# 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* — *Aug 2019 - Jul 2024*

- Advisor: *Dr. Tim Menzies*
- Dissertation: *Guidelines for the Application of Neural Technologies in Software Analytics (or: How to Do More with Less in SE)*

**PES University** — Bangalore, India
*B.E. Computer Science – Advisor: Dr. Snehanshu Saha* — *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% and reduced peak memory usage duration by 21.3%.
- **Drafting improvements:** Helped improve complaint drafting results by 115% and motion drafting by 28.4% of usefulness.
- **Workflow improvement:** Led initiative to use Bayesian optimization for prompt tuning.
- **Innovation:** Developed fast passage filtering approach based on contrastive loss and tabu search.
- **Technology:** Python, Litestar, React, TypeScript, Tailwind

**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 SOTA and is 36.4% 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) compared to prior SOTA.
- **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)

**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

## RECENT PUBLICATIONS

See full list on Google Scholar.

1. **Yedida, R.**, & Menzies, T. (2025). Is Hyper-Parameter Optimization Different for Software Analytics? *IEEE Transactions on Software Engineering, doi: 10.1109/TSE.2025.3550103.*

2. 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.*

3. **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, 49(4), 2856-2872.*

4. **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, 219, 119673.*

5. **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).*

## Funding

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

## Service to Profession

**Guest Editor**, IEEE Software SI: The Impact of AI on Productivity and Code 2025; ASEj SI: Replications and Negative Results (RENE) 2025; EMSE SI: Replications and Negative Results (RENE) 2025

**Co-Chair**, ASE 2025 Workshop on Intelligent Software Engineering; ASE 2024 Workshop on Replications and Negative Results (RENE)

**PC Member**, ICSE 2026; AI Foundation Models and Software Engineering (FORGE) 2024; ICSME Artifact Evaluation Track, 2021-2023; ASE Artifact Evaluation Track, 2022; International Conference on Modeling, Machine Learning, and Astronomy (MMLA), 2019

**Reviewer**, NeurIPS 2023, 2025; EMSE 2021, 2025; ASEj 2025; AAAI 2025; ICML 2024-2025; ICLR 2024-2025; NCAA 2023-2025; TMLR 2024; Neural Processing Letters 2023-2024; Artificial Intelligence Review 2023; Journal of Big Data, 2023; ASE 2023; IEEE SSCI 2020

## Honors

**Google Developer Expert - ML and GCP**                             Mar 2025 - Present

The Google Developer Experts program is a global network of highly experienced technology experts, influencers, and thought leaders who have expertise in Google technologies, are active leaders in the space, natural mentors, and contribute to the wider developer and startup ecosystem.

**Google Cloud Champion Innovator - Cloud AI/ML**                    Oct 2022 - Mar 2025

Champion Innovators are a global network of more than 500 professionals, nominated by Googlers, who are technical experts in Google Cloud products and services. Each Champion specializes in one of nine different technical categories.

## Relevant Projects

**RAG Agent for Zotero**                                             Oct 2024 - Present
*Rust, SQLite*                                                              [GitHub]

WIP: Building a PDF parser and a RAG agent for academic papers by linking to the user's Zotero library.

**RAISE**                                                           Aug 2020 - Present
*Python, Keras*                                                        [GitHub :: PyPI]

Sole developer for a PEP8-compliant, ML Python package used by our research lab and others for replicable results. Downloaded 55k times.

**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.

**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.

## Skills

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

**Frameworks**: Flask, Node.js, React

**ML/AI**: Keras, PyTorch, Hyper-parameter Optimization, Bayesian Optimization

**Databases**: MySQL, MongoDB, DynamoDB

**Cloud**: Google Cloud (Compute Engine, Cloud Storage, Vertex AI), AWS (EC2, S3, DynamoDB)