

Puneet Mathur

✉ pmathur3k6@gmail.com | 🌐 themadaiguy.github.io | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#) | 📍 San Jose, USA

Education

University of Maryland, College Park	College Park, MD, USA
Ph.D. and M.S. in Computer Science; GPA: 3.97/4.0	Aug 2019 – Dec 2023
Advised by Dr. Dinesh Manocha	
Netaji Subhas Institute of Technology (NSIT)	New Delhi, India
B.E. in Computer Engineering; GPA: 3.8/4.0	Aug 2014 – May 2018

Professional Experience

Adobe Research	San Jose, CA, USA
Research Scientist, Agentic Intelligence Research (AIR Lab)	Sept 2023 – Present
- Full-duplex audio language modeling: Building robust evaluation benchmarks for full-duplex speech agents (document grounding, hallucination, and instruction-following), and post-training full-duplex audio LMs to think, reason, and retrieve during conversation for grounded question answering; mitigating latency via learned audio fillers and GRPO post-training for reasoning and retrieval, alongside efficient inference optimizations for omni-modal models. - Partial-policy gradients for RL in LLMs: Novel credit assignment technique for offline policy gradients that scales sub-horizon planning for multi-step LLM trajectories, with scaling laws for policy gradient finetuning and improved long-context persona-following in LLM roleplaying. - Reasoning-based LLM clustering via RL/GRPO: Post-trained Qwen models via GRPO to reason over complex user queries for organizing large text collections, improving query-aware clustering by +14% V-score (at par with GPT-o3). - Roleplaying agents: Multi-agent orchestration layer with planning, behavior alignment, and role verification for consistent specialist personas; RL reward modeling and curriculum learning for tutor agents. - Self-evolving agents: Led development of a self-evolving tutor agent using TextAdam, a conversational prompt optimization technique yielding a 23% gain in instruction following over TextGrad and GEPA. - Agentic memory and continual learning: Novel memory architectures that let LLM agents learn from past interactions and reduce recurring errors; deployed in the Common Agentic Framework and used by 5+ Adobe products. - LLM attribution and RAG: Post-hoc attribution agents providing citations for QA systems over document text, tables, charts, and visual artifacts via multimodal retrieval and agentic verification; improved table attribution in Acrobat AI Assistant by 12%.	

Research Internships

Microsoft Research	Redmond, WA, USA
Research Scientist Intern, Visual Document Understanding	May 2023 – Aug 2023
Developed a Transformer-based model for chart information extraction that recovers chart structure and infills underlying data values, improving F1 score by 15% over prior methods.	
Meta AI (Facebook AI)	Menlo Park, CA, USA
Research Scientist Intern, NLP & Speech Team	Sept 2022 – May 2023
AI for speech recognition and language model personalization: retrieval-augmented language modeling for NLP+speech, generative NLP for factual consistency and domain adaptation, and embedding-based retrieval for query understanding, improving word-error rate in ASR across several datasets. Papers accepted to EMNLP 2023 and Interspeech 2023.	
Adobe Research	San Jose, CA, USA
Research Scientist Intern, Document Intelligence Lab	June 2022 – Aug 2022
Automated document editing based on natural language prompt instructions. Paper accepted to AAAI 2023.	
Verisk AI	New Jersey, USA
AI/ML Research Intern	Jan 2022 – June 2022
Natural language inference on legal documents using optimal evidence selection. Paper accepted to EMNLP 2022.	
Adobe Research	San Jose, CA, USA (Remote)
Research Scientist Intern, Document Intelligence Lab	June 2021 – Aug 2021

- End-to-end extraction of hierarchical document structures from semi-structured forms and invoices. Paper accepted at WACV 2023.

Dataminr Inc.

New York, USA

AI Research Intern

June 2020 – Aug 2020

- Duplicate detection of news headlines as a tiered (event-entity-phrase level) paraphrase identification task, deployed as an internal system.

Selected Publications on Full-Duplex Audio & Speech - [All Publications](#)

1. Duplex-R1: Full-Duplex Audio LMs that Reason, Retrieve, and Speak While Searching
Puneet Mathur, Nedim Lipka, Zeyu Jin, Dinesh Manocha
In submission, 2026 ([Coming soon](#))
2. DuplexSpeechBench–Document Grounding: Benchmarking Document Grounding and Hallucinations in Full-Duplex Voice Agents
Puneet Mathur, Nedim Lipka, Zeyu Jin, Dinesh Manocha
In submission, 2026 ([Coming soon](#))
3. DuplexSpeechBench–IFEval: Instruction Following Evaluation for Full-Duplex Voice Agents
Puneet Mathur, Manan Suri, Dinesh Manocha
In submission, 2026 ([Coming soon](#))
4. Omni2LoRA: Coherence-Preserving Parametric Memory for Efficient Omni Language Models
Puneet Mathur, Manan Suri, Dinesh Manocha
In submission, 2026 ([Coming soon](#))

Selected Publications on Agentic AI & Reinforcement Learning - [All Publications](#)

1. [ARGUS: Structured Uncertainty guided Clarification for LLM Agents](#)
Manan Suri, Puneet Mathur, Nedim Lipka, Franck Deroncourt, Ryan A. Rossi, Dinesh Manocha
ACL 2026
2. [Partial Policy Gradients for RL in LLMs](#)
Puneet Mathur, Branislav Kveton, Subhojyoti Mukherjee, Viet Dac Lai
In submission, 2026
3. [Cluster-R1: Large Reasoning Models Are Instruction-following Clustering Agents](#)
Peijun Qing, Puneet Mathur, Nedim Lipka, Varun Manjunatha, Ryan A. Rossi, Franck Deroncourt, Saeed Hassanpour, Soroush Vosoughi
In submission, 2026

Selected Publications on Document Intelligence & RAG - [All Publications](#)

1. [ChartEval: LLM-Driven Chart Generation Evaluation Using Scene Graph Parsing](#)
Kanika Goswami, Puneet Mathur, Franck Deroncourt, Ryan A. Rossi, Vivek Gupta, Dinesh Manocha
AAACL 2025
2. [ChartLens: Fine-grained Visual Attribution in Charts](#)
Manan Suri, Puneet Mathur, Nedim Lipka, Franck Deroncourt, Ryan A. Rossi, Dinesh Manocha
ACL 2025
3. [Follow the Flow: Fine-grained Flowchart Attribution with Neurosymbolic Agents](#)
Manan Suri, Puneet Mathur, Nedim Lipka, Franck Deroncourt, Ryan A. Rossi, Vivek Gupta, Dinesh Manocha
EMNLP 2025
4. [VisDoM: Multi-Document QA with Visually Rich Elements Using Multimodal Retrieval-Augmented Generation](#)
Manan Suri, Puneet Mathur, Franck Deroncourt, Kanika Goswami, Ryan A. Rossi, Dinesh Manocha
NAACL 2025
5. [MoDS: Moderating a Mixture of Document Speakers to Summarize Debatable Queries in Document Collections](#)
Nishant Balepur, Alexa Siu, Nedim Lipka, Franck Deroncourt, Tong Sun, Jordan Lee Boyd-Graber, Puneet Mathur
NAACL 2025

6. [MATSA: Multi-Agent Table Structure Attribution](#)
Puneet Mathur, Alexa Siu, Nedim Lipka, Tong Sun
EMNLP 2024
7. [DocEdit-v2: Document Structure Editing Via Multimodal LLM Grounding](#)
Manan Suri, Puneet Mathur, Franck Dernoncourt, Rajiv Jain, Vlad I Morariu, Ramit Sawhney, Preslav Nakov, Dinesh Manocha
EMNLP 2024
8. [DocPilot: Copilot for Automating PDF Edit Workflows in Documents](#)
Puneet Mathur, Alexa Siu, Varun Manjunatha, Tong Sun
ACL 2024
9. [DocEdit: Language-guided Document Editing](#)
Puneet Mathur, Rajiv Jain, Jiuxiang Gu, Franck Dernoncourt, Dinesh Manocha, Vlad Morariu
AAAI 2023
10. [TIMERS: Document-level Temporal Relation Extraction](#)
Puneet Mathur, Rajiv Jain, Franck Dernoncourt, Vlad Morariu, Quan Hung Tran, Dinesh Manocha
ACL 2021
11. Several other works on document temporal-relation extraction, layout understanding, and text-to-SQL robustness, published at EMNLP/NAACL/WACV/ACL: [see full list here](#)
Puneet Mathur et al.

Selected Publications on Multimodal AI - [All Publications](#)

1. [PersonaLM: Language Model Personalization via Domain-distributed Span Aggregated K-Nearest N-gram Retrieval Augmentation](#)
Puneet Mathur, Zhe Liu, Ke Li, Yingyi Ma, Gil Keren, Zeeshan Ahmed, Dinesh Manocha, Xuedong Zhang
EMNLP 2023 (Findings)
2. [DOC-RAG: ASR Language Model Personalization with Domain-Distributed Co-occurrence Retrieval Augmentation](#)
Puneet Mathur, Zhe Liu, Ke Li, Yingyi Ma, Gil Keren, Zeeshan Ahmed, Dinesh Manocha, Xuedong Zhang
COLING 2024
3. [DocLayoutTTS: Dataset and Baselines for Layout-informed Document-level Neural Speech Synthesis](#)
Puneet Mathur, Franck Dernoncourt, Quan Hung Tran, Jiuxiang Gu, Ani Nenkova, Vlad Morariu, Rajiv Jain, Dinesh Manocha
Interspeech 2022
4. [MONOPOLY: Financial Prediction from MONetary POLicY Conference Videos Using Multimodal Cues](#)
Puneet Mathur, Atula Tejaswi Neerkaje, Malika Chhibber, Ramit Sawhney, Fu-Ming Guo, Franck Dernoncourt, Sanghamitra Dutta, Dinesh Manocha
ACM Multimedia 2022
5. [Affect2MM: Affective Analysis of Multimedia Content Using Emotion Causality](#)
Trisha Mittal, Puneet Mathur, Aniket Bera, Dinesh Manocha
CVPR 2021
6. Several other works on speech emotion recognition, financial risk forecasting, and early vision/social-media NLP, published at EMNLP/Interspeech/ICASSP/AAAI: [see full list here](#)
Puneet Mathur et al.

Skills

Programming Languages: Python, C++, Java, SQL

Data Science Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, Jupyter Notebook, NumPy, Pandas

Software Technologies: Flask, Kafka, AWS, Git