Kewen Peng

Cell: 336-251-9877 |**Email:** kpeng@ncsu.edu | **GitHub:** kpeng2019 | **ResearchGate**

Education

North Carolina State University — Raleigh, NC
Ph.D. in Computer Science | Advisor: Dr. Tim Menzies | Lab: RAISE Lab — Aug. 2019 – Feb. 2024

Wake Forest University — Winston-Salem, NC
B.S. in Computer Science | B.A. in Mathematics — Aug. 2015 – May 2019

Skills

General Expertise: Deep Learning, Machine Learning Fairness, Supervised/Unsupervised Modeling, Optimization
Programming: Python (5+ years), Java/C/C++ (3+ years), JavaScript, SQL
DevOps: Docker, Jenkins, Ansible, Travis-CI, Vagrant
Coursework: DevOps | Data Structure | Algorithm | Data Mining | Advanced AI | Automated SE

Work Experience

Data Scientist Intern — May. 2022 – Aug. 2022
Indeed Inc. — Remote

- Owned the recommendation models (multi-label logistic regression) in 11 new international markets using the ETL pipeline (Airflow); Obtained significant improvement by over 12%.
- Conducted data analysis using Python for economic trade-offs in the job seeker recommendation system.
- Managed the team-owned data index in Java; Developed a new weighting strategy for hierarchical classification.

Research Assistant — Jan. 2020 – Present
North Carolina State University, RAISE Lab — Raleigh, NC

- Proposed SOTA algorithms aiming for more explainable and fairness-aware machine learning software. Published papers in peer-reviewed journals and conferences.
- Led experiments with Active Learning and Curriculum Learning tactics in the NLP domain (both classification and generative tasks). Proposed a router-based pipeline that reduces training and inference costs for deep learning models (e.g., BERT, LSTM) to 10%.

Selected Research Experience

Machine Learning Fairness — Oct. 2019 – Present
RAISE Lab, research project, funded by LAS and NSF — Raleigh, NC

- Explored fairer results in machine learning software via different approaches of bias mitigation. Achieved significant reduction in bias by 67% without compromising model performance.
- Explored reliable and robust explanation generation tools for fairer SE. Designed the recommendation system that enhances explanation generation with actionable analysis.

Taming Deep Learning — Jan. 2021 – May. 2023
RAISE Lab, research project, funded by LAS — Raleigh, NC

- Designed a model-agnostic method (via Keras) to reduce the training time of deep learning models to 50% on NetFlow attack detection data.
- Designed a router-based method (via Keras) to reduce the inference cost of deep learning models to 10% on NLP tasks, such as emotion classification and semantic textual similarity.

Persona and Knowledge Empowered Conversational AI — Jan. 2023 – May. 2023
Research project — Raleigh, NC

- Designed and implemented a social-chat conversational model, which generates personalized answers to a user's question based on the user's persona information and knowledge base.
- Implemented hierarchical cross attention in grounding phases to locate context-relevant candidates.
- Implemented in PyTorch, experimented with BART and GPT-2, the new approach reaches comparable performance to baseline methods and obtained 20% reduction in training and inference cost.

Publications

- Luigi Ferraro, Ellen Kirkman, W Frank Moore, **Kewen Peng**, *On the Noether Bound for Noncommutative Rings*, **PAMS journal** (Accepted).
- Joymallya Chakraborty, **Kewen Peng**, Tim Menzies, *Making fair ML software using trustworthy explanation*, **ESEC/FSE 2020** (Accepted).
- **Kewen Peng**, Tim Menzies, *Defect Reduction Planning (using TimeLIME)*, **TSE journal** (Accepted).
- **Kewen Peng**, Christian Kaltenecker, Norbert Siegmund, Sven Apel, Tim Menzies, *VEER: Disagreement-Free Multi-objective Configuration*, **EMSE journal** (Accepted).
- **Kewen Peng**, Joymallya Chakraborty, Tim Menzies, *FairMask: Better Fairness via Model-based Rebalancing of Protected Attributes*, **TSE journal** (Accepted).
- Tim Menzies, **Kewen Peng**, Andre Lustosa, *Fairer Software Made Easier (using "Keys")*, **ASE 2021 RAISE Workshop** (Accepted).
- Junfeng Liu, Zhuocheng Mei, **Kewen Peng**, Ranga Raju Vatsavai *Context Retrieval via Normalized Contextual Latent Interaction for Conversational Agent*, **IEEE ICDM 2023 WAIN** (Accepted).

Honors and Service

- Member in Upsilon Pi Epsilon Honor Society, 2017
- Honorable Mention in ICPC Mid-Atlantic Regional, 2018
- Honorable Mention in COMAP MCM Contest, 2018
- Undergraduate Summer Research Fellowship, Wake Forest University, 2018
- Journal Reviewer, Empirical Software Engineering (EMSE), 2021
- Keynote Speaker, International Conference on Software Engineering (ICSE), 2022
- Web Chair, International Conference on Software Engineering (ICSE) Fairware Workshop, 2023
- Journal Reviewer, Transaction on Software Engineering (TSE), 2023