

Yuting Guo, M.S.

✉ yguo262@emory.edu

🐦 @YutingGuo12

🌐 yuting-guo-53955b13b

🐙 <https://github.com/yguo0102>



Education

2021 – Now	📖 Emory University , Atlanta, GA, USA <i>Doctor of Philosophy in Computer Science</i>	GPA: 3.8/4
2019 – 2021	📖 Emory University , Atlanta, GA, USA <i>Master of Science in Computer Science</i>	GPA: 3.6/4
2012 – 2016	📖 Beijing Language and Culture University , Beijing, China <i>Bachelor of Computer Science and Technology</i>	GPA: 3.3/4
Coding Skills	📖 Python, Large Language Models (LLMs), Generative AI, PyTorch, Java, C++	

Research Projects

CHAMPS	📖 Applied LLMs in clinical narratives including verbal autopsy, clinical abstractions, and hospital summaries to predict the cause of death of neonates. 📖 Applied LLMs to recognize and redact personally identifiable information, including names, locations, and dates, in clinical narratives.
LLM	📖 Investigated and benchmarked different strategies to use GPT3.5, GPT4, SVM, and RoBERTa on health-related text classification tasks with social media text. 📖 Applied LLMs in Pubmed scientific publications to extract blood pressure values.
Transformer	📖 Transgender and natal sex classification based on clinical notes from GA, CA and MA. 📖 Detection of patients with Fontan, CHD and ASD based on clinical notes.

Research Presentations

Oral	📖 ALTA Workshop 2021 Virtual Benchmarking of Transformer-Based Pre-Trained Models on Social Media Text Classification Datasets 📖 AMIA Informatics Summit 2023 Seattle, WA, USA Generalizable Natural Language Processing Framework for Migraine Reporting from Social Media Automatic Detection of Intimate Partner Violence Victims from Social Media for Proactive Delivery of Support
Poster	📖 AMIA Annual Symposium 2023 New Orleans, LA, USA Overview of Shared Task 2: Multi-Class Classification of Sentiment Associated with Therapies in English Tweets AMIA Annual Symposium 2025 Atlanta, GA, USA Benchmarking Open-Source Large Language Models on Healthcare Text Classification Tasks

Honors & Awards

2025  BMI Award: Excellence in Research

Scholarly Works

Journal Articles

- 1 L. C. Ivey, C. L. Raskind-Hood, F. H. Rodriguez III, *et al.*, “Survival by heart failure phenotype in a fontan surveillance population, 2010–2023,” *Pediatric Cardiology*, pp. 1–8, 2026.
- 2 S. Das, Y. Ge, **Y. Guo**, *et al.*, “Two-layer retrieval-augmented generation framework for low-resource medical question answering using reddit data: Proof-of-concept study,” *J Med Internet Res*, vol. 27, e66220, Jan. 2025, ISSN: 1438-8871.  DOI: 10.2196/66220.
- 3 **Y. Guo**, “Leveraging few-shot learning and large language models for analyzing blood pressure variations across biological sex from scientific literature,” *Computers in Biology and Medicine*, 2025, Article reference: CBM_111128.
- 4 **Y. Guo**, H. Shi, W. M. Book, *et al.*, “Machine learning and natural language processing to improve classification of atrial septal defects in electronic health records,” *Birth Defects Research*, vol. 117, 3 Mar. 2025, ISSN: 2472-1727.  DOI: 10.1002/bdr2.2451.
- 5 **Y. Guo**, A. Ovadje, M. A. Al-Garadi, and A. Sarker, “Evaluating large language models for health-related text classification tasks with public social media data,” *Journal of the American Medical Informatics Association*, vol. 31, pp. 2181–2189, 10 Mar. 2024, ISSN: 1527-974X.  DOI: 10.1093/jamia/ocae210.
- 6 Y. Ge, **Y. Guo**, S. Das, M. A. Al-Garadi, and A. Sarker, “Few-shot learning for medical text: A review of advances, trends, and opportunities,” *Journal of Biomedical Informatics*, vol. 144, p. 104 458, 2023, ISSN: 1532-0464.  DOI: <https://doi.org/10.1016/j.jbi.2023.104458>.
- 7 **Y. Guo**, S. Das, S. Lakamana, and A. Sarker, “An aspect-level sentiment analysis dataset for therapies on Twitter,” *Data in Brief*, p. 109 618, 2023, ISSN: 2352-3409.  DOI: <https://doi.org/10.1016/j.dib.2023.109618>.
- 8 **Y. Guo**, M. A. Al-Garadi, W. M. Book, *et al.*, “Supervised text classification system detects fontan patients in electronic records with higher accuracy than ICD/i codes,” *Journal of the American Heart Association*, vol. 12, no. 13, Jul. 2023.  DOI: 10.1161/jaha.123.030046.
- 9 A. Sarker, S. Lakamana, **Y. Guo**, *et al.*, “# Chronicpain: Automated building of a chronic pain cohort from twitter using machine learning,” *Health Data Science*, vol. 3, p. 0078, 2023.
- 10 M. Al-Garadi, A. Sarker, **Y. Guo**, E. Warren, Y.-C. Yang, and S. Kim, “Automatic Identification of Intimate Partner Violence Victims From Social Media,” *Injury Prevention*, vol. 28, no. 1353-8047, A48–A48, 2022.  DOI: 10.1136/injuryprev-2022-SAVIR.123.
- 11 M. A. Al-Garadi, S. Kim, **Y. Guo**, *et al.*, “Natural Language Model for Automatic Identification of Intimate Partner Violence Reports from Twitter,” *Array*, p. 100 217, 2022.
- 12 M. A. Al-Garadi, Y.-C. Yang, **Y. Guo**, *et al.*, “Large-Scale Social Media Analysis Reveals Emotions Associated with Nonmedical Prescription Drug Use,” *Health Data Science*, vol. 2022, 2022.
- 13 **Y. Guo**, Y. Ge, Y.-C. Yang, M. A. Al-Garadi, and A. Sarker, “Comparison of Pretraining Models and Strategies for Health-Related Social Media Text Classification,” *Healthcare (Basel)*, vol. 10, no. 8, p. 1478, Aug. 2022.
- 14 S. Kim, M. Al-Garadi, **Y. Guo**, *et al.*, “Building a Social Media Platform for Surveillance and Support of Intimate Partner Violence and Its Victims During COVID-19 and Beyond: Preliminary Findings,” *Journal of Advanced Nursing*, pp. 26–27, 2022.

Conference Proceedings

- 1 **Y. Guo**, Y. Ge, and A. Sarker, "Detection of Medication Mentions and Medication Change Events in Clinical Notes Using Transformer-Based Models," in *MedInfo 2023*, 2023.
- 2 **Y. Guo**, S. Kim, E. Warren, Y.-C. Yang, L. Sahithi, and A. Sarker, "Automatic Detection of Intimate Partner Violence Victims from Social Media for Proactive Delivery of Support," in *AMIA 2023 Informatics Summit*, 2023.
- 3 **Y. Guo**, S. Rajwal, S. Lakamana, *et al.*, "Generalizable Natural Language Processing Framework for Migraine Reporting from Social Media," in *AMIA 2023 Informatics Summit*, 2023. DOI: 10.48550/ARXIV.2212.12454.
- 4 **Y. Guo** and A. Sarker, "SocBERT: A Pretrained Model for Social Media Text," in *The 2023 Workshop on Insights from Negative Results in NLP*, 2023.
- 5 J. M. Powell, **Y. Guo**, A. Sarker, and J. L. McKay, "Classification of fall types in parkinson's disease from self-report data using natural language processing," in *Artificial Intelligence in Medicine*, J. M. Juarez, M. Marcos, G. Stiglic, and A. Tucker, Eds., Cham: Springer Nature Switzerland, 2023, pp. 163–172, ISBN: 978-3-031-34344-5.
- 6 Y. Ge, **Y. Guo**, Y.-C. Yang, M. A. Al-Garadi, and A. Sarker, "A Comparison of Few-Shot and Traditional Named Entity Recognition Models for Medical Text," in *2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)*, IEEE, 2022, pp. 84–89.
- 7 D. Weissenbacher, J. Banda, V. Davydova, *et al.*, "Overview of the Seventh Social Media Mining for Health Applications (#SMM4H) Shared Tasks at COLING 2022," in *Proceedings of The Seventh Workshop on Social Media Mining for Health Applications, Workshop & Shared Task*, Association for Computational Linguistics, 2022, pp. 221–241. URL: <https://aclanthology.org/2022.smm4h-1.54>.
- 8 **Y. Guo** and J. Choi, "Enhancing Cognitive Models of Emotions with Representation Learning," in *The NAACL Workshop on Cognitive Modeling and Computational Linguistics 2021*, 2021.
- 9 **Y. Guo**, Y. Ge, M. A. Al-Garadi, and A. Sarker, "Pre-trained Transformer-based Classification and Span Detection Models for Social Media Health Applications," in *Proceedings of the Sixth Social Media Mining for Health (# SMM4H) Workshop and Shared Task*, 2021, pp. 52–57.
- 10 **Y. Guo**, Y. Ge, R. Liao, and A. Sarker, "An Ensemble Model for Automatic Grading of Evidence," in *Proceedings of the The 19th Annual Workshop of the Australasian Language Technology Association*, 2021, pp. 213–217.
- 11 **Y. Guo**, X. Dong, M. A. Al-Garadi, A. Sarker, C. Paris, and D. Mollá-Aliod, "Benchmarking of Transformer-Based Pre-Trained Models on Social Media Text Classification Datasets," in *Proceedings of the The 18th Annual Workshop of the Australasian Language Technology Association*, 2020, pp. 86–91.
- 12 **Y. Guo**, M. A. Al-Garadi, and A. Sarker, "Emory at WNUT-2020 Task 2: Combining Pretrained Deep Learning Models and Feature Enrichment for Informative Tweet Identification," in *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, 2020, pp. 388–393.