

Max David Gupta
mg7411@princeton.edu · (484) 522-8282

EDUCATION

Princeton University

M.S.E., Computer Science

Princeton, NJ

August 2024–Present

Graduate researcher in the machine learning department under the supervision of Prof. Tom Griffiths (AI Lab and Computational Cognitive Science lab).

Relevant Coursework: Machine Learning and Pattern Recognition, Foundations of Probabilistic Modeling, AI Safety, Cognitive Psychology, Great Moments in Computing

Research Interests: Relational reasoning, metalearning, and inductive biases in neural networks

Columbia University

B.A., Applied Mathematics

New York, NY

September 2017–May 2021

GPA: 3.51 — Major GPA: 3.71

Relevant Coursework: Natural Language Processing, Applied Deep Learning, Data Structures, Discrete Math, Linear Algebra, Intro to Statistics, Discrete Math, Complex Variables, Analysis and Optimization, Calculus III & IV

PRESENTATIONS & PAPERS

- **Gupta M.**, Rane S., McCoy T., Griffiths T. *Convolutional Neural Networks Can (Meta)-Learn the Same-Different Relation*, To Appear in the Proceedings of the 47th Annual Conference of the Cognitive Science Society
- Bencomo G., **Gupta M.**, Marinescu I., McCoy T., Griffiths T. *Teasing Apart Architecture and Initial Weights as Sources of Inductive Bias in Neural Networks*, To Appear in the Proceedings of the 47th Annual Conference of the Cognitive Science Society
- **Gupta, M.**, Franke M., Hawkins R. *Pragmatic Vigilance Inoculates Social Networks Against Misinformation*, Pre-print, under review
- Azizi E., Azad T., **Gupta M.**, Nazaret A. *Ensembling in Variational Autoencoder Architectures for Effective Posterior Distribution of Cell State Estimation*. Presentation: Computational Cancer Lab, Columbia University: May 2021
- **Gupta M.**, Nika J., Carsault T. *Multi-Step Chord Prediction for Human-AI Improvisation*. Presentation: Columbia Undergraduate Research Symposium: September 2020
- **Gupta M.**, Malt M. *Musical Markov Chains*. Presentation: Reid Hall Research Seminar, Paris, France: May 2020

AWARDS & HONORS

- Hackathon Winner: Cognitive Modeling of Humans vs. Large Language Models, University of Tübingen – 2023
- Dean’s List – Columbia University 2020–2021
- Heinrich Research Fellowship, Columbia University (\$2500) – 2020
- Spritz Family Research Grant, Columbia University (\$3000) – 2020
- PSAT National Commended Scholar – 2017
- Canadian Gold Medalist in both French (2016–2017) and Latin (2014–2017) National Language Exams
- Greville Smith Scholarship, McGill University (\$48,000, not availed) – 2017
- Hugh M. Brock National Entrance Scholarship, University of British Columbia (\$30,400, not availed) – 2017

RESEARCH EXPERIENCE

Max Planck Institute – Center For Humans and Machines

AI Behaviorist - Research Assistant

Berlin, Germany

May–August 2024

- Engineered online JS experiments to simulate multi-generational human–Large Language Model (LLM) communications in transmission networks of varying sizes and types. Built full-stack applications for research published in human-AI interaction conferences.
- Contributed visualizations and NLP analyses of public sentiments on AI risks across 400 human participants from India and the US. Compared attitudes via sociological coding.

University of Tübingen - Computational Pragmatics Lab

Research Intern – Supervisors: Michael Franke, Robert Hawkins, Charley Wu

Tübingen, Germany

Oct 2023–Present

- Built a multi-agent extension of the Rational Speech Act (RSA) framework to study the effects of Bayesian belief updating of persuasive social goals on belief diffusion in social networks. Simulated multi-agent RSA communications in random networks using Python and R. Derived conditions for belief convergence under varying network topologies and agent actions.
- First-authored a paper (under review) illustrating conditions under which listener vigilance inoculates networks against misinformation.

Columbia University - Azizi Lab

Research Assistant - Supervisor: Elham Azizi

New York, NY

Jan–August 2021

- Analyzed the effects of aggregating variational auto-encoder (VAE) outputs to form ensembled predictions for differential gene expression. Compared deep ensembling and batch ensembling VAE’s trained on single-cell human brain and heart data.
- Trained Bayesian neural networks and VAE’s in PyTorch; analyzed posteriors across random initializations. Presented methods and findings to the computational cancer biology lab.

IRCAM, Centre Pompidou

Research Fellow – Supervisors: Mikhail Malt, Jérôme Nika

Paris, France

Jan–Sep 2020

- Built language models (RNNs, LSTMs) trained in PyTorch on musical data from live jazz for human-AI improvisations.
- Authored and presented a report and poster on generative music with language models at Columbia’s 2020 research symposium. Authored a manuscript on stochastic models in music, published at Reid Hall’s research symposium.

PROFESSIONAL EXPERIENCE

Weill Cornell Medical, Cornell University

Research Software Engineer

New York, NY

May 2022–Nov 2022

- Worked on efficiency and indexing of the main NLP pipeline, using OCR to parse doctor notes into machine-readable text.
- Built an AWS-hosted ETL pipeline with Docker, Python, Java, and SQL to securely geocode address data from hospital patients, increasing geocoding accuracy and runtime efficiency by 15%.

Infosys Consulting

Business Analyst – AI & Automation

New York, NY

August 2021–May 2022

- Built and deployed an NLG model from open-source to automate 85% of credit loan risk report writing at a top 3 US bank.
- Engineered several NLP models for financial document classification, summarization, and generation in NLTK and JS.

- Assisted executive advising at 2 of the top 5 US banks in AI automation, chatbot implementation, and process mining. Wrote concise technical guides on each topic, shared with clients and the firm at large.

TEACHING EXPERIENCE

Princeton University

Princeton, NJ

Assistant Instructor: COS 126, COS 240

Sep. 2024–Present

Weekly teaching and grading for a cohort of 20 undergraduates in coding and theory problem sets and exams. (COS 126 is an intro to Java class and COS 240 is a proof-based math class for CS majors.)

University of Tübingen

Tübingen, Germany

Teaching Assistant: Reinforcement Learning for Language Model Training

Nov. 2023–Present

Grading and coding support for over 50 graduate students using TensorFlow for problem sets and RL research projects.

CompTIA

New York, NY

Head Data Science Instructor

Jan 2022–Present

Head instructor for an online data science and coding bootcamp (Python and SQL) and assistant teacher for web development with React JS. Designed and delivered written and technical curriculum on computer science, statistics, and data analysis.

Columbia University

New York, NY / Remote

Teaching Assistant: Calculus IV

Jan–May 2021

Graded assignments and held office hours for approximately 80 students in Professor Daniela De Silva’s class. Coordinated grading across sections and wrote technical guides on calculus topics.

Community and Outreach

Journal Clubs: Meta-Learning and Mechanistic Interpretability

Organized journal clubs for post-doc and graduate students to meet bi-weekly to discuss topics in meta-learning and mechanistic interpretability (supported by the Natural and Artificial Minds initiative at Princeton).

National Deep Inference Fabric (NDIF) Pilot

Virtual

Feb 2025 – Running mechanistic interpretability analyses on Llama 405-B with NN-sight, an NDIF-run library.

IICSSS 2023 (Interdisciplinary Computational Cognitive Science Summer School) Tübingen, Germany

1st Place: Hackathon for Cognitive Modeling

September 2023

Coursework: Comparing language models to humans; computational modeling for learning; human language models

ESSLI 2023 (European Summer School in Language, Logic & Information) Ljubljana, Slovenia

Coursework: Probabilistic Language of Thought; Formal Language Theory and Neural Networks; deep language learning from raw speech; logic, data, examples, and learning

Center for AI Safety

Berkeley, CA (Remote)

Intro to Machine Learning Safety Fellow

June–August 2023

Grant-funded student. Coursework covering mechanistic interpretability, machine ethics, systemic AI safety, adversarial robustness, and existential risk from future AI systems.

SKILLS, LANGUAGES & INTERESTS

Programming Languages: Python, Java, R, SQL, JavaScript, HTML, Bash, MATLAB

Technical Skills: Machine learning, data visualization, data analysis and statistical insights, computational modeling, scientific communication, experimental design, web design, teaching and curriculum design

Spoken Languages: English (fluent); French (intermediate)

Interests: Cognitive science, running, camping, reading