

VASUDEV SYAMSUNDER MENON

919-798-7081 | vasumenon@gmail.com | vasumenon.com
linkedin.com/in/vmenon04 | github.com/vmenon04

EDUCATION

North Carolina State University

Raleigh, NC

B.S. in Computer Science & B.S. in Mathematics

Expected Spring 2026

Cumulative GPA: 4.0/4.0

Relevant Coursework: Deep Learning, Artificial Intelligence, Data Science, Linear Algebra, Data Structures and Algorithms, Operating Systems, Combinatorics, Real Analysis, Differential Equations, Discrete Mathematics

Honors: Phi Beta Kappa, Dean's List (All Semesters), Provost's Professional Experience Program Researcher

Eötvös Loránd University

Budapest, Hungary

Budapest Semesters in Mathematics (BSM)

Fall 2025

Relevant Coursework: Deep Learning, Graph Theory, Advanced Combinatorics (Hypergraph Theory)

RESEARCH INTERESTS

Mechanistic interpretability of neural networks; transformer theory and geometry; differential privacy in federated learning; mathematical foundations of deep learning (metastability, spectral methods, high-dimensional geometry).

MANUSCRIPTS IN PREPARATION

- Menon, V., et al. *Mechanistic Interpretability of Transformers on Simple Games*. Manuscript in preparation.

HONORS THESIS

- Menon, V., & Murray, R. *Metastable Dynamics in Transformer Particle Systems*. Undergraduate Honors Thesis, North Carolina State University, 2026.
Thesis Advisor: Ryan Murray (NCSU)

SELECTED ACADEMIC WRITING

- Menon, V. *Some Structures in Transformer Architectures*. Directed Reading Program paper, Budapest Semesters in Mathematics, January 2026. 16 pp.
drp.elte.hu/wordpress/wp-content/uploads/2026/01/Vasudev_Menon.pdf
- Menon, V., Gulhane, S., & Mehta, G. *Differential Privacy in Federated Learning on MIMIC-IV*. Course project report, 2024.
Advisors: Dr. Olivera Kotevska (ORNL), Dr. Mansoor Haider (NCSU)
[drive.google.com/...](https://drive.google.com/)

- **Mechanistic Interpretability & Simple Games.**

Budapest Semesters in Mathematics Undergraduate Research Symposium, Budapest, Hungary, December 11, 2025.

Co-presented with S. Reiter, J. Tran, Q. Tran, B. Whitsett, and P. Zsámboki.

Covered single-layer transformer theory for the (n, k) row-completion task; presented theoretical characterization of BOS, explicit attention solutions, experimental verification via SGD ensemble training, and truncated-SVD dimensionality reduction results.

RESEARCH EXPERIENCE

Metastable Dynamics in Transformer Particle Systems

2025–2026

Undergraduate Honors Thesis, North Carolina State University

Advisor: Ryan Murray (NCSU)

Co-authored thesis on metastable behavior in transformer self-attention dynamics on S^1 and S^2 . Characterized a universal three-phase evolution: rapid bipodal clustering, a prolonged metastable plateau, and consensus collapse. Proved via linearization that bipodal states are unstable saddle points with one positive eigenvalue $e^{-\gamma}$. Confirmed numerically that the metastable timescale scales logarithmically.

Mechanistic Interpretability of Transformers on Simple Games

Fall 2025

Budapest Semesters in Mathematics, Eötvös Loránd University

Advisor: Pál Zsámboki (Rényi Institute)

Designed and conducted interpretability research comparing hardcoded versus gradient-learned transformer solutions on controlled game tasks (the (n, k) row-completion / tic-tac-toe family). Established theoretical characterizations of minimal attention solutions; trained ensembles of 10^2 – 10^4 single-layer models via SGD; analyzed learned embedding geometry, attention structure, and failure modes at varying key dimensions d_k .

Some Structures in Transformer Architectures

Fall 2025

Budapest Semesters in Mathematics Directed Reading Program

Advisor: Attila Jung (ELTE)

Authored 16-page paper on transformer mechanistic interpretability. Expressed single-head self-attention as a Kronecker product operator, revealing separable sequence/feature structure. Decomposed attention into circuits, interpreting one-layer transformers as bigram + skip-trigram ensembles. Connected transformer superposition to Johnson–Lindenstrauss-style high-dimensional geometry.

Differential Privacy in Federated Learning (FL) on MIMIC-IV

2023–2024

North Carolina State University

Shrawani Gulhane, Gaurav Mehta, Dr. Olivera Kotevska (ORNL), Dr. Mansoor Haider (NCSU)

Implemented Distributed Differential Privacy (DDP) in federated learning using the Flower framework to protect sensitive healthcare data. Simulated heterogeneous clients using racial groupings from the MIMIC-IV ICU dataset. Analyzed privacy–accuracy trade-offs under Gaussian noise and gradient clipping. Processed 100GB+ of data using SQL and Power Query.

Batch Normalization in Deep Learning

May–August 2023

NSF-Funded Research Internship, North Carolina State University

Automated deep learning experiments via shell scripting and R to analyze the performance of batch normalization across neural network architectures. Trained and evaluated neural networks with statistical analysis and metric visualizations (precision, recall, runtime). Contributed to NSF-

funded research through experiment design and literature synthesis.

INDUSTRY EXPERIENCE

Fidelity Investments

Software / Machine Learning Intern

Durham, NC

June–August 2025

Engineered a Retrieval-Augmented Generation (RAG) system to automate translation of pension plan documentation into executable Java code. Built scalable ML infrastructure using AWS Bedrock, LangChain, HuggingFace, and Boto3. Designed a Neo4j graph database encoding pension rule relationships and dependencies for code generation. Optimized LLM accuracy through prompt engineering, validation pipelines, and context management. Deployed an end-to-end Streamlit application integrating the RAG pipeline, graph database, and code generation workflows.

TECHNICAL SKILLS

Programming Languages: Python, Java, C/C++, R, SQL

AI / ML: PyTorch, TensorFlow, Keras, Scikit-learn, HuggingFace, Federated Learning (Flower), Differential Privacy

Systems & Tools: Docker, AWS (Bedrock, Boto3), Git, CI/CD, Power Platform, Linux, REST APIs

Frameworks: FastAPI, Flask, Django, Streamlit

Other: Data Visualization (Matplotlib, Pandas), L^AT_EX