YASH GUPTA

📞 +1 (412) 909 6040 | 🌐 github.io | ✉ yashgup2@cs.cmu.edu | in yashgupta1427 | ⌨ yashgupta-7 | 🎓 scholar.google.com

EDUCATION

Carnegie Mellon University — **Pittsburgh, PA**
M.S. in Machine Learning, GPA: 4.28/4.33 (Major), 4.23 (Overall) — Aug 2022 – Dec 2023
- Courses: Advanced Machine Learning, Statistics, Regression, Deep Learning, Graphical Models, Convex Optimization

Indian Institute of Technology (IIT), Bombay — **Mumbai, India**
B.Tech (Honors) in CS with Minor in Data Science, GPA: 9.92/10.00, Rank: 5/1025 — Aug 2018 – May 2022
- Institute Academic Prize (Top 10), Excellent TA Award, Bronze Medal at Inter-IIT Tech Meet
- Courses: Probability, Statistics, Machine Learning, Deep Learning, NLP, Data Structures, Linear Algebra, Algorithms

AWARDS

Competitions Citadel Midwest Datathon (**1st**, won $10,000, 2023) 📄; IIT (India Rank 19/1.2M (Main), 42/200K (Advanced), 2018);
Olympiads Asian Physics Olympiad (**Bronze** Medal, Vietnam, 2018); International Junior Science Olympiad (**Gold** Medal, South Korea, 2016); Physics & Astronomy Indian Team Selection Camps (Top 39/200K, 2018)
Scholarships Narotam Sekhsaria Fellow (Top 13/5,000, $60K No Interest Loan, 2022); KVPY Fellow (India Rank 4/50K, 2018)

EXPERIENCE

Susquehanna International Group — **Philadelphia, PA**
Incoming Quantitative Research Intern — Jun 2023 – Aug 2023

Haptik AI — **Mumbai, India**
Machine Learning Intern — Jun 2022 – Aug 2022
- Enhanced chatbot performance by incorporating table retrieval and question answering, increasing F1 score by 7%
- Engineered ensemble models for dealing with diverse types of structured data and tabular matrices in client documents

Tower Research Capital — **Gurugram, India**
Quantitative Strategist Intern — Dec 2021 – Jan 2022
- Developed automated strategies to classify individual market signals into different feature classes
- Designed algorithms to learn data transformations for each class, and improved linear regression fits across several assets

Samsung Research — **Suwon, South Korea**
Machine Learning Intern — Jun 2021 – Jul 2021
- Modelled recommender system techniques to develop systems for predicting user-item interactions for Samsung smart appliances

Technische Universitat Braunschweig — **Braunschweig, Germany**
Algorithms Research Intern — May 2020 – Jul 2020
- Advanced on proving NP-hardness of geometrically packing a set of squares into a disc by a reduction from 3-PARTITION
- Resolved an open problem in tile assembly, about sufficiency of straight 2-cuts for dividing polyominoes into convex sub-regions

RESEARCH PROJECTS & PUBLICATIONS

Responsible AI Games & Ensembles (Submitted to NeurIPS 2023) — **Statistical ML Group, CMU**
- Developed a minimax framework to solve games arising in responsible AI (RAI) contexts such as robustness, fairness, & tail risk
- Proposed algorithms for solving these RAI games motivated by the classical statistical literature on boosting, and regression

Question Answering over Tables & Text (Published at **ACL 2023**) — **IBM Research**
- Devised robust scalable systems for question answering over both structured tabular and unstructured textual forms
- Introduced a training strategy for multiple instance & multiple answers based weak supervision, improving performance by 21%

Tapestry: Pooled RT-PCR Testing (Published at **IEEE OJSP 2021**, IEEE SPS Top 25) — **IIT Bombay**
- Engineered a single-round pooled testing approach along with a team of 15, leading to quick testing & conserved reagents
- Improved over two-round pooling by a factor of 2 in the number of tests required, got recognized by Nature Magazine & DCGI
- Finalist Team (Top 10 in Open Innovation) at the **$6 Million XPRIZE** Rapid Covid Testing Competition

Pre-Training for Text Summarization (Published at **NAACL SDP 2021**) — **Verisk Analytics**
- Designed techniques to achieve state-of-the-art summarization task performance on low resource datasets
- Derived 4% absolute improvement using intermediate pre-training & custom reinforcement learning rewards to encourage saliency

SKILLS

Languages (*Intermediate*) C++, Python, (*Familiar*) SQL, MATLAB, Racket, Javascript, LaTeX, HTML, Yacc, Bash
Tools/Frameworks Pytorch, Pandas, Numpy, Tensorflow, ReactJS, Django, Git