . Created the **EKS Template** within the tool
  - Deployed a portfolio optimization model to AWS SageMaker, automating data ingestion, training, and real-time inference via managed endpoint using a Blue-Green CI/CD Pipeline.
  - Onboarded an internal trading application onto the automation framework, reducing AWS cloud costs by **36%**.

PES Innovation Lab (formerly Microsoft Innovation Lab) May 2023 - July 2023

Development of a new Robotic motion planning and search algorithm

- Developed BIT* Connect, **a novel sampling-based motion planning and search algorithm** that blends BIT*'s elliptical heuristic with RRT-Connect's bi-directionality. Algorithm Implemented in Open Motion Planning Library.

PES University, SME (Natural Language Processing) February 2023 - April 2023

Introduction to Machine Learning and Natural Language Processing

- Taught a **4 week course** to a class of 40 students with curated content to familiarize beginners with the fundamentals of NLP (Transformers, Word2Vec, Tokenizers, NER , Summarization)

Data Analytics Teaching Assistant August 2024 - December 2024

R , Python

- Took hands - on workshops on R programming
- Curated teaching material for
- graded assignments and hosted doubt clearing sessions

Projects

- **Transitive Closure based Caching and Indexing in Google Zanzibar like Distributed Systems**: Implemented dynamic transitive closure based indexing to improve query performance and scalability in Zanzibar-like distributed systems. Improved read - latencies (in a mixed workload of 99% reads and 1% writes) by 57% compared to vanilla SpiceDB.
  Tech: SpiceDB,Go,Python
- **Semantic Search Engine**: Built a semantic search engine using Haystack's text to image search pipeline; Multimodal retiever to retrieve images as answers to a text query for the WikiArt dataset. Tech: OpenAI CLIP Embedding , FAISS Index, Elasticsearch , Flutter, Firebase
- **RAG based Code Optimisation and Review Evaluation**: Developing a RAG-based system that recommends suitable Golang package APIs to replace existing user-written code for enhanced efficiency and maintainability. Novelty in using Abstract Syntax Trees to optimise code flow .
  Tech : Llama Index, Claude, Pinecone
- **Fuzzy Search and Semantic Alignment of Multi-Perspective Narratives**: Built a fuzzy and semantic search engine for Mahabharata retellings (e.g., "Draupadi's PoV," "Karna's PoV") . Levenshtein distance for typo tolerance and TF-IDF/ BM25 for keyword-based ranking. Implemented semantic search with sentence embeddings and cosine similarity for PoV-aware retrieval, indexing character-specific versions. Applied vector space models for re-ranking and cross-version thematic comparison.
- **Project Roots**: Built a flutter app with personalized cognitive games to be used by cognitive specialists for better memory reinforcement through cognitive exercises while treating patients. **(this project received a 5 Lakh funding from Cisco Thingqubator)**
  Tech: Flutter, Firebase, Figma

Accomplishments

- Selected to conduct a workshop on distributed graphs at **RootConf'25**
- **CNR Rao Scholarship** Academic Excellence awardee of **INR 20,000** in my 3rd year - PES University
- Ranked in **top 1%** from 27,000 applicants , awarded 100 % scholarship for the **Women Engineers** program by Google and TalentSprint.
- Granted startup incubation and seed funding by **Cisco Thingqubaotr** (5 Lakhs).
- Selected as a **Harvard HPAIR** delegate for 2024.
- **JP Morgan Code for Good Hackathon** . Among 545 students nationwide ,**top 4** - university.
- Top **2.5%** from all over India in the **Google Girl Hackathon** 2023.
- **1639** rank in **Google Hash Code** 2022.

Communities

- **WiDS Ambassador Women in Data Science (WiDS) at Stanford University**:
  Helped in hosting the WiDS regional event at Indian Institute of Management Bangalore. (January 2022 - Present)
- **GDSC Lead** : Conducted workshop on RAG (Aug 2023 - Sep 2024)
- **Head of PES Innovation Lab(formerly Microsoft Innovation Lab) for 2**: Conduct research internship over the summer, mentored 3+ research projects , conducted workshops , hackathons and CTFs. (023-2024)