Guillermo Marr

Contact: GUILLERMOMARR3@GMAIL.COM

linkedin: https://www.linkedin.com/in/guillermo-marr-942925190/

EDUCATION

FLORIDA STATE UNIVERSITY

BACHELORS DEGREE: COMPUTER SCIENCE & STATISTICS   *Expected graduation Dec 2025*

MAJOR GPA (CS): 3.7 · MAJOR GPA (STATS): 4.0

Relevant Coursework: Data Structures and Algorithms I & II, Probability, Mathematical Statistics, High Performance Computing (CUDA, OpenMP, MPI), Deep and Reinforcement Learning, Pattern Recognition, Computational Probabilistic Modeling, HITM Research, Machine Learning, Software Engineering

EXPERIENCE

UNDERGRADUATE COMPUTER SCIENCE RESEARCHER (QUANTITATIVE RESEARCH ON CRYPTOCURRENCY MARKETS)   *January 2025 - Present*

- Conducted topological data analysis (TDA) on limit order book (LOB) data downsampled to 1-minute bars, extracting persistent homology features alongside smart price indicators and microstructure metrics (e.g., order imbalance, depth asymmetry).

- Modeled future expected returns, mid-price at t+1, and bid-ask spread using extracted features; applied HDP-HMM (Hierarchical Dirichlet Process Hidden Markov Model) to identify latent regimes in microstructure dynamics via Gaussian-distributed features.

- Utilized non-linear models (XGBoost/LightGBM) to visualize feature importance and relationships, predicting mid-price, expected value, and spread; incorporated meta-labeling to identify optimal trade entry/exit points based on prediction confidence.

- Performed exploratory data analysis (EDA) and statistical modeling on features as part of a full quantitative research pipeline, including hypothesis testing, correlation analysis, and regime-aware validation.

- Investigated supplementary factors such as bridge volume and sentiment signals to model potential liquidity gaps and their probabilities, enhancing regime detection for volatility forecasting (ongoing).

SOFTWARE ENGINEER INTERN   *June 2024 - August 2024*

L3 Harris Technologies

- Developed high-performance backend systems in C++ within a Scrum Agile environment, utilizing object-oriented design and advanced data structures to support distributed aerospace applications.

- Engineered low-latency inter-component communication using C++ socket programming and XML-based configuration, processing real-time data streams and reducing latency by 15%.

- Implemented Git-based version control and CI/CD pipelines (Jenkins, Google Test) to maintain code quality and reliability in mission-critical software.

- Optimized system performance through memory management and concurrency techniques, enhancing scalability for large-scale data processing.

- Received Full Time Return Offer (Declined to Pursue Dual Major Stats and Honors Research)

AI/ML RESEARCH PARTICIPANT AI4ALL IGNITE   *September 2024 - December 2024*

AI4ALL

- Collaborated with a team to train a Convolutional Neural Network (CNN) on a dataset of 50,000 medical images, achieving 92% accuracy in tumor detection using Python and TensorFlow.

- Partnered with industry mentors from DoorDash and Databricks to apply best practices in data ingestion, statistical analysis, and model optimization for large-scale datasets.

- Participated in weekly sessions on advanced AI topics, including time series analysis and reinforcement learning, enhancing my quantitative research skills.

PROJECTS

Regime-Aware Anomaly Detector

HMM, DBSCAN, Isolation Forest

Built a hybrid detector that first segments price/flow data into latent regimes (HMM), then runs DBSCAN + Isolation Forest inside each regime to slash false-positives in whale-dump alerts by ≈40%. Demonstrates that context-aware anomaly filters beat global thresholds on volatile assets.

Distributed ML Optimiser

CUDA, MPI, PyTorch

Prototyped gradient-aggregation kernel that balances GPU/CPU load across a 4-node cluster; cuts wall-clock training time 30% on a ResNet-50 while holding accuracy. Showcases HPC tool-chain for data-heavy models.