William Hunter McNichols (he/him)

wmcnichols@umass.edu linkedin.com/in/wmcnicho hunter-codes.com Google Scholar

Research Interests

My research centers on developing and studying large language model–based systems for classroom use, with an emphasis on intelligent tutoring and modeling student–LLM interactions. I work extensively with open-source LLMs and deep learning frameworks to prototype, deploy, and evaluate NLP systems at classroom scale. AI's capacity to help us learn fascinates me—as our learning, in turn, shapes AI into a better teacher.

Education

University of Massachusetts Amherst Amherst, MA
Ph.D. in Computer Science (Expected June 2026) *GPA: 3.95*

– Advisor: Prof. Andrew Lan, ML4ED Lab

University of Southern California Los Angeles, CA
M.S. Computer Science, 2016 *GPA: 3.64*

University of Southern California Los Angeles, CA
B.S. Computer Science, 2016 *GPA: 3.73, Magna Cum Laude*

Research Experience

NLP and Generative AI in Education University of Massachusetts Amherst : Sep 2022 – Present

– **The StudyChat Dataset: Student Dialogues with ChatGPT (Under Review, 2025)**
Built and deployed a classroom-scale LLM-based tutoring platform for an AI course for 300 students. Collected and analyzed 16K+ student dialogues to investigate student–LLM behaviors and their relationship to learning outcomes. Released the StudyChat dataset publicly to advance research on educational LLM interactions.
– **Algebra Error Classification with Large Language Models (AIED 2023)**
Developed a novel framework leveraging large language models (LLMs) to automatically classify algebraic errors from open-ended student responses. Conducted comparative analyses demonstrating improvements in accuracy and robustness over traditional rule-based methods.
– **Can LLMs Replicate ITS Feedback on Open-Ended Math Questions? (EDM 2024)**
Investigated the potential for LLMs to generate pedagogically valuable feedback traditionally provided by human-built Intelligent Tutoring Systems (ITS). Designed experiments and conducted qualitative analyses comparing model-generated feedback against expert annotations.
– **Automated Distractor Generation for Math MCQs Using LLMs (NAACL 2024)**
Explored multiple LLM-based methods, including fine-tuning, retrieval, and in-context learning, to generate plausible distractors for math multiple-choice questions. Conducted human evaluations and quantitative experiments comparing machine-generated distractors with educator-designed alternatives.

Industry Experience

Applied Scientist Intern, Software Application Inc. San Francisco : May 2025 – August 2025

– Built an LLM training and deployment pipeline within the company's software ecosystem.
– Mentored a junior research engineer on a distributed evaluation project, teaching principles of software design, experimental methodology, and applied machine learning.
– Demonstrated leadership through collaborative planning, project management, and the translation of research goals into mentorship opportunities.

Software Engineer, Siri Distributed Platform, Apple Inc. Cupertino, CA : 2016 – 2019

– Developed and maintained internal frameworks to support scalable Siri dialogue systems.
– Worked across platforms using Java, TypeScript, Python, and Objective-C
– Created internal tools and extensions for VSCode, significantly simplifying the developer experience.
– Modernized JavaScript runtime integration within Objective-C systems.
– Automated developer-facing workflows in the Siri personality pipeline, increasing throughput by over 500%.

Founder & Freelance Software Engineer, HunterCodes Remote : 2019 – 2022

- Founded and operated an independent software development business
- Led all aspects of product development from client discovery to implementation and deployment.
- Partnered with startups and small businesses to design and deploy full-stack web applications and data pipelines.
- Delivered custom, full-stack software solutions across a wide range of frameworks such as React and AWS

Publications

Tran, P., Feng, W., Sireci, S. G., **McNichols, H.**, and Lan, A. *Using Item Features to Calibrate Educational Test Items: Comparing Artificial Intelligence and Classical Approaches. American Journal of Education and Learning*, 10(2), 178–189, 2025.

Lee, J., **McNichols, H.**, and Lan, A. *Exploring Automated Keyword Mnemonics Generation with Large Language Models via Overgenerate-and-Rank.* Findings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024.

Feng, W.*, Lee, J.*, **McNichols, H.***, Scarlatos, A.*, Smith, D., Woodhead, S., Ornelas, N., and Lan, A. *Exploring Automated Distractor Generation for Math Multiple-Choice Questions via Large Language Models.* Findings of the Association for Computational Linguistics: NAACL, 2024.

McNichols, H., Lee, J., Fancsali, S., Ritter, S., and Lan, A. *Can Large Language Models Replicate ITS Feedback on Open-Ended Math Questions?* Proceedings of the 17th International Conference on Educational Data Mining (EDM), 2024.

McNichols, H.*, Feng, W.*, Lee, J.*, Scarlatos, A., Smith, D., Woodhead, S., and Lan, A. *Automated Distractor and Feedback Generation for Math Multiple-choice Questions via In-context Learning.* GAIED Workshop at NeurIPS, 2023.

Kumar, N. A., Feng, W., Lee, J., **McNichols, H.**, Ghosh, A., and Lan, A. *A Conceptual Model for End-to-End Causal Discovery in Knowledge Tracing.* Proceedings of the 16th International Conference on Educational Data Mining (EDM), 2023.

McNichols, H., Zhang, M., and Lan, A. *Algebra Error Classification with Large Language Models.* International Conference on Artificial Intelligence in Education (AIED), 2023.

*Denotes equal contribution.

Teaching Experience

University of Massachusetts Amherst Amherst, MA : 2022 – Present

- **CS 383 Artificial Intelligence (Fall 2024, Spring 2025) — Instructor**
  - Taught classrooms of 180 undergraduate students, significantly scaling a curriculum initially developed for smaller classrooms.
  - Developed extensive coursework, including new assignments and assessments, designed for independent student learning.
  - Managed a teaching team of graders and TAs; coordinated grading, feedback, and classroom logistics.
  - Conducted classroom research evaluating student interactions with GPT-powered Intelligent Tutoring Systems.
- **CS 240 Reasoning Under Uncertainty (Fall 2025) — Teaching Assistant**
  - Led discussion sections, held office hours, and reviewed student work.
- **CS 326 Web Programming (Fall 2022) — Teaching Assistant**
  - Provided instructional support, held office hours, and guided students through course projects.

City College of New York Manhattan, NY : Jan 2020 – Sep 2022

- **CS 44800 Artificial Intelligence — Lecturer**
  - Designed and taught an advanced undergraduate elective to roughly 30 students, emphasizing foundational AI concepts and recent industry trends.
  - Created hands-on coding assignments in Python, including Support Vector Machines, Neural Networks, and other machine learning algorithms.
  - Adapted coursework regularly to incorporate emerging technologies and reflect the latest industrial applications.

University of Southern California Los Angeles, CA : 2014

- **CS 101/102 Introduction to Programming — Undergraduate Teaching Assistant**
  - Assisted students in office hours, emphasizing coding best practices and foundational programming concepts.

Skills

Languages: Python, JavaScript, Java, Objective-C, C++, TypeScript

Frameworks: PyTorch, SciKit Learn, React, AWS

Pedagogical: Curriculum design, classroom scaling and coordination, instructional technology, learning analytics

Honors and Awards

NAEP Math Scoring Challenge (2023) – Grand Prize Winner

NeurIPS 2022 Causal Edu Competition - 3rd place

University of Southern California Trustee Scholar

Service and Leadership

Reviewer: NAACL Building Education Applications (BEA) Workshop 2024,2025

Independent Study Mentor, ML4ED Lab

HackUMass XII Judge