Rohith Pudari

PHD STUDENT, ELECTRICAL AND COMPUTER ENGINEERING

✉ r.pudari@mail.utoronto.ca | ⌂ www.eecg.utoronto.ca/ rpudari/ | ⚇ rohithpudari | ⚇ rohithpudari | ⚇ rohithpudari

Education

PhD. Electrical and Computer Engineering Toronto, Canada
University of Toronto 2022 - Ongoing

M.Sc. Computer Science Victoria, Canada
University of Victoria,
Thesis: AI Supported Software Development: Moving Beyond Code Completion. [PDF] 2020 - 2022

B.Tech. Information Technology Hyderabad, India
Jawaharlal Nehru Technological University, GPA 9.2/10 2015 - 2019

Research and Experience

University of Toronto. Graduate Researcher Toronto, Canada
Advisor: Dr. Shurui Zhou 2022-present

- Worked on organising Stack-Overflow data using Documentations.
- A structured organization of Stack-Overflow posts can help readers branch out of their initial questions and explore other content, which is especially helpful when a developer is unaware of the specific items they are looking for. The project will focus on classifying questions on python into classes that are defined by the official standard library documentation of python.

University of Victoria. Graduate Researcher Victoria, Canada
Advisor: Dr. Neil Ernst 2020-2022

- Explored the current limitations of large language models for code such as Copilot, and created a taxonomy to moving beyond code completion to AI-supported software engineering.
- Performed a qualitative study on finding the best ways to identify and model the topics of software design by mining public question answering platforms(stack-overflow and stack-exchange), which can be used to identify changing design approaches, validating ml-based code generation for design smells and recommending related design discussions.
- Created a server to conduct a survey for collecting insights on how software developers interact with various levels of autonomous bots on GitHub.
- Gained experience with the entire research process including developing new research ideas and hypotheses, experimental design, data analysis and interpretation, and writing scientific papers.

Deloitte, Data Scientist Hyderabad, India
Advisor: Dr. Anitha Jadhav 2019 - 2020

- Created architecture for a server cluster to support access and analysis of data to employees spread across the world.
- Worked on data of fortune 500 pharma company, made a custom model to predict risks and demand of products using various factors and add it to the reports.
- Combined models through ensemble modelling, Present information using data visualisation techniques, Proposed solutions and strategies to business challenges.

Deloitte, Research Scientist Intern Hyderabad, India
Advisor: Dr. Anitha Jadhav 2019

- Created tests and wrote documentation for a internal tool for detecting code smells and test coverage report.
- Implemented a ML application where risks associated for a product is calculated using custom Questionnaires thereby providing risk mitigation plans.
- Worked under Governance, Risk and Compliance - Data Collection and Analysis

Google, Software Engineering Intern
Bangalore, KA
Advisor: Praveen Singh
2018

- Created the visual flow builder for the DialogueFlow web interface.
- Worked on creating Dialogue-flow intent generation process and scaling up the payment gateway for the google payment application in India.
- Gave multiple talks on the usages and ways to integrate DialogueFlow into existing softwares at conferences and meetups.

Life Of Girl, Software Engineering Intern
Hyderabad, TS
Advisor: Dr. Sikinder Khan
2017

- Created a Chat bot, to connect the local police department (SHE team) to help women safety, This received a government funding and eventually got implemented in local police network.
- Deployed a ML model to create personalised responses for the chat bot, where it further used DialogueFlow as its backend.

Sreenidhi Institute of Science and Technology, Undergraduate Researcher
Hyderabad, India
Advisor: Dr. Subhani Shaik
2015-2019

- Performed a comparative study based on optimization techniques for software cost estimation, which improved the accuracy of software cost estimations by coupling Bayesian multi-class algorithm with existing optimization techniques. The developed model is empirically validated using different evaluation metrics through a statistical framework.
- Created multiple projects to perform facial and expression detection and made them scalable for real world use
- Designed a parallel computing algorithm to perform pattern mining in datasets.

Publications

1. R. Pudari, S. Zhou, I. Ahmed, Z. Dai and S. Zhou, *"Aligning Documentation and Q&A Forum through Constrained Decoding with Weak Supervision,"* 2023 IEEE International Conference on Software Maintenance and Evolution (ICSME), Bogotá, Colombia, 2023, pp. 346-351, doi: 10.1109/ICSME58846.2023.00043. [LINK].

2. Rohith Pudari, Neil Ernst, *"From Copilot to Pilot: Toward AI Supported Software Engineering".*
   arXiv preprint arXiv:2303.04142 (2023)

3. Roshan Lasrado, Rohith Pudari, Neil Ernst, *"What Do Developers Discuss About Design? Exploring Design Knowledge in Stack Overflow and GitHub".* arXiv preprint, (2021)

Teaching and Mentorship

ECE 444 - Software Engineering
University of Toronto
Teaching Assistant- [Course Link]
Fall 2023, Fall 2022

- This course explores these issues broadly covering the fundamentals of modern software engineering. The course combines rigorous foundations (guiding principles, precise terminology, well-defined techniques) with extensive opportunities for the development of practical skills using state-of-the art tools and techniques based on the latest research and practice in software engineering.

SENG 321 - Requirements Engineering
University of Victoria
Teaching Assistant
Spring 2022, Spring 2021

- This course is designed to address the issues of requirements management throughout the software development life cycle, teaching techniques for requirements elicitation, analysis and modelling, as well as formal specification, negotiation and decision making

SENG 275 - Software Testing
University of Victoria
Teaching Assistant
Summer 2021

- A Course on the practical application of testing in a modern software development environment, which includes a variety of industry standard languages, tools and frameworks while applying testing techniques to improve the quality of multi-component software product.

SENG 350 - Software Architecture and Design
University of Victoria

Teaching Assistant
Fall 2021

- An introductory course to analysis and design of software architectures with UML (Unified Modelling Language) and their subsequent synthesis at the program level.Topics include requirements analysis, analysis and design of static and dynamic view points of architectures and model driven engineering. Design patterns are introduced and applied as solutions to recurring design problems.

SENG 330 - Object-Oriented Software Development
University of Victoria

Teaching Assistant
Fall 2020

- A introduction to the discipline of software design, with a focus on object-oriented design. Software design can easily be discounted as a secondary concern when solving small computing problems such as assignments and scripts. However, it assumes a critical role when attempting to build realistic and high-quality software applications.

Machine Learning Crash Course (MLCC) workshop
Hyderabad, TS

Organizer
Fall 2018

- I served as a Organizer for a one day workshop, which included talks from industry experts showing how to learn and apply fundamental machine learning concepts

Explore ML workshop
Bangalore, KA

Organizer
Fall 2019

- I planned and implemented the delivery of learning and development experiences to students all over India on campus in a 'flipped classroom' environment.

Facebook Developer Circles, Hyderabad
Hyderabad, TS

Mentor
Fall 2020

- Serving as a mentor for beginners in computer science and helping them engage with the community and learn from the events hosted by the organisation.

Climate Change AI (CCAI) workshop
Long Beach, CA

Mentor
Fall 2019

- I served as a Research Mentor as part of the Climate Change AI (CCAI) ICML workshop program. I provided feedback, advice, and research discussions to a mentee interested in submitting to the workshop

Awards

2019	**Recipient**, Gold Medal Award at SNIST	Hyderabad, TS
2017	**Recipient**, Hyderabad Best Coder at JNTUH	Hyderabad, TS
2016	**Scholarship**, Outstanding student fellowship, SNIST	Hyderabad, TS

Languages and Tools

Python, Tensorflow, Pytorch, Ruby, Django, Javascript, Matlab, java, C, C++, C#, Scipy, Scikit-learn, Pandas, swift, AWS, Numpy, Gensim, OpenCV, Homebrew, Qiskit, Julia, Theano, Google App Engine, dialogueflow, Weka, Git, SVN.

Open Source Projects

PySyft, Openmined
https://github.com/OpenMined/PySyft

- PySyft decouples private data from model training, using Federated Learning, Differential Privacy, and Encrypted Computation (like Multi-Party Computation (MPC) and Homomorphic Encryption (HE)) within the main Deep Learning frameworks like PyTorch and TensorFlow
- The Syft ecosystem seeks to change this system, allowing you to write software which can compute over information you do not own on machines you do not have (total) control over. This not only includes servers in the cloud, but also personal desktops, laptops, mobile phones, websites, and edge devices. Wherever your data wants to live in your ownership, the Syft ecosystem exists to help keep it there while allowing it to be used privately for computation.

SwiftSyft, Openmined

https://github.com/OpenMined/SwiftSyft

- SwiftSyft makes it easy for you to train and inference PySyft models on iOS devices. This allows you to utilize training data located directly on the device itself, bypassing the need to send a user's data to a central server.
- OpenMined set out to build the world's first open-source ecosystem for federated learning on web and mobile. SwiftSyft is a part of this ecosystem, responsible for bringing secure federated learning to iOS devices. You may also train models on Android devices using KotlinSyft or in web browsers using syft.js.

Smith-Waterman algorithm optimization

https://github.com/HarisSmajlovic/smith-waterman-optimization

- performance optimizations for Linear gap Smith-Waterman algorithm, which is hard to parallelize due to its sequential nature of instructions.
- In comparison to our base non-parallel optimized algorithm, The final algorithm is on the magnitude of 74x faster. Discounting the overhead experienced from PCIe transfers, using a GPU significantly sped up the algorithm.
- Took a base implementation of the Smith-Waterman algorithm, and iteratively improved the data and task parallelism of the algorithm, improved memory access patterns, added SIMD and multicore, and GPU usage to increase algorithmic performance by more than 74x.

Summarizer

https://github.com/rohithpudari/summariser

- A personal use machine learning model developed using Pytorch and Python
- This project was using data from research papers published on arxiv from 2000 to 2017, using the abstract as the summary of the paper, iteratively including human summaries of articles related to computer science from various news source websites.
- Given a article or a research paper, the model could prioritise sentences and generate the summary of it. I developed it to save time in reading mandatory readings of classes and to quickly skim through many research papers to decide which one is worth reading for my current research direction.

Invited Talks and Panels

Major League Hacking   Hyderabad, TS

Speaker   2020

- Gave a talk on best practices in developing machine learning algorithms and ways to tackle bias.

Advances in SE4AI   Bangalore, KA

Panel Member   2019

- Discussion on how can we use current software engineering methods in the process of machine learning development process.

DevFest Hyderabad by Google   Hyderabad, TS

Speaker   2018

- Introduced the ways to use Dialogueflow in every application to support chatbots