PARAM BIYANI

prmbiy.github.io parambiyani8@gmail.com +91 8425935760 Google Scholar

EDUCATION

BITS Pilani - K.K. Birla Goa Campus *August 2019 - July 2023*
B.E. in Computer Science, Minor in Data Science CGPA: **8.96/10**

RESEARCH INTERESTS

AI for Code, Artificial Intelligence, Software Engineering, Human-AI Collaboration, AI for Math, Formal Reasoning

WORK EXPERIENCE

Qualcomm Hyderabad, India
Senior AI Engineer *December 2025 - present*

- Developing an automated NL to spec generator (autoformalization) for Qualcomm's chip software verification pipelines.
- Led development of a coding assistant for low-resource domain specific languages, used by 15000+ C developers daily.

Supervised Program for Alignment Research Remote
Research Fellow *February 2025 - present*

- Developing a benchmark to evaluate agent shutdownability and assess the extent to which agents exhibit evaluation awareness.
- Investigating when agents fail to comply with shutdown and developing mitigation strategies to improve reliability.

Microsoft, PROSE Team Bangalore, India
Research Fellow *October 2023 - October 2025*

- Led research on an autonomous **agentic AI** system that does whatever a human developer can: task planning, code exploration and editing; contributions shipped to **GitHub Copilot**, with a 15% gain on swe-bench-verified over the baseline swe-agent.
- Improved agentic performance across long term interaction benchmarks (SpreadSheetBench, τ-bench, OSWorld (autonomous computer-use)) tasks by 10–20% using a novel bootstrapping algorithm that learns knowledge bases from past experiences.
- Developed a system to autoformalize informal olympiad math problems to Lean 4 theorems, achieving a 95% automation.
- Designed a rubric-based learning algorithm to infer human satisfaction and task relevance from conversational data, deployed across **Visual Studio Copilot** and **GitHub** assistants evaluating millions of interactions per week.
- Contributed to an interactive debugger, achieving 3x improvement in bug-resolution rates. Deployed in Visual Studio IDE.

Adobe, Adobe Exchange Bangalore, India
Member of Technical Staff *July 2023 - October 2023*
Product Intern *May 2022 - August 2022*

- Integrated code vulnerability checks for Adobe Photoshop and InDesign plugin package upload service.
- Implemented asynchronous file uploads, resulting in 60% accelerated upload speeds on Adobe Exchange.
- Worked on observability on Entitlements Service in Adobe IO using Java SpringBoot, Grafana, Prometheus.

American Express, AI Labs (Data Science) Bangalore, India
Analyst Intern *January 2023 - June 2023*

- Built a toolkit for tracking feature importance across decision-tree models, enabling production model behaviour monitoring.
- Developed a custom Gini gain function that improved (1) regularization on data with high missingness and (2) generalization on multi-market models. Improved an internal version of XGBoost, aimed at credit default/fraud prediction.
- Ideated a novel imputation strategy that dynamically selects methods based on feature missingness type, outperforming standard imputation on high-missingness financial data and improving precision by 2%.

Speech and Language Lab, NTU | NTU, Singapore
Research Intern *August 2022 - January 2023*

- Trained convolution-augmented vision transformers to improve the acoustic feature representations of audio signals, improving F1 score our audio analysis pipeline by 10%. Developed a real-time Voice Activity Detection (VAD) system.
- Developed vision transformer models with CRNNs and self-attention to secure top five in DCASE 2023 Challenge Task4b.
- Used by Singapore Technologies for emergency audio event detection in real-world deployments.

PAPERS AND PUBLICATIONS

C = CONFERENCE, W = WORKSHOP

C1. 🏆 Best Paper Presentation **Param Biyani***, Yasharth Bajpai*, Arjun Radhakrishna, Gustavo Soares, Sumit Gulwani – **RUBICON: Rubric-Based Evaluation of Domain-Specific Human AI Conversations** *AIware 2024*

C2. 🏆 Best Paper Y. Bajpai*, B. Chopra*, **P. Biyani**, C. Aslan, S. Gulwani, D. Coleman, C. Parnin, A. Radhakrishna, G. Soares – **Let's Fix this Together: Conversational Debugging with GitHub Copilot** *VL/HCC 2024*

W1. **Param Biyani**, Shashank Kirtania, Yasharth Bajpai, Sumit Gulwani, Ashish Tiwari – **INDIMATHBENCH: Autoformalizing Mathematical Reasoning Problems with a Human Touch** *P-AI-FM at AAAI 2026; Under Submission ICLR 2026*

W2. Shashank Kirtania, **Param Biyani**, Priyanshu Gupta, Yasharth Bajpai, Roshni Iyer, Sumit Gulwani, Gustavo Soares – **Improving Language Agents Through BREW** *MTI-LLM at NeurIPS 2025; Under Submission ICLR 2026*

W3. B. Chopra*, Y. Bajpai*, **P. Biyani**, G. Soares, A. Radhakrishna, C. Parnin, and S. Gulwani – **Exploring Interaction Patterns for Debugging: Enhancing Conversational Capabilities of AI-assistants** *HCI and NLP at NAACL 2024*

SELECTED RESEARCH PROJECTS

IndiMathBench: Formal theorem proving benchmark and autoformalization study | July 2025 - Sept 2025

- Created a 312 sized benchmark for Automated Theorem Proving benchmark in Lean4 from Indian Mathematical Olympiads.
- Conducted a systematic study of Human-AI collaboration on formalization tasks.
- Created a VS code extension to speed up Lean annotation by 3.5x, using multi-LLM formalization and group reflection.

AutoDev: Autonomous Programming Agent June 2024 - March 2025

- Designed and led a 6+ month research project on automating software development using multi-agent AI systems.
- Built specialized agents for planning, retrieval, code generation, and review, enabling hierarchical task execution.
- Running and quickly iterating on experiments to develop the agent system support and break SOTA on 20+ benchmarks including swe-bench. Directly informed key system-design decisions for GitHub Copilot.

RUBICON: Human-AI conversation evaluation | | October 2023 - April 2024

Under Dr. Gustavo Soares and Dr. Arjun Radhakrishna *Microsoft PROSE*

- Devised a learning technique that generates user satisfaction and domain relevant assertions from conversational data.
- Designed a custom scoring function and greedy based algorithm to select top n assertions from a larger pool of generations.
- Improved classification rate with confidence of > 0.9 by 56%, and F1 score by 16%.

TEACHING ASSISTANTSHIPS

Machine Learning, Deep Learning, Foundations of Data Science, Object Oriented Programming, Computer Programming

EXTRACURRICULARS

CS Department Student Body, BITS Goa | *Academic Coordinator*

- Founded the ASCII Mentorship Programme during undergrad. A student led set of projects aimed at research output. Involves 24 projects offered by 37 seniors and 10 professors. More than 300 students are involved in the projects.

TECHNICAL SKILLS

Languages:	Python, C, C++, Lean 4, Java, C#, JavaScript
AI:	Prompt Engineering, RL, Fine-Tuning, Developing Agentic Systems, RAG, Transformers
Softwares & Tools:	Pytorch, Tensorflow, All popular LLM usage/APIs, Blender 3D, Pandas, Matplotlib, sklearn, Numpy

RELEVANT COURSES

Computer Science:	Data Structures and Algorithms, Operating Systems, Object Oriented Programming, Database Management Systems, Computer Networks, Design and Analysis of Algorithms
ML/AI:	Deep Learning, Machine Learning, Reinforcement Learning, Linguistics, Applied Statistical Methods, Cognitive Neuroscience, Computational Learning Theory (Audit)

ACHIEVEMENTS

- Best Paper Award, VL/HCC 2024; Awarded by IEEE Symposium on Visual Languages and Human-Centered Computing.
- Best Paper Presentation Award, AIware 2024; Awarded by ACM International Conference on AI-Powered Software.
- Merit Scholarship for being in the **top 3%** of students based on Academic Performance.