Puneet S. Bagga

(913)-548-6785 | psbagga17@gmail.com | psbagga.com | LinkedIn

Education

Georgia Institute of Technology — Atlanta, GA

B.S. in Computer Science, | GPA: 4.00/4.00 — *Aug 2022 – Dec 2024*

Selected Courses: Stochastics 1 (PhD), Fin. Optimization (grad), Machine Learning, Automata Theory, Algorithms

Research Experience

Industrial Engineering and Operations Research, Columbia University — New York, NY

Staff Associate I (Visiting Researcher/Student) – Advised by Dr. Tianyi Peng — *Sep 2025 – Present*

- Optimization methods for products in LLM recommendation systems, working on prompt meta-optimization (1)
- AI Agents Reading group, Social Simulation LLM project, observing *Generative AI* (PhD), *Machine Learning and High-Dimensional Data in OR* (PhD), with planned additional graduate coursework in Math and OR

Industrial and Systems Engineering, Georgia Institute of Technology — Atlanta, GA

Research Assistant (paid) – Advised by Dr. Arthur Delarue — *Jan 2025 – Apr 2025*

- Continuation of prior work

Undergraduate Researcher – Advised by Dr. Arthur Delarue — *Aug 2022 – Dec 2024*

- Core focus: Solving combinatorial optimization problems with machine learning. Applied ML & RL approaches, optimization techniques, & neural architectures to solve instances with advantageous tradeoffs and augment solvers
- Designed neural methods using Deep Sets and Set Transformer for multi-dimensional knapsacks. Used as a pre-solve, 90+% recovery and 2 OOM speedups vs Gurobi for large instances, capacity for generalizability (2)
- Trained dual pointer network with sequential formulation of Quadratic Assignment Problem with A2C & PPO, benchmarked against classical heuristics (3), oral presentation at UROP Symposium (Fall 2023, Spring 2024)
- Sole technical contributor on both projects. Work included: data generation, model architectures, training & GPU parallelism, statistical analysis, identifying key experiments, investigations. Drove research direction

GT AI Safety Initiative (in collaboration with EleutherAI) — Atlanta, GA

Volunteer Researcher – Advised by Curtis Huebner — *May 2023 – Dec 2023*

- Contributed to unpaired text to image & image to text generation project. Worked with VAEs, VQ-VAEs, and vision transformers for multimodal generation

Research Projects & Preprints

1. **Bagga**, PS., Farias, V., Korkotashvili, T., Peng, T., & Wu, Y. arXiv | github: E-GEO: A Testbed for Generative Engine Optimization in E-Commerce (submitted to ACM Web Conference Short Papers Track)
2. **Bagga**, PS., Delarue, A., & Pauphilet, J. Learning Knapsack Decision Rules using Permutation-Invariant Neural Architectures. (in preparation)
3. **Bagga**, PS. & Delarue, A. (2023). arXiv | github: Solving the Quadratic Assignment Problem using Deep Reinforcement Learning. (preprint)

Work Experience

Mercury — New York, NY

Software Engineer — *Apr 2025 – Present*

- International Wires & AI Agents; invited by VP of Data Science to help drive project on AI agents for Support

Software Applications Inc. (acquired by OpenAI) — San Francisco, CA

Research Intern — *May 2024 – Jul 2024*

- Led experiments and built infra for agentic LLMs, enabling on-device inference & function calling via fine-tuning
- Developed FT pipeline with custom loss and GPU parallelism for 10k+ samples, around token enforcement specs
- Designed & trained on domain-specific language for datetime representation, achieving ∼95% on translation evals

Mercury — San Francisco, CA

Machine Learning Intern – Risk Intelligence — *May 2023 – Aug 2023*

- Developed risk-aware decision engine to automate spend limit requests, pioneering first ML application at company
- Achieved 97% accuracy on custom risk and confidence thresholds, saving 16+ hours/month on spend limit requests
- Boosted intermediate performance 15% via email data processing using GPT-4 API and N-shot prompting

Honors & Awards

- Graduated with Highest Honors – Georgia Institute of Technology (Dec 2024)
- Dean's List Faculty Honors – All semesters (2022-2024)
- Best Use of Google Cloud – GT Hacklytics, NLP-powered climate news search engine (Feb 2023)
- USA Computing Olympiad – Platinum Division (Mar 2022)

Skills

Languages: Python (PyTorch, numpy, pandas, sklearn, XGBoost, FastAPI), Haskell, SQL, Matlab, Julia, C++
Expertise: Machine Learning, Mathematical Optimization