Soham Gupta

224-283-5180 | sohamgupta@college.harvard.edu | www.linkedin.com/in/sohamgupta21 | github.com/SohamGupta21

Education

Harvard University Aug. 2023 – May 2027

Bachelor of Arts in Computer Science, Statistics *GPA: 3.8*

Relevant Coursework: Probability (Stat 110), Introduction to Statistical Inference (Stat 111), Introduction to Machine Learning (MIT 6.390), Topics in Machine Learning (CS 282r), Data Structures and Algorithms (CS 124), Systems Programming and Machine Organization (CS 61)

Experience

Quantitative Developer and Research Intern May 2024 – Present

Zimmer Partners *New York, NY*

- Engineered and deployed a low-latency web platform leveraging Redis, cutting data retrieval times by 500x and enhancing real-time analytics for a $10B+ hedge fund. Used Python, TypeScript and Java.
- Developed Airflow batch pipelines to ingest and process large-scale Bloomberg and alternative datasets, leveraging dynamic task mapping to cut runtime from 2 hours to under 2 minutes.
- Designed a model to create options volatility surfaces to price OTC options, reducing analysis time by over 5x.

President Sep. 2023 – Present

Harvard Sports Analytics Collective *Cambridge, MA*

- Conducted advanced statistical modeling on sports datasets, applying regression analysis, time-series forecasting, and Bayesian inference to evaluate player and team performance.
- Published 2 data-driven research articles using machine learning techniques (SciKit-Learn, PyTorch) to generate predictive trends and identify market inefficiencies.
- Organized and hosted 15+ speaker events featuring executives from top organizations, including ESPN, NBA, and NFL, fostering industry connections and professional development for members.

Software Engineering Intern Aug. 2024

Onyx Odds *New York, NY (Remote)*

- Launched a Sports Sweepstakes product backed by XFund, achieving 1,000+ user signups within the first month.
- Deployed a high-performance mobile-web application by integrating React Native with MongoDB and Node.js, reducing load times by 30%.
- Implemented AI identity verification system, bringing user authentication accuracy to above 95%.

Data Science Contractor Sep. 2023 – Dec. 2023

Boeing, SharkNinja *Remote*

- Conducted market analysis and growth forecasting using advanced methods such as Monte Carlo simulations on market data in 15+ industries across dozens of verticals.
- Partnered with C-Suite executives to analyze over 6 million consumer inquiries, leveraging NLP techniques such as sentiment analysis and concept extraction to inform decisions and reducing the volume of data by 74%.
- Led a 6-member Data Science team on a 4-month project, insights improved operational efficiency by over 24%.

Projects

SpikeTrade: Stock Market Simulation | *Python, PyTorch, BindsNET, Neuroscience* Jan. 2025 – May 2025

- Built a large-scale stock market simulator integrating Spiking Neural Network (SNN) agents, rule-based traders, and stochastic noise traders to mimic real-world investor behavior.
- Modeled retail decision-making using biologically inspired SNNs trained on price, sentiment, and earnings data; achieved over 70% R^2 in Finance and Consumer sectors.
- Designed a real-time Streamlit dashboard for interactive experimentation with trader composition and simulation parameters.

Systematic Bias-Probing with LLMs | *Python, LLMs, Numpy, Flask* Oct. 2024 – Dec. 2024

- Developed a mathematical framework to integrate 3 bias probing methods, analyzing over 1M tokens.
- Built an interactive web tool using React and Flask, enabling users to identify bias in pre-trained models across predefined attributes within 5 minutes, improving evaluation speeds by 200x.
- Presented research on bias detection in LLMs to 50+ attendees, demonstrating a contrastive framework leveraging Bayesian statistics to identify systematic bias.

Technical Skills

Languages: Python, JavaScript, Swift, C++, Java, TypeScript, HTML/CSS, R, OCaml
Frameworks: React, Node.js, Flask, Django, FastAPI, React Native, JavaFX, Next.js, Nest.js
Developer Tools: Git, Docker, GCP, AWS, Azure, Redis, SQL, MongoDB, Firebase
Libraries: Tensorflow, PyTorch, pandas, NumPy, Matplotlib, SciKit-Learn, Keras, NLTK