Tianpei Xia

http://xiatianpei.com • 360-808-3650 • txia4@ncsu.edu • arennax

EDUCATION

North Carolina State University, Raleigh, NC — Aug. 2016 - Dec. 2020
Ph.D. in Computer Science | Adviser: Dr. Tim Menzies

The University of Texas at Dallas, Richardson, TX — Aug. 2013 - Dec. 2015
M.S. in Computer Science

Nanjing University of Posts and Telecom., Nanjing, China — Sep. 2009 - Jun. 2013
B.S. in Electrical Engineering

SKILLS AND INTERESTS

• Experience in machine learning, software analytics, hyperparameter optimization and software development.
• Proficient in *Python*, familiar with *Java and R*, good at ML tools: Scikit-learn, Pytorch, Tensorflow.
• Interested in back-end/infrastructure development as well as machine learning and computer vision engineering and research positions.

SELECTED PROJECTS

NSF Funded: Search-based Software Engineering Research Aug, 2017 - Present
Research Assistant Under Dr. Tim Menzies, North Carolina State University, USA

- **Evolutionary Algorithms for Hyperparameter Optimization:** Developed a hyperparameter optimization architecture called OIL (Optimized Inductive Learning), which applied evolutionary algorithms (e.g. Differential Evolution and NSGA-II) for software analytic tasks. OIL is tested on a wide range of optimizers using data from 945 projects. In experiment, OIL improved the performance of effort estimation in accuracy (won 16 out of 18 cases) and efficiency (reduced runtime from days to hours).
- **Sequential Model Optimization for Software Effort Estimation:** Applied and developed a sequential model based method (also known in the machine learning literature as an active learner) named "FLASH" for the first time in software effort estimation domain. With the constraints of iterations and explorations, FLASH can efficiently find good configurations for machine learning algorithms (e.g. CART), then improve the performance of software effort estimation tasks (Accuracy is improved by 11% on average).
- **Effort Estimation for Agile Projects on GitHub:** Developed and applied a software effort estimation architecture named ROME (Rapid Optimization Methods for Effort-estimation) on massive agile project datasets. The data is collected on web-based repository hosting service named GitHub by using its APIs and self-defined feature selection module. In experiment, ROME can achieve comparable performance with much less computing resources.

System migration for educational computer programming game Jan, 2017 - May, 2017
Research Project, Game2Learn Lab, North Carolina State University, USA

- **Game Platform Migration:** Helped to migrate an educational purpose programming game named "BOTS" from its original developing platform "Unity 4" to "Unity 5" by using JavaScript. After migration, more potential features are enabled for game's future extension and development. BOTS is an serious puzzle game designed to teach programming fundamentals for novice computer users.

Satellite images change detection by using Gaussian Mixture Model Jan, 2017 - May, 2017
Graduate Course Project, North Carolina State University, USA

- **Image Change Detection:** Applied and developed a Gaussian mixture model to identify landscape changes using high resolution satellite images. This grid-based method has competitive performance in Bi-temporal change detection. Given two very high resolution satellite images from the same landscape area, it can achieve the same level of change detecting ability as human eyes.

Design, develop, and test a Laboratory Management System Aug, 2016 - Dec, 2016
Graduate Course Project, North Carolina State University, USA

- **Database Design:** Design, develop, and test ABC Laboratory Management System using MySQL. The system can support up to 400 people to do experiments, workshops and financial transaction based on combined database. The project is in four parts: conceptual design and requirements analysis (Phase I), database design requirements (Phase II), normalization (Phase III), and final demonstration (Phase IV).

ACADEMIC EXPERIMENCE

Teaching Assistant | **NC State, Raleigh, NC**

- CSC540 Database Management Concepts and Systems — Aug, 2016 - Dec, 2016
- CSC226 Discrete Mathematics for Computer Scientists — Jan, 2017 - May, 2017
- CSC236 Concepts and Facilities of Operating Systems — Aug, 2017 - Dec, 2017
- CSC495 Programming Languages and Modeling — Jan, 2018 - May, 2018

SELECTED PUBLICATIONS

- **Tianpei Xia**, Rahul Krishna, Jianfeng Chen, George Mathew, Xipeng Shen, Tim Menzies, *Hyperparameter Optimization for Effort Estimation*. **EMSE(Under Review)**, 2018. → tiny.cc/txiaHyper.
- **Tianpei Xia**, Rui Shu, Tim Menzies, *Sequential Model Optimization for Effort Estimation of Contemporary Software Projects*. **In preparation**, 2019.
- Rui Shu, **Tianpei Xia**, Tim Menzies, *Application of Data Balancing and Hyperparameter Optimization in Security Bug Report Prediction*. **In preparation**, 2019.