Priyanshu Gupta

Associate Researcher II
PROSE Group, Microsoft Redmond

priyansgupta@microsoft.com | priyanshu.42g@gmail.com
Google Scholar : Priyanshu Gupta
Github : www.github.com/purug2000
+91 9559006855 | guptapriyanshu.com

INTERESTS

Primary: Large Language Models, Natural Language Processing, Agent Based Systems
Others: AI for Software Engineering, Developer Productivity

EDUCATION

Indian Institute of Technology Kanpur   *India*
Bachelor of Technology, Computer Science and Engineering   *July '17 - May '21*
CPI: 9.6/10

Aalto University School of Science   *Finland*
Exchange Studies, Computer Science   *Jan '20 - June '20*

EXPERIENCES

Microsoft
Researcher IC3   *2021- present*, **Bangalore, India**

Github Copilot Chat for Visual Studio and Visual Studio Code
- Contributed extensively on Copilot Chat by implementing the testing infra., owning chat modes and bug fixes.
- Built a simple bespoke framework to enable agent based contextual multi-turn conversations within copilot.
- Developed an LLM-based multi-agent autonomous system for assisting users upgrade their Java codebase within **Github Copilot Chat**. Copilot for Java Upgrade was featured at **Github** Universe 24.   [Blog]
- Built and owned an LLM based chat experience in **Github Copilot** to assist users with **cloud migration** tasks. The experience helps 1000 users daily to migrate their projects to azure. Featured in Build 23.   [Blog]

Intelligent Next Code Edit Prediction
- Evaluated LLM's capabilities to predict future edits given prior edits by experimenting with different prompting styles and fine-tuning a small Language Model for edit prediction.   *GrACE; Gupta et.al. FSE' 23*
- Designed a novel technique to automatically learn patterns from code editing data and using them for next edit prediction. Shipped feature in **Intellicode** with 2000 daily hits.   *Overwatch; Zhang et.al. OOPSLA' 22*

Conversational Advanced Data Analytics in Microsoft Excel
- Developed a novel synthetic data generation pipeline for creating data analysis queries from raw data science articles.
- Evaluated various Open source LLMs and agentic design patterns for the conversational data analysis tasks.

Misc. Research on Language Agents
- Proposed a technique that equips Language Agents with a memory of past failures, resulting in a 5-25% performance improvement in production systems.   *MetaReflection; Gupta et.al. EMNLP'24*
- Developed a technique to fix errors in Retrievel Augmented Generation(RAG) databases using RAG system failures that improves performance by upto 9% across internal and external scenarios.   *StackFeed; Gupta et.al. arxiv.*

Adobe Research India
Research Internship   *Summer'20*, **Adobe Research Labs, Bengaluru**

Improving Experience Interface - Fact Extraction and Verification
- Proposed a multi-stage retrieval and reasoning technique to verify factoid claims over a free-text Wikipedia corpus.
- Beat the existing State of the Art performance on the ACL FEVer dataset by 1 percent in terms of the FEVer score with a **10 percent** improvement on claims requiring multiple evidences.

(Patent Granted)

PATENTS

Granted

- ML Techniques for augmenting E-Documents with data verification Indicators   [US11687716B2]
Navita Goyal, Vipul Shankhpal, **Priyanshu Gupta**,Ishika Singh, Baldip Singh Bijlani, Anandha velu Natarajan

Published

- A System and Method for Detection of Keystroke Using Defocus   [App Number 201911044242]
Priyanshu Gupta,Tushar Goswamy, Himanshu Kumar, K S Venkatesh

Publications

- METAREFLECTION: Learning Instructions for Language Agents using Past Reflections [Link]
Priyanshu Gupta, Shashank Kirtania, Ananya Singha, Arjun Radhakrishna, Sherry Shi, Gustavo Soares
Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)

- Grace: Language Models Meet Code Edits [Link]
Priyanshu Gupta*, Avishree Khare*, Yasharth Bajpai, Saikat Chakraborty, Sumit Gulwani, Aditya Kanade,
Arjun Radhakrishna, Gustavo Soares, Ashish Tiwari
ACM Symposium on the Foundations of Software Engineering (FSE), 2023

- Overwatch: Learning Patterns in Code Edit Sequences [Link]
Priyanshu Gupta*, Yuhao Zhang*, Yasharth Bajpai*, Ameya Ketkar, Miltiadis Allamanis, Titus Barik, Sumit
Gulwani, Arjun Radhakrishna, Mohammad Raza, Gustavo Soares, Ashish Tiwari
Proceedings of the ACM on Programming Languages OOPSLA, 2022

- TST^R: Target Similarity Tuning Meets the Real World [Link]
Anirudh Khatry, Sumit Gulwani, **Priyanshu Gupta**, Vu Le, Ananya Singha, Mukul Singh, Gust Verbruggen
Findings of the Association for Computational Linguistics: EMNLP 2023

- LOGIC-LM++: Multi-Step Refinement for Symbolic Formulations [Link]
Shashank Kirtania, **Priyanshu Gupta**, Arjun Radhakrishna
Proceedings of the 2nd Workshop on Natural Language Reasoning and Structured Explanations @ ACL 2024

- A Defocus Based Novel Keyboard Design [Link]
Priyanshu Gupta, Tushar Goswamy, Himanshu Kumar, K.S. Venkatesh
International Conference on Human Computer Interaction (HCII), 2020

- Convex Surrogates for Unbiased Loss Functions in Extreme Classification With Missing Labels
Erik Schultheis, Mohammadreza Qaraei, **Priyanshu Gupta**, Rohit Babbar [Link]
Presented at *Workshop on eXtreme Classification, ICML 2020*
Accepted at *The Web Conference*, 2021

- A Probabilistic Framework for Knowledge Graph Data Augmentation [Link]
Priyanshu Gupta*, Jatin Chauhan*, Pasquale Minervini
Presented at Data-centric AI Workshop @ *NeurIPS 2021*

- Source-Free Unsupervised Domain Adaptation using Class Prototypes [Link]
Priyanshu Gupta*, Jinang Shah*, Nidhi Hegde*, Harshit Kumar*,Vaibhav Jindal, Ashutosh Modi
Presented at SemEval workshop @ *ACL-IJCNLP 2021*

Pre-prints

- STACKFEED: Structured Textual Actor-Critic Knowledge Base Editing with FeedBack [Link]
Naman Gupta, Shashank Kirtania, **Priyanshu Guta**, Krishna Kariya, Sumit Gulwani, Arun Iyer, Suresh
Parthasarathy, Arjun Radhakrishna, Sriram K. Rajamani, Gustavo Soares

- Augmented Embeddings for Custom Retrievals [Link]
Anirudh Khatry, Yasharth Bajpai, **Priyanshu Gupta**, Sumit Gulwani, Ashish Tiwari

Skills

Languages:	Python, C, .NET *with familiarity in* C++, Java and TypeScript
Frameworks/libraries:	PyTorch, Scikit-learn, Transformers

Honors

2024	Served as a reviewer for ACL, EMNLP and ICLR conferences.
2023	Served as a reviewer for ACL, EMNLP and ICSE conferences.
2019	**2nd** Prize at the KGMU-IITK Medical Hackathon.
2018/19	Awarded **Academic Excellence Award** twice for academic proficiency by IIT Kanpur.
2017	Secured All India Rank(AIR) **599** among 1.5M applicants in **JEE-Advanced, 2017**.
2016	Secured AIR **21** among 100K applicant in **KVPY -SX** examination.