Wenxin Jiang

Ph.D. Student
Elmore Family School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47906

jiang784@purdue.edu
765–409–1715
https://wenxin-jiang.github.io

RESEARCH THEME

My research interests lie in *Software Engineering for Machine Learning*, especially the reusability, reproducibility, and trustworthiness of deep learning pre-trained models (PTMs).

EDUCATION

Ph.D, Electrical and Computer Engineering, *GPA: 3.7/4.0* 2020–present
Purdue University, West Lafayette, IN

B.Sc. Applied Physics, *GPA: 3.3/4.0* 2016–2020
Southeast University, Jiangsu, China

Study Abroad Program, Engineering Physics, *GPA: 3.8/4.0* 2019
University of California, Santa Barbara, CA

PROFESSIONAL EXPERIENCE

Graduate Research Assistant 2021–present
Purdue University — Advised by James C. Davis

TensorFlow Model Developer 2021–present
Purdue University × Google

REFEREED CONFERENCE PUBLICATIONS

[1] **Jiang**, Synovic, Hyatt, Schorlemmer, Sethi, Lu, Thiruvathukal, and Davis. *An Empirical Study of Pre-Trained Model Reuse in the Hugging Face Deep Learning Model Registry*. Proceedings of the ACM/IEEE 45th International Conference on Software Engineering (**ICSE'23**).

[2] **Jiang**, Synovic, Jajal, Schorlemmer, Tewari, Pareek, Thiruvathukal, and Davis. *PTMTorrent: A Dataset for Mining Open-source Pre-trained Model Packages*. Proceedings of the 20th Annual Conference on Mining Software Repositories — Data and Tool Showcase Track (**MSR-Data'23**).

[3] Davis, Jajal, **Jiang**, Schorlemmer, N. Synovic, and G.K. Thiruvathukal. *Reusing Deep Learning Models Challenges and Directions in Software Engineering*. Proceedings of the IEEE John Vincent Atanasoff Symposium on Modern Computing (**JVA'23**).

[4] Montes, Peerapatanapokin, Schultz, Guo, **Jiang**, and Davis. *Discrepancies among Pre-trained Deep Neural Networks: A New Threat to Model Zoo Reliability*. Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering — Ideas, Visions, and Reflections track (**ESEC/FSE-IVR'22**).

REFEREED WORKSHOPS, DEMONSTRATIONS, AND COMPETITIONS

[1] **Jiang**, Synovic, Sethi, Indarapu, Hyatt, Schorlemmer, Thiruvathukal, and Davis. *An Empirical Study of Artifacts and Security Risks in the Pre-trained Model Supply Chain*. Proceedings of the 1st ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (**SCORED'22**).

[2] Synovic, Hyatt, Sethi, Thota, Shilpika, Miller, **Jiang**, Amobi, Pinderski, Laufer, Hayward, Kingensmith, Davis, and Thiruvathukal. *Snapshot Metrics Are Not Enough: Analyzing Software Repositories with Longitudinal Metrics*. Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering — Demonstrations track (**ASE-Tool Demonstrations'22**).

[3] Veselsky, West, Ahlgren, Thiruvathukal, Klingensmith, Goel, **Jiang**, Davis, Lee, and Kim. *Establishing trust in vehicle-to-vehicle coordination: a sensor fusion approach*. Proceedings of the 23rd Annual International Workshop on Mobile Computing Systems and Application (**HotMobile'22**).

TECHNICAL REPORTS

[1] **Jiang**, Jones, Yasmin, Synovic, Sashti, Chen, Thiruvathukal, Yuan, and Davis. *PeaTMOSS: Mining Pre-Trained Models in Open-Source Software*. `https://arxiv.org/pdf/2310.03620`. 2023.

[2] **Jiang**, Cheung, Thiruvathukal, and Davis. *Exploring Naming Conventions (and Defects) of Pre-trained Deep Learning Models in Hugging Face and Other Model Hubs*. `https://arxiv.org/pdf/2310.01642`. 2023.

[3] Jajal, **Jiang**, Tewari, Woo, Lu, Thiruvathukal, and Davis. *Analysis of Failures and Risks in Deep Learning Model Converters: A Case Study in the ONNX Ecosystem*. `https://arxiv.org/abs/2303.17708`. 2023.

[4] **Jiang**, Banna, Vivek, Goel, Synovic, Klingensmith, Thiruvathukal, and Davis. *Challenges and Practices of Deep Learning Model Reengineering: A Case Study on Computer Vision*. `https://arxiv.org/abs/2303.07476`. 2023.

[5] Banna, Chinnakotla, Yan, Vegesana, Vivek, Krishnappa, **Jiang**, Lu, Thiruvathukal, and Davis. *An Experience Report on Machine Learning Reproducibility: Guidance for Practitioners and TensorFlow Model Garden Contributors*. `https://arxiv.org/abs/2107.00821`. 2021.

POSTERS

[1] Schorlemmer, **Jiang**, and Davis. *Machine Learning Supply Chain Security*. 2023 Purdue CERIAS Symposium (**CERIAS'23**). ***Award: Best Poster — 2nd-place***.

[2] **Jiang**, Schorlemmer, and Davis. *Trustworthy Re-use of Pre-trained Neural Networks*. 2023 Purdue CERIAS Symposium (**CERIAS'23**).

TEACHING ASSISTANT

ECE 595 – Advanced Software Engineering Spring 2022
Purdue University

INVITED TALKS

An Empirical Study of Pre-Trained Model Reuse in the Hugging Face Deep Learning Model Registry 2023
Purdue University Programming Languages Group, Seminar

Deep Learning Model Reengineering: An Exploratory Case Study on Computer Vision 2022
Purdue University Programming Languages Group, Seminar

AWARDS AND RECOGNITION

ACM SIGSOFT CAPS Travel Grant (ICSE'23)	2023
Purdue Graduate Student Government and the Graduate School Travel Grant (ICSE'23)	2023
ACM SIGSOFT CAPS Travel Grant (ESEC/FSE'22)	2022
Study Abroad Fellowship, Southeast University	2019
Second prize, Vision Guided Robot Competition, Southeast University	2019
Distinction Award, Southeast University	2018
Third prize, Structual Innovation Invitation Competition, Southeast University	2017

ACTIVITIES AS A REFEREE

Sub-Reviewer: LCTES'23, ESEC/FSE'23, ASE'22	2023

PROFESSIONAL MEMBERSHIPS

Student member, Association for Computing Machinery (ACM)

Student member, Institute of Electrical and Electronics Engineers (IEEE)