

VARUN SINGHAL

San Diego, CA | (858) 444-0091 | varunraisinghal12@gmail.com | linkedin.com/in/varunraisinghal/ | github.com/varunraisinghal/

EDUCATION

University of Southern California Masters of Science in Computer Science: Artificial Intelligence Concentration: Artificial Intelligence, Machine Learning, Quantum Computing + Cryptography	Los Angeles, CA January 2025 – December 2026
University of California, San Diego Bachelors of Science: Computer Science GPA: 3.83/4.00, Provost Honors (2020 - 2024)	San Diego, CA October 2020 - June 2024

PROFESSIONAL EXPERIENCE

GEICO Machine Learning Engineer	San Francisco, CA May 2024 - August 2025
<ul style="list-style-type: none">Designed and developed quote estimation service leveraging vehicle/driver features and historical premiums; improved latency from ≈30s to <13s, cut compute costs by 50+%, and met <15s Client SLAProactively initiated and guided a cross-team effort to design and implement scoring models for the quoting service, enabling precise premium estimates at scale and enhancing overall system performanceBuilt automated commission-calculation tool to parse spreadsheets, apply pricing algorithms, and generate cost/commission breakdowns with 0% error; reduced manual turnaround from ~3-4 days to <1 hourProductionized models and pipelines with AWS SageMaker, Lambda, DynamoDB, REST APIs, and CI/CD for versioned deployments; improved API throughput, reliability, and latencyAuthored API docs, internal wiki pages, video tutorials, and conducted trainings to onboard partner-team engineers and new hires; reduced questions, onboarding time, and accelerated adoption across 5+ teamsRecognized by VP of Engineering for leading quote-to-bind workflow initiatives and delivering measurable savings	
Amazon: Amazon Web Services (AWS) Software Engineering Intern	Seattle, WA May 2023 - August 2023
<ul style="list-style-type: none">Provisioned production-grade databases and storage pipelines for enterprise clients (e.g., Fortune-500) using AWS services; ensured seamless integration with clients' existing environmentsLed end-to-end onboarding and trainings for 5 client workloads: set up schema, access controls, data pipelines; ran integration tests to validate data flow and compliance before go-liveImplemented automated health-checks and logging with CloudWatch and GuardDuty; reduced failure detection time by ~40% and manual triage effort for support teamEarned AWS Certified Cloud Practitioner and AWS Certified Developer credentials during internship	

TECHNICAL PROJECTS

Project: Low Latency Arbitrage Bot [C++]

- Engineered an optimized graph-theory algorithm to detect profitable cryptocurrency arbitrage opportunities across major exchanges (Binance, Coinbase, Kraken) with **<10 ms detection latency**
- Built a **custom order-management system (OMS)** leveraging exchange APIs to deliver real-time market-data retrieval and **<300 ms order-execution latency**, maximizing trade value across exchanges

RELEVANT SKILLS

- Machine Learning/AI:** Python (NumPy, Pandas), PyTorch, TensorFlow, Keras, scikit-learn, XGBoost, LightGBM, Hugging Face Transformers (BERT, RoBERTa), spaCy, NLTK, MLflow, LangChain, FAISS/vector databases, LLM fine-tuning, Prompt Engineering
- Computer Vision:** OpenCV, CNNs, Faster R-CNN, U-Net / Mask R-CNN, object detection & segmentation, Albumentations
- Data Science/Statistics:** Feature Engineering, EDA, A/B Testing, Bayesian Estimation, Regression & Classification, Gradient Boosting, Clustering (K-Means, GMM), Dimensionality Reduction (PCA, t-SNE, UMAP), Recommendation Systems
- Cloud/MLOps:** AWS, Docker, Kubernetes, Terraform, Git Action, CI/CD, Agile/Scrum
- Programming Languages:** Python, C++, Java, SQL, R, Go, JavaScript, Solidity, Rust
- Data Engineering/Backend:** REST APIs, GraphQL, Kafka, Redis, ETL Pipelines, Spark/Hadoop
- Algorithms/CS Fundamentals:** Graph, Tree, Hash Table, Priority Queue, Dynamic Programming, A*, Monte Carlo Tree Search
- Tools/Visualization:** Jupyter Notebooks, VS Code, Git/GitHub, Matplotlib, Seaborn, Plotly