Rahul Yedida

Website :: GitHub :: LinkedIn :: Google Scholar

hello@ryedida.me
+1 (206) 660-7542

Education

North Carolina State University — Raleigh, NC
Ph.D. Computer Science – Advisor: Dr. Tim Menzies — *Aug 2019 - Jul 2024*

PES University — Bangalore, India
B.E. Computer Science — *Aug 2015 - May 2019*

Employment

LexisNexis Legal & Professional — Raleigh, NC
Senior Data Scientist I — *May 2024 - Present*

- **Performance improvements:** Improved customer-facing product runtime by 24.8% through various inefficiency fixes based on profiling results.
- **Drafting improvements:** Helped improve complaint drafting results by 46 percentage points of usefulness through various improvements
- **Workflow improvement:** Led initiative to use Bayesian optimization for prompt tuning.
- **Technology:** Python, Litestar, React, TypeScript, Tailwind

Amazon — New York, NY / Bellevue, WA
Software Dev Engineer Intern — *May 2023 - Aug 2023*

- Implemented profile locks for Prime Video on Echo Show devices.
- **Technology:** React Native, TypeScript

Software Dev Engineer Intern — *May 2022 - Jul 2022*

- Developed a full-stack system to publish announcements in scorecards used by delivery service partners (DSPs).
- **Technology:** React/Redux, TypeScript, Redux Saga, DynamoDB, Java Spring

North Carolina State University — Raleigh, NC
PhD Student — *Aug 2019 - Jul 2024*

- **State-of-the-art hyper-parameter optimization:** Proposed a novel hyper-parameter optimization method that outperforms prior work and is 200-700% faster.
- **Better, faster deep learning for SE:** Improved defect prediction by up to 123% (F-1 score), code smell detection by up to 30% (AUC), issue lifetime prediction by up to 76% (accuracy), automated microservice partitioning by up to 285% (modularity)
- **Semi-supervised learning:** Achieved state-of-the-art results (up to 100% improvement in AUC) on static code warnings analysis using 10% of the labels.
- **Teaching:** Teaching assistant for 830 students in total, over 5 semesters, for CSC 230 (C and Software Tools), CSC 510 (Software Engineering), and CSC 591/791 (Automated Software Engineering)

Recent Publications

See full list on Google Scholar.

1. Baldassarre, M. T., Ernst, N., Hermann, B., Menzies, T., & **Yedida, R.** (2023). (Re)use of Research Results (is Rampant). *Communications of the ACM, 66(2), 75-81.*

2. **Yedida, R.**, Kang, H. J., Tu, K., Lo, D., & Menzies, T. (2023). How to Find Actionable Static Analysis Warnings: A Case Study with FindBugs. *IEEE Transactions on Software Engineering, (01), 1-17.*

3. **Yedida, R.**, Krishna, R., Kalia, A., Menzies, T., Xiao, J., & Vukovic, M. (2023). An Expert System for Redesigning Software for Cloud Applications. *Expert Systems with Applications.*

4. **Yedida, R.**, Menzies, T. (2022). How to Improve Deep Learning for Software Analytics (a case study with code smell detection). In *2022 IEEE/ACM 19th International Conference on Mining Software Repositories (MSR). IEEE, 2022.*

5. **Yedida, R.**, & Menzies, T. (2021). On the Value of Oversampling for Deep Learning in Software Defect Prediction. *IEEE Transactions on Software Engineering, doi: 10.1109/TSE.2021.3079841*

6. Agrawal, A., Yang, X., Agrawal, R., **Yedida, R.**, Shen, X., & Menzies, T. (2021). Simpler Hyperparameter Optimization for Software Analytics: Why, How, When?. *IEEE Transactions on Software Engineering, doi: 10.1109/TSE.2021.3073242*

7. Yang, X., Chen, J., **Yedida, R.**, Yu, Z., & Menzies, T. (2021). Learning to recognize actionable static code warnings (is intrinsically easy). *Empirical Software Engineering, 26(3), 1-24.*

8. **Yedida, R.**, Krishna, R., Kalia, A., Menzies, T., Xiao, J., & Vukovic, M. (2021). Lessons learned from hyper-parameter tuning for microservice candidate identification. *Proceedings of the thirty-sixth IEEE/ACM International Conference on Automated Software Engineering (ASE)*

9. **Yedida, R.**, & Saha, S. (2021). Beginning with Machine Learning: A Comprehensive Primer. *The European Physical Journal Special Topics*, 230(10), 2363-2444.

10. Saha, S., Nagaraj, N., Mathur, A., **Yedida, R.**, & Sneha, H. R. (2020). Evolution of novel activation functions in neural network training for astronomy data: habitability classification of exoplanets. *The European Physical Journal Special Topics, 229(16), 2629-2738.*

11. **Yedida, R.**, Saha, S., & Prashanth, T. (2020). LipschitzLR: Using theoretically computed adaptive learning rates for fast convergence. *Applied Intelligence, 1-19.*

12. Sridhar, S., Saha, S., Shaikh, A., **Yedida, R.**, & Saha, S. (2020, July). Parsimonious Computing: A Minority Training Regime for Effective Prediction in Large Microarray Expression Data Sets. In *2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.*

13. Khaidem, L., **Yedida, R.**, & Theophilus, A. J. (2019, November). Optimizing Inter-nationality of Journals: A Classical Gradient Approach Revisited via Swarm Intelligence. In *International Conference on Modeling, Machine Learning and Astronomy (pp. 3-14). Springer, Singapore.*

Funding

$5,000, Google Cloud Academic Research Grant, Feb 2022

Service to Profession

Guest Editor, EMSE Special Issue on Replications and Negative Results (RENE) 2025

Co-Chair, Workshop on Replications and Negative Results (RENE) at ASE 2024

Reviewer, TMLR 2024; ICML 2024; Neural Processing Letters 2023, 2024; Neural Computing & Applications (NCAA), 2023, 2024; Artificial Intelligence Review 2023; ICLR 2024; NeurIPS 2023; Journal of Big Data, 2023; Automated Software Engineering (ASE), 2023; Empirical Software Engineering (EMSE), 2021; IEEE Symposium Series on Computational Intelligence (SSCI) 2020

PC Member, AAAI 2025; AI Foundation Models and Software Engineering (FORGE) at ICSE 2024; Automated Software Engineering (ASE) Artifact Evaluation Track, 2022; International Conference on Software Maintenance and Evolution (ICSME) Artifact Evaluation Track, 2021, 2022, 2023; International Conference on Modeling, Machine Learning, and Astronomy (MMLA), 2019

Honors and Awards

Google Cloud Champion Innovator - Cloud AI/ML, Jul 2023

Google Cloud Research Innovators Mentor, Dec 2022

Google Cloud Champion Innovator, Oct 2022

Google Cloud Research Innovator, Feb 2022

Relevant Projects

Programmable Resumes Aug 2023 - Present
Python *GitHub*

Developed a specification and implementation for modular, customizable resumes with support for two popular LaTeX templates. Wrote this resume using this tool.

pysh May 2021 - Dec 2024
C++, TMLanguage *GitHub*

Developed a superset of Python that allows running Shell code natively, with a VS Code syntax highlighting extension.

RAISE Aug 2020 - Nov 2024
Python, Keras *GitHub :: PyPI*

Sole developer for a PEP8-compliant, ML Python package used by our research lab and others for replicable results. Downloaded 43k times.

Threaded Discussions Website Feb 2021 - Jun 2021
MongoDB, Node.js, React *GitHub*

Companion website for video calls that allows for Reddit-style, threaded discussions.

Google/Meta Data Mining Feb 2021 - May 2021
Python, Keras *GitHub*

Data science project to use Google Takeout and Meta user data to suggest products to advertise to a user from Amazon best-sellers using DistilGPT-2, and achieved 0.6 F-1 score.

NearConnect Nov 2020 - Mar 2021
iOS, SwiftUI *GitHub :: App Store*

iOS app to connect with people nearby using multicast peer-to-peer connections.

Novel Drug Repurposing Hypotheses Oct 2019 - Feb 2020
Python, PyTorch *GitHub*

Identified novel drug repurposing hypotheses using text mining of radio transcripts, and verified results using a knowledge graph.

Personalized Chatbot May 2019 - May 2019
Python, Keras *GitHub*

Fine-tuned a GPT-2 345M model on 730k messages from Telegram logs to create a personalized chatbot.

Intelligent Tutoring System Sep 2018 - May 2019
Python *GitHub*

Implemented an Intelligent Tutoring System backend using Bayesian Knowledge Tracing and a novel question selection algorithm.

JournalBear Jan 2017 - Feb 2019
JavaScript, Electron *GitHub :: Softpedia*

Cross-platform journal application with AES-256 encryption. Rated 4/5 by Softpedia.

Astronomy Image Restoration Aug 2018 - Nov 2018
Python, Keras *GitHub*

Developed a machine learning approach to restore astronomical images affected by PSF anisotropy and smearing in crowded-field photometry, improving data retention and accuracy in differential imaging analysis of long-baseline optical time series.

Human Activity Data Project Oct 2018 - Nov 2018
Python, Keras *GitHub*

Collected personal activity data for 9 months, grouped tasks into 21 categories. Analyzed most productive hours of the day.

Results Scraper Mar 2018 - Aug 2018
MongoDB, Express.js, React, Node.js *GitHub*

Website for scraping university examination results and displaying charts and printable reports, with caching.

Video Sharing Website Oct 2017 - Dec 2017
MySQL, Express.js, React, Node.js, Sass, Elasticsearch *GitHub*

Simplified implementation of a video-sharing website with subscriptions and custom searching.

Web development projects Jan 2017 - Nov 2017
MongoDB, Express.js, React, Node.js, Sass, D3.js *CodePen*

Projects include URL shortener, rogue-like dungeon crawler game, voting application, Simon game, land surface temperature heatmap, and mapping meteorite impacts across the globe.

Xtreme Calculations Apr 2013 - Oct 2017
VB.NET, Python *Softpedia*

Windows math software to solve scientific and mathematical problems, with over 30k downloads across multiple sites.

Skills

Languages: Python, TypeScript, Java, C++, Gleam

Frameworks: Flask, Keras, PyTorch, Node.js, React

Databases: MySQL, MongoDB, DynamoDB

Cloud: Google Compute Engine, S3, Google Cloud Storage, EC2