Aravind Kannappan

408-591-5449 | aravinds.kannappan@gmail.com

Education

New York University — **New York, NY**
Master of Science in Applied Statistics, Specialization in Machine Learning — *Expected Graduation May 2026*

Baylor University — **Waco, TX**
Bachelor of Science in Statistics, Minor in Biology — *Aug 2020 – Aug 2024*

Work Experience

Machine Learning Engineer Intern — **New York, NY**
Icahn School of Medicine at Mount Sinai — *Apr 2025 – Present*

- Processed 200K+ patient records with SQL and Python ETL for insurance modeling, improving data reliability
- Built a PyTorch transformer–based personalized insurance recommender optimizing coverage for 15K+ patients
- Implemented SHAP-based explainability highlighting key cost drivers across patient cohorts
- Developed a RAG system over insurance policy documents to deliver policy-aware recommendations to customers
- Built open-source tool for automated insurance plan comparisons across thousands of coverage–cost configurations.

Biostatistics Intern — **Houston, TX**
Baylor College of Medicine | *Publication* — *July 2022 – Dec 2024*

- Analyzed 2,000+ genomic profiles using Python to identify biomarkers correlated with differential treatment outcomes
- Benchmarked regression, random forest, and gradient boosting models in scikit-learn using nested cross-validation
- Selected a final predictive model achieving 0.81 AUC and contributed to a peer-reviewed publication in *The Oncologist*

Research Experience

CSCI 2271 Computer Vision — **New York, NY**
NYU Courant Institute — *Aug 2025 – Dec 2025*

- Built diffusion-based world model on 737K+ Super Mario Bros frames, propagating gameplay dynamics
- Trained VAE for state-to-observation mapping, achieving 15.7 dB PSNR and 0.71 SSIM visual fidelity across frames
- Developed CNN reward model predicting game outcomes, achieving 0.61 AUC, enabling gradient-based difficulty
- Trained PPO agents (10M steps) for trajectory collection and policy optimization in diffusion-simulated environments
- Incorporated Hidden Markov Models for skill-based level design, enabling adaptive transitions between difficulty levels

Projects

Fantasy Oracle - Web App — *Nov 2025*

- Built Fantasy Oracle, an AI co-pilot for fantasy sports strategy, drafts, and trades across major leagues
- Conducted 1K user interviews with commissioners and power-users to validate product-market fit
- Built multi-agent LLM system with league integrations for autonomous waiver, trade, and lineup optimization

PersonaDx - Open Source — *Apr 2025*

- Built LangChain multi-agent LLM pipeline for diagnostic insight extraction from patient narratives
- Implemented agent-coordination logic in Python, integrating OpenAI GPT-4 and BioBERT for symptom extraction
- Developed React/TypeScript and D3.js dashboard for actionable clinical recommendations from complex patient cases
- Open-sourced FastAPI/Streamlit and Docker stack for plug-and-play EHR integration

TrafficFlowOpt - IOS App — *Oct 2024*

- Built city-scale traffic simulator in C++17 with CUDA-parallel graph processing, delivering 15–25× speedup vs CPU
- Designed PyTorch autograd signal-timing optimizer, cutting network-wide delay 23% on real Chicago datasets
- Engineered Vue.js + Mapbox GL dashboard for live congestion heatmaps and real-time signal-phase recommendations
- Containerized full pipeline with CMake, Docker, and NVIDIA runtime for one-click municipal deployment

Technical Skills

Programming Languages: **Advanced**: Python, R, SQL, **Intermediate**: Typescript, C++, CUDA
ML/Analytics: PyTorch, Tensorflow, Scikit-learn, XGBoost, Pandas, NumPy, Tableau, Power BI, Hugging Face
Cloud/Tools: AWS (S3, SageMaker, Lambda), Docker, Github, Airflow, Apache Spark
Microsoft Proficiency: Excel(Advanced), Powerpoint(Advanced), Word(Advanced)