

Yang Shi

☎ +1 706 207 8251 • ✉ yshi26@ncsu.edu • 🌐 yshi.info

Education

- **North Carolina State University** **Raleigh, NC**
Ph. D., Computer Science (GPA 4.0), Graduate Merit Award *August, 2019 – May 2024*
Dissertation Committee: Thomas Price (Chair), Tiffany Barnes, Min Chi, Ken Koedinger (CMU), Eric Wiebe
Dissertation Title: Trustworthy Code-Informed Student Modeling for CS Education
Supervised Students: Poorvaja Penmetsa (21' MSCS NCSU), Krupal Shah (20' MCS NCSU)
- **University of Georgia** **Athens, GA**
M. S., Computer Science, Graduate Assistantship *August, 2015 – December, 2017*
Thesis Committee: Wenzhan Song (Chair), Hamid Arabnia, Tianming Liu
- **Central South University** **Changsha, China**
B. Eng., Automation, Undergraduate Excellence Award *September, 2011 – June, 2015*
Thesis Advisor: Fang Liu

Employment

- **Research Intern**
Microsoft Research, Redmond, WA. *March 2022 – August 2022*
- **Research Scientist Intern**
Stratifyd Inc., Charlotte, NC. *May 2019 – August 2019*

Awards and Funding

Personal Awards and Certifications.....

- **Inclusive Teaching Certificate, NCSU** (*In Progress, September 2022*)
- **Preparing the Professoriate, NCSU** (*In Progress, September 2022*)
- **Best Paper Award, EDM 2021** (*July 2021*)
- **Summer Graduate Merit Award, NCSU** (*May 2021*)
- **AWS Cloud Credit Award for Research, Amazon AWS** (*Spring 2020*)
- **Graduate School Travel Award, University of Georgia** (*Summer 2018*)

Research Funding.....

- NSF #2013502: Generalizing Data-Driven Technologies to Improve Individualized STEM Instruction by Intelligent Tutors, Improving Undergraduate STEM Education (IUSE), Graduate Research Assistant, (*Fall 2020 - Spring 2021, Fall 2022*)
- Department of Labor: Artificial Intelligence Academy (AIA): North Carolina's Apprenticeships for Innovation, Course Instructor, (*Fall 2021 - Spring 2022*)

Publications (All Peer Reviewed)

(*Italics are supervised Master students*)

Book Chapters.....

- [B1 (In Press)]. **Continuous Student Modeling for Programming in the Classroom: Challenges, Methods, and Evaluation.**
Ye Mao, Samiha Marwan, Preya Shabrina, Yang Shi, Thomas Price, Min Chi and Tiffany Barnes
Handbook of Artificial Intelligence In Education, 2023.

Journal Papers.....

- [J2]. **Enhanced Cyber-physical Security in Internet of Things through Energy Auditing.**
Fangyu Li, Yang Shi, Shinde Aditya, Jin Ye and WenZhan Song
IEEE Internet of Things Journal. 2019.
- [J1]. **System Statistics Learning based IoT Security: Feasibility and Suitability.**
Fangyu Li, Shinde Aditya, Yang Shi, WenZhan Song and Xiang-Yang Li
IEEE Internet of Things Journal. 2019.

Conference Papers (All oral presentation).....

- [C8]. **Code-DKT: A Code-based Knowledge Tracing Model for Programming Tasks.**
Yang Shi, Min Chi, Tiffany Barnes and Thomas Price
The 15th International Conference on Educational Data Mining (EDM 2022), 2022. (Acceptance Rate 28.6%, 26/91 Full Papers)
- [C7]. **Identifying Common Errors in Open-ended Machine Learning Projects.**
James Skripchuk, Yang Shi and Thomas Price
The 53rd ACM Technical Symposium on Computing Science Education (SIGCSE 2022), 2022. (Acceptance Rate 27.9%, 144/516 Full Papers)
- [C6]. **More With Less: Exploring How to Use Deep Learning Effectively through Semi-supervised Learning for Automatic Bug Detection in Student Code.**
Yang Shi*, Ye Mao*, Tiffany Barnes, Min Chi and Thomas Price
The 14th International Conference on Educational Data Mining (EDM 2021), 2021. (Combined Acceptance Rate 27.2%, 44/162 Short Papers)
- [C5]. **Knowing both when and where: Temporal-ASTNN for Early Prediction of Student Success in Novice Programming Tasks.**
Ye Mao, **Yang Shi**, Samiha Marwan, Thomas Price, Tiffany Barnes and Min Chi
The 14th International Conference on Educational Data Mining (EDM 2021), 2021. (Acceptance Rate 22%, 22/100 Full Papers)
- [C4]. **Just a Few Expert Constraints Can Help: Humanizing Data-Driven Subgoal Detection for Novice Programming.**
Samiha Marwan, Yang Shi, Ian Menezes, Min Chi, Tiffany Barnes and Thomas Price
The 14th International Conference on Educational Data Mining (EDM 2021), 2021. (Acceptance Rate 22%, 22/100 Full Papers, **Best Paper Award**)
- [C3]. **Toward Semi-Automatic Misconception Discovery Using Code Embeddings.**
Yang Shi, *Krupal Shah*, Wengran Wang, Samiha Marwan, *Poorvaja Penmetsa* and Thomas Price
The 11th International Conference on Learning Analytics & Knowledge (LAK 21), 2021. (Acceptance Rate 29.3%, 29/99 Short Papers)
- [C2]. **Energy Audition based Cyber-Physical Attack Detection System in IoT.**
Yang Shi, Fangyu Li, Wenzhan Song, Xiang-Yang Li and Jin Ye
ACM SigMobile China, 2019.
- [C1]. **Dynamic Time-frequency Feature Extraction for Brain Activity Recognition.**
Yang Shi, Fangyu Li, Tianming Liu, Fred R. Beye and WenZhan Song
2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018.

Workshop Papers.....

- [W3]. **Investigate Effectiveness of Code Features in Knowledge Tracing Task on Novice Programming Course.**
Poorvaja Penmetsa, Yang Shi and Thomas Price
Work-In-Progress Track, CSEDM Workshop @ EDM'21. 2021.
- [W2]. **TEST_POSITIVE at W-NUT 2020 Shared Task-3: Joint Event Multi-task Learning for Slot**

Filling in Noisy Text.

Chacha Chen, Chieh-Yang Huang, Yaqi Hou, Yang Shi, Enyan Dai and Jiaqi Wang
2020 The 6th Workshop on Noisy User-generated Text (W-NUT) at EMNLP 2020. 2020.

- [W1]. **Comparing Feature Engineering Approaches to Predict Complex Programming Behaviors.** Wengran Wang, Yudong Rao, Yang Shi, Alexandra Milliken, Chris Martens, Tiffany Barnes and Thomas W Price
CSEDM Workshop @ EDM'20. 2020.

Research Presentations

- **Code-DKT: A Code-based Knowledge Tracing Model for Programming Tasks.** Department of Computer Science, UNC Charlotte. July 2022. (Invited Presentation)
- **Toward Semi-Automatic Misconception Discovery Using Code Embeddings.** Department of Computer Science, North Carolina State University. April 2021.

Teaching Experience

Instructor.....

North Carolina State University

- CSC 522 Automated Learning and Data Analysis *Spring 2023*
- AI Academy Data Mining *Fall 2021*
A remote course funded by the Department of Labor under the AI Academy project at NC State University, designed for teaching professionals AI skills from various backgrounds. The course introduces a wide spectrum of concepts in data mining, including exploratory data analysis, and all kinds of algorithms for data mining, including supervised learning (Naive Bayes, Neural Networks, Support Vector Machine, etc.) and unsupervised learning (Dimension Reduction, Clustering, etc.). Each class is followed by a hands-on practice section led by TAs to introduce how to write Python code for specific concepts using Scikit-learn.
- CSC 116 Intro to Programming – Java *Summer 2020, Summer 2021*
I taught CSC 116, Introduction to Java course over the summer semesters of 2020 and 2021. This is the first introductory course of CS and non-CS majors for computer programming. The class size has been about 60 students, while the summer semesters are three months. The course content covers the basics of Java programming, including math operations, logic, loops, objects, and file management.

Teaching Assistant.....

North Carolina State University

- CSC 422/522 Automated Learning and Data Analysis *Spring 2020*
- CSC 510 Software Engineering *Fall 2019*

University of Georgia

- CSCI 3360 Data Science I *Spring 2018, Spring 2017*
- CSCI/PHIL 4550/6550 Introduction to Artificial Intelligence *Fall 2016*

Academic Services

Workshop Organizer/Co-Organizer/Chair.....

- [O4]. Bitu Akram, Thomas Price, **Yang Shi**, Peter Brusilovski, Sharon I-Han Hsiao, Juho Leinonen: Educational Data Mining in Computer Science Education (CSEDM) Workshop, LAK'23. 2023.
- [O3]. Bitu Akram, Thomas Price, **Yang Shi**, Peter Brusilovski, Sharon I-Han Hsiao: Educational Data Mining in Computer Science Education (CSEDM) Workshop, EDM'22. 2022.
- [O2]. Bitu Akram, Thomas Price, **Yang Shi**, Peter Brusilovski, Sharon I-Han Hsiao: Educational Data Mining in Computer Science Education (CSEDM) Workshop, EDM'21. 2021.
- [O1]. Thomas Price, Peter Brusilovski, Sharon I-Han Hsiao, Ken Koedinger, **Yang Shi**: Educational

Data Mining in Computer Science Education (CSEDM) Workshop, EDM'20. 2020.

Journal Reviewer.....

- Journal of Educational Data Mining (2022)
- Research and Practice in Technology Enhanced Learning (2022)
- IEEE Internet of Things Journal (2018 - 2022)
- IEEE Transactions on Network and Service Management (2020 - 2021)
- IEEE Transactions on Signal and Information Processing over Networks (2020)
- IEEE Open Journal of the Computer Society (2020)
- IEEE Access (2020)

Conference Program Committee/Reviewer.....

- The International Conference on Learning Analytics & Knowledge (LAK) (2023) (PC)
- Symposium on Educational Advances in Artificial Intelligence (EAAI) (2023) (PC)
- Technical Symposium on Computer Science Education (SIGCSE) (2023) (PC)
- AAAI Conference on Artificial Intelligence (AAAI) (2023) (PC)
- International Conference on Educational Data Mining (EDM) (2022) (PC)
- ACM Innovation and Technology in Computer Science Education (ITiCSE) (2022) (PC)
- ACM CHI Conference on Human Factors in Computing Systems (2022) (Reviewer)

Conference Subreviewer:.....

SPLICE workshop, ACM ICER, AIED, IEEE INFOCOM, IEEE FWC

Professional Affiliations

IEEE Student Member, ACM Student Member.