Vasudev Menon

919-798-7081 | vmenon2@ncsu.edu | vmenon04.github.io | linkedin.com/in/vmenon04 | github.com/vmenon04

Education

North Carolina State University — Raleigh, NC
B.S. in Computer Science and B.S. in Mathematics | Cum. GPA: 4.0/4.0 — *Expected Spring 2026*
- Coursework: Deep Learning, Artificial Intelligence, Data Science, Linear Algebra, Data Structures and Algorithms, Operating Systems, Combinatorics, Real Analysis, Differential Equations, Discrete Mathematics
- Honors and Awards: Phi Beta Kappa, Dean's List All Semesters

Eötvös Lorand University — Budapest, Hungary
Budapest Semesters in Mathematics (BSM) — *Fall 2025*
- Coursework: Deep Learning, Graph Theory, Advanced Combinatorics (Hypergraph Theory)

Experience

Fidelity Investments — Durham, NC
Software/Machine Learning Intern - Summer 2025 — *June 2025 – August 2025*
- Engineered **RAG** system to automate translation of pension plan documentation into executable Java code
- Used **AWS Bedrock**, **LangChain**, **HuggingFace**, and **Boto3** to build scalable **ML infrastructure**
- Designed **Neo4j** graph database to encode pension rule relationships and dependencies for code generation
- Optimized LLM accuracy through **prompt engineering**, validation pipelines, and context management techniques
- Deployed end-to-end **Streamlit** application integrating RAG pipeline, graph database, and code generation workflows

Fidelity Investments — Durham, NC
Software Engineer Intern - Summer 2024 — *June 2024 – August 2024*
- Engineered an enterprise assessment system used by 4,000+ company leads to evaluate application reliability
- Implemented a dashboard integrated with **Power Platform API**
- Built and deployed **Backstage** developer portal plugins in **Node.js** and **TypeScript**, enhancing internal tooling
- Utilized **Docker** for **DevDocs** integration to provide documentation for users

North Carolina State University — Raleigh, NC
NSF-Funded Deep Learning Research Intern — *May 2023 – August 2023*
- Automated **deep learning** experiments via **shell scripting** and **R**, analyzing the performance of **batch normalization**
- Trained/evaluated **neural nets** with statistical analysis and metric visualizations (precision, recall, time)
- Contributed to NSF-funded research through experiment design and literature synthesis

Rice University — Remote
Data Science Intern (NSF Project) — *June 2021 – Feb. 2022*
- Conducted research on an NSF-funded project to analyze and extrapolate COVID-19 sequence data using **R**
- Utilized numerous state-of-the-art statistical algorithms to examine DNA sequence data
- Analyzed **100,000+** COVID-19 strains using statistical models, uncovering key genomic differences
- Presented research at the **Junior Science and Humanities Symposium**

Projects

Mechanistic Interpretability of Transformers on Simple Games | *PyTorch, NumPy, HuggingFace*
- Conducted **interpretability research** comparing hardcoded vs learned **transformer models** on controlled game tasks
- Designed novel experimental framework for enabling direct comparison of hand-engineered and gradient-learned solutions
- Analyzed transformer internals to understand how **neural networks** learn **algorithmic reasoning**

Differential Privacy in Federated Learning (FL) on MIMIC-IV | *AI/ML, Python, PyTorch, Flower, Big Data*
- Implemented **Distributed Differential Privacy (DDP)** in FL using **Flower** to protect sensitive healthcare data
- Simulated heterogeneous clients using racial groupings from the **MIMIC-IV** ICU dataset to reflect real-world skew
- Analyzed trade-offs between privacy and model accuracy, with **Gaussian noise** and **clipping** techniques
- Overcame large-scale (**100GB+**) data processing challenges with **SQL** and **Power Query**
- Collaborated with Dr. Kotevska (ORNL) and Dr. Haider (NCSU); authored and presented findings

Rapster: Full-Stack Music Sharing Platform | *AI/ML, FastAPI, Supabase, Docker, Essentia, Cloudflare R2*
- Developed a full-stack app to upload and analyze audio tracks with ML-powered **feature extraction** (BPM, key, MFCCs)
- Deployed ML workloads using **Essentia** containerized with **Docker** and orchestrated via **Docker Compose**
- Stored user files in **Cloudflare R2** and metadata in **Supabase**, using signed URLs for access control

Technical Skills

Programming: Python, Java, C/C++, R, SQL
AI/ML: PyTorch, TensorFlow, Keras, Scikit-learn, HuggingFace, Federated Learning, Differential Privacy
Systems & Tools: Docker, AWS, Boto3, Git, CI/CD, Power Platform, Linux, REST APIs
Frameworks: FastAPI, Flask, Django, Flower, Streamlit
Other: Data Visualization (Matplotlib, Pandas), LaTeX