

Neha Verma

Email : nverma7@jhu.edu
Website: nverma1.github.io
GitHub : github.com/nverma1

EDUCATION

Johns Hopkins University Ph.D. in Computer Science – <i>Advisors: Kevin Duh & Kenton Murray</i>	Baltimore, MD August 2021–current
Johns Hopkins University M.S.E. in Computer Science – <i>Advisors: Kevin Duh & Kenton Murray</i>	Baltimore, MD August 2021–May 2023
Yale University B.S. in Computer Science & Mathematics – Distinction in the Major – Magna Cum Laude	New Haven, CT August 2017–May 2021

EXPERIENCE

Google DeepMind & YouTube Research Intern – <i>Hosts: Nikhil Mehta & Naijing Zhang</i>	Mountain View, CA May 2025–August 2025
Meta - Fundamental AI Research Research Intern – <i>Host: Maha Elbayad</i>	Menlo Park, CA May 2023–September 2023
Language and Information Learning Lab (LILY) at Yale Research Assistant – <i>Advisor: Dragomir Radev</i>	New Haven, CT May 2018–August 2021
CVS Health/Aetna Data Science Intern	Remote May 2020–August 2020
Summer Undergrad Math Research at Yale (SUMRY) Research Assistant – <i>Advisor: Karamatou Yacoubou Djima</i>	New Haven, CT May 2019–August 2019

PAPERS

- [1] **N. Verma**, K. Murray, K. Duh. “DOTResize: Reducing LLM Width via Discrete Optimal Transport-based Neuron Merging”, *preprint*, 2025.
- [2] **N. Verma**, K. Murray, K. Duh. “Merging feed-forward sublayers for compressed Transformers”, *preprint*, 2025.
- [3] **N. Verma**, M. Elbayad. “Merging text transformer models from different initializations”, to appear in *Transactions on Machine Learning Research (TMLR)*, 2024 and *Workshop on High-dimensional Learning Dynamics at ICML (HiLD@ICML)*, 2024.
- [4] **N. Verma**, K. Murray, K. Duh. “Exploring geometric representational disparities between multilingual and bilingual translation models”, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024.
- [5] E. Salesky, **N. Verma**, P. Koehn, M. Post. “Multilingual Pixel Representations for Translation and Effective Cross-lingual Transfer”, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.

- [6] H.L. Xinyuan*, **N. Verma***, B. Odoom, U. Pradeep, M. Wiesner, S. Khudanpur. “JHU IWSLT 2023 Multilingual Speech Translation System Description”, *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT)*, 2023.
- [7] A. Hussein, C. Xiao, **N. Verma**, T. Thebaud, M. Wiesner, S. Khudanpur. “JHU IWSLT 2023 Dialect Speech Translation System Description”, *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT)*, 2023.
- [8] K. Marchisio, **N. Verma**, K. Duh, P. Koehn. “*IsoVec*: Controlling the Relative Isomorphism of Word Embedding Spaces”, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- [9] **N. Verma**, K. Murray, K. Duh. “Strategies for Adapting Multilingual Pre-training for Domain-Specific Machine Translation”, *Proceedings of the 15th biennial conference of the Association for Machine Translation in the Americas (AMTA)*, 2022.
- [10] I. Li, J. Pan, J. Goldwasser, **N. Verma**, W. P. Wong, M. Nuzumlah, B. Rosand, Y. Li, M. Zhang, D. Chang, R. A. Taylor, H. M. Krumholz, D. Radev. “Neural Natural Language Processing for Unstructured Data in Electronic Health Records: a Review”, *Computer Science Review*, 2022.
- [11] L. Nan, C. Hsieh, Z. Mao, X. Lin, **N. Verma**, R. Zhang, W. Kryściński, N. Schoelkopf, R. Kong, X. Tang, M. Mutuma, B. Rosand, I. Trindade, R. Bandaru, J. Cunningham, C. Xiong, D. Radev. “FeTaQA: Free-form Table Question Answering”, *Transactions of the Association for Computational Linguistics (TACL)*, 2021.
- [12] L. Nan, D. Radev, R. Zhang, A. Rau, A. Sivaprasad, C. Hsieh, X. Tang, A. Vyas, **N. Verma**, P. Krishna, Y. Liu. “DART: Open-domain structured data record to text generation”, *Proceedings of the NAACL-HLT*, 2021.
- [13] R. Zhang, C. Westerfield, S. Shim, G. Bingham, A. Fabbri, W. Hu, **N. Verma**, and D. Radev. “Improving Low-Resource Cross-lingual document Retrieval by Reranking with Deep Bilingual Representations”, *Proceedings of the Association for Computational Linguistics (ACL)*, 2019.

SKILLS

- **Programming Languages:** Python, Bash, SQL
- **Frameworks/Toolkits:** PyTorch, Jax, fairseq, HuggingFace Transformers, numpy, Dask, L^AT_EX, Git

TEACHING

- **Teaching Assistant**, Johns Hopkins University
 - Introduction to Human Language Technology Fall 2025
- **Undergraduate Learning Assistant**, Yale University
 - Language and Computation (LING 227) Spring 2021
 - Abstract Algebra (MATH 350) Fall 2019 & 2020
 - Linear Algebra & Vector Calculus II (MATH 231) Spring 2019

- **Peer Tutor**, Yale University

- Fields & Galois Theory (MATH 370) Spring 2020 & 2021
- Calculus of Functions of Several Variables (MATH 120) Fall 2018

FELLOWSHIPS AND AWARDS

- Phi Beta Kappa 2021
- Yale College First-Year Summer Research Fellowship in the Sciences & Engineering 2018

GRADUATE COURSES

- Machine Translation, Advanced Natural Language Processing, Machine Learning, Machine Learning for Trustworthy AI, Self-Supervised Statistical Models, Parallel Computing for Data Science, Vision as Bayesian Inference, Information Extraction from Speech

SERVICE

- **Reviewing**
 - ICLR (2025, 2026)
 - NeurIPS (2025)
 - ACL Rolling Review (2023, 2024, 2025)
 - Computational Linguistics Journal (2022)
 - AMTA (2022)
- **Organizing**
 - Machine Translation Marathon of the Americas (MTMA) Organizer 2024
- **Outreach**
 - JHU CLSP Application Support Program Organizer 2022–2025
 - Skype a Scientist 2024
 - Regional NACLO Organizer @ Yale University 2020–2021
 - Yale Association for Women in Mathematics (AWM) Chapter President 2020–2021