Saisrijith Reddy Maramreddy

shrey.maramreddy@gmail.com | +1 929-305-8774 | LinkedIn | GitHub | Portfolio

Education

Baruch College – Zicklin School of Business — New York, NY
M.S. Statistics & Data Science | GPA 3.95 — *08/2024 – 05/2026*

Imperial College Business School — London, UK
M.Sc. Investment & Wealth Management | Merit Classification — *09/2021 – 09/2022*

Pennsylvania State University — University Park, PA
B.S. Industrial Engineering | GPA 3.83 — *08/2016 – 05/2020*

Experience

CUNY Research Foundation | Research Assistant, Prof. Zeda Li | New York, NY — 11/2025 – Present

- Built **ANNOTA**, a PDF-to-code research pipeline: dual-parser ingestion (PyMuPDF + IBM Docling), GPT-based method normalization, ChromaDB vector retrieval, and auto-generated LaTeX study guides — producing 266 structured method records and 261 runnable Python/NumPy estimators behind a unified `BaseMethod` interface (867 registry entries across 7 statistical families).
- Built Monte Carlo simulation + benchmark infrastructure across 7 simulator tracks; ran the `spectral_transfer` benchmark on the Docling-extracted estimators — 6,000 synthetic brain-connectivity graphs over 8 experimental regimes (sample size, domain shift, source contamination) comparing 5 transfer-learning variants, scoring F1, precision-recall, SHD, and normalized Frobenius error.

BCITS Pvt Ltd | Data Analyst Intern | Remote — 09/2022 – 07/2023

- Engineered ETL pipelines over 50K+ billing, IoT, and transaction records in R/SQL, eliminating reconciliation errors and lifting billing accuracy 15%.
- Shipped KPI dashboards that cut manual reporting ∼8 hrs/week and drove a 20% customer-satisfaction gain post-rollout.

Projects

AI Blocks: "Scratch for AI engineering" — DAG-based composer of 505 adaptive blocks that auto-maps a repo, wires files to blocks, flags missing infra, and cuts LLM tokens ∼80% vs. Opus 4.6 via embedding-grounded retrieval. Built in TypeScript/Next.js with MCP. *1st place, VibeForward × Lovable Hackathon (NYC 2026).*

Situational Intelligence: Real-time AI emergency monitor streaming WebRTC frames to Claude Vision for motion/threat scoring, with speech-triggered dispatch and Leaflet-based responder routing — built end-to-end in 6 hours. *1st place, Grayscale Hackathon (NYC 2026).*

Trace: 24/7 autonomous supply-chain agent monitoring inventory and disruption signals, orchestrating vendor payments via API with human-in-the-loop approval gates. *JavaScript, LLM Agents, Payments API.*

MIRA Stylist: Luxury AI fashion companion with conversational style onboarding, computer-vision virtual try-on, and animated editorial look generation. *Python, OpenAI, Computer Vision, FastAPI.*

M.I.R.A: Multi-agent voice assistant with wake-word detection, Whisper streaming ASR, and Cartesia TTS; an LLM router dispatches sub-agents including a Playwright browser agent for natural-language control of live web workflows. *Python, OpenAI, Whisper, Cartesia, Playwright.*

CorVas: Senior-friendly cardiac-recovery PWA (live in production) with 12-week rehab plans, AI symptom check-ins, medication tracking, and care-team coordination. *Next.js, TypeScript, LLM APIs.*

Octagon Intel: Calibrated UFC prediction platform with prefight-only XGBoost classifier, probability calibration, leak-free temporal splits, and Kelly-criterion sizing against live sportsbook odds; Next.js dashboard backed by FastAPI inference. *Next.js, FastAPI, XGBoost, scikit-learn.*

WalkWithMe: Pedestrian-first iOS navigation with on-device CoreML hazard detection, ARKit turn-by-turn overlays, conversational "Loop Assistant," and GPX-backed themed walks. *SwiftUI, CoreML, ARKit, FastAPI.*

Statistical Learning Projects: Quant Alpha Pipeline, Property Valuation Modeling, FairStacks NBA Draft Predictor, and Ridge Bias-Variance Simulation — applied ML across equity alphas, real-estate valuation, fairness-aware classification, and regularization theory. *Python, XGBoost, LightGBM, Optuna, SHAP, CVXPY, HMM.*

Elo-Based Forecasting: Sports outcome prediction combining dynamic Elo ratings with time-series modeling for live win-probability estimation. *Python, Time Series, Elo.*

CineSeq: Decay-aware multivariate Seq2Seq box-office forecasting with attention-augmented LSTM-GRU, exponential-decay priors, and trailer-emotion embeddings. *PyTorch, LSTM/GRU, Attention.*

Skills

Languages: Python, TypeScript, Swift, R, SQL, SAS
ML/AI: PyTorch, scikit-learn, XGBoost, LightGBM, Optuna, SHAP, Seq2Seq, LSTM/GRU, Attention, Graphical Lasso, Ridge/Lasso, SVM, LDA
LLM Systems: OpenAI, Claude Vision, LangChain, LangGraph, Whisper, Cartesia, RAG, Embeddings, MCP, Playwright agents
Infra & Product: FastAPI, Next.js, React, SwiftUI, ARKit, CoreML, Streamlit, Vercel, PostgreSQL, WebRTC, Git
Data: Pandas, NumPy, feature engineering, time-series, multivariate methods, calibration, Matplotlib/Seaborn