ERIC LEHMAN

EDUCATION

Massachusetts Institute of Technology, Cambridge, MA   *September 2020 - May 2024*
Electrical Engineering & Computer Science
PhD in Computer Science

Northeastern University, Boston, MA   *September 2016 - May 2020*
Khoury College of Computer Science
Bachelor of Science in Computer Science

PUBLICATIONS

1. Adams, G., Fabbri, A., Ladhak, F., **E. Lehman**, and Elhadad, N. (2023). From sparse to dense: Gpt-4 summarization with chain of density prompting

2. **E. Lehman***, Zack*, T., Suzgun, M., Rodriguez, J. A., Celi, L. A., Gichoya, J., Jurafsky, D., Szolovits, P., Bates, D. W., Abdulnour, R.-E. E., Butte, A. J., and Alsentzer, E. (2023). Coding inequity: Assessing gpt-4's potential for perpetuating racial and gender biases in healthcare. *Lancet Digital Health*

3. **E. Lehman**, Hernandez, E., Mahajan, D., Wulff, J., Smith, M. J., Ziegler, Z., Nadler, D., Szolovits, P., Johnson, A., and Alsentzer, E. (2023). Do we still need clinical language models? *Conference on Health, Inference, and Learning (CHIL)*

4. **E. Lehman**, Jain, S., Pichotta, K., Goldberg, Y., and Wallace, B. C. (2021). Does bert pretrained on clinical notes reveal sensitive data? *North American Chapter of the Association for Computational Linguistics (NAACL).*

5. Nye, B. E., DeYoung, J., **E. Lehman**, Nenkova, A., Marshall, I. J., and Wallace, B. C. (2021). Understanding clinical trial reports: Extracting medical entities and their relations. *Proceedings of the American Medical Informatics Association*, 2021:485–494.

6. DeYoung*, J., Jain*, S., Rajani*, N. F., **E. Lehman**, Xiong, C., Socher, R., and Wallace, B. C. (2020). Eraser: A benchmark to evaluate rationalized nlp models. *Association for Computational Linguistics (ACL).*

7. **E. Lehman**, DeYoung, J., Barzilay, R., and Wallace, B. C. (2019). Inferring which medical treatments work from reports of clinical trials. *North American Chapter of the Association for Computational Linguistics (NAACL).*

8. **E. Lehman**, Krishnan, R. G., Zhao, X., Mark, R. G., and Lehman, L. H. (2018). Representation learning approaches to detect false arrhythmia alarms from ECG dynamics. *Machine Learning for Healthcare (MLHC).*

9. Ren, O., Johnson, A. E. W., **E. Lehman**, Komorowski, M., Aboab, J., Tang, F., Shahn, Z., Sow, D., Mark, R., and Lehman, L. H. (2018). Predicting and understanding unexpected respiratory decompensation in critical care using sparse and heterogeneous clinical data. *IEEE International Conference on Healthcare Informatics (ICHI).*

WORKSHOP PAPERS

10. **E. Lehman**, Lialin, V., Legaspi, K. E., Sy, A. J., Pile, P. T., Alberto, N. R., Ragasa, R. R., Puyat, C. V., Talino, M. K., Alberto, I. R., Alfonso, P. G., Moukheiber, D., Wallace, B., Rumshisky, A., Liang, J., Raghavan, P., Celi, L. A., and Szolovits, P. (2022). Learning to ask like a physician. *Proceedings of the 4th Clinical Natural Language Processing Workshop*, pages 74–86

11. Liang, J. J., **E. Lehman**, Iyengar, A. S., Mahajan, D., Raghavan, P., Chang, C. Y., and Szolovits, P. (2022). Towards generalizable methods for automating risk score calculation. *Proceedings of the 21st SIGBioMed Workshop on Biomedical Language Processing*

12. **E. Lehman***, DeYoung*, J., Nye, B., Marshall, I., and Wallace, B. C. (2020). Evidence inference 2.0: More data, better models. *Biomedical Natural Language Processing (BioNLP) Workshop.*

EXPERIENCE

Head of Clinical NLP *June 2022 - Present*
OpenEvidence

- Machine learning lead for building retrieval-augmented generation (RAG) technology that powers ClinicalKey AI (partnership with Elsevier Health) and OpenEvidence (6 million yearly physician queries)
- Managed and mentored team of six machine learning scientists and engineers to drive quality improvements of ClinicalKey AI and OpenEvidence

Research Assistant *September 2020 - May 2024*
EECS, Massachusetts Institute of Technology
Supervised by Professor Peter Szolovits

- Examined how in-domain pretraining and tokenization affects performance and efficiency of language models on clinical NLP tasks
- Quantified biases of GPT-4 towards certain demographic populations and how LLMs might disproportionately impact clinical care
- Investigated the efficacy of using masked language modeling, probing, and text generation approaches to extract protected health information from encoder-only models trained on sensitive data

Research Assistant *January 2018 - September 2020*
Khoury College of Computer Science, Northeastern University
Supervised by Professor Byron Wallace

- Designed and used multi-layer annotation pipeline for construction of a novel corpus for automatic extraction of results from clinical trial reports
- Trained hierarchical encoder-only transformer models for Population, Intervention, Comparator, and Outcome (PICO) extraction on PubMed full text reports

Machine Learning Researcher *June 2017 - August 2018*
IMES, Massachusetts Institute of Technology
Supervised by Distinguished Professor Roger Mark

- Developed and evaluated machine learning techniques intended to reduce the frequency of false ventricular tachycardia alarms

TEACHING EXPERIENCE

Teaching Assistant   *Jan 2023 - May 2023*

Machine Learning for Healthcare, MIT

- Designed course syllabus, homeworks, and final exam
- Gave guest lecture on state-of-the-art clinical NLP models

Teaching Assistant   *Jan 2017 - May 2017*

Fundamentals of Computer Science I, Northeastern University

- Tutored and held weekly office hours

ACADEMIC HONORS

- Best Paper Award at Conference on Health, Inference, and Learning (CHIL), 2023
- Computing Research Association Undergraduate Honorable Mention (2019)
- Computing Research Association Undergraduate Nominee (2018)
- Huntington 100 Award (2020): Honored among top 100 students at Northeastern for outstanding leadership, academic excellence, and community contributions.