VINAY JOGANI

joganivinay@gmail.com | (617)-992-4663 | San Francisco, CA | Linkedin | Github | Portfolio

EDUCATION

Northeastern University, Boston, MA   *September 2023 - August 2025*
Master of Science in Information Systems
Focused Coursework: Advance Data Science & Architecture, Parallel Machine Learning & AI, LLM with Knowledge Graph DB, Natural Language Engineering, AI Generative Modeling with focus in Finance

Veermata Jijabai Technological Institute, Mumbai, India   *August 2019 - June 2023*
Bachelor of Technology in Information Technology
Focused Coursework: Data Structures and Algorithms, Linear Algebra, Discrete Mathematics, Artificial Intelligence, Machine Learning

SKILLS

Programming Languages: Python, SQL, Objective C, Swift, C++, Java, R, JavaScript, MATLAB, Cypher
Machine Learning & AI: TensorFlow, PyTorch, Keras, Scikit-learn, XGBoost, LightGBM, Stable-Baselines3, OpenAI Gym, Gymnasium, Supervised Learning, Unsupervised Learning, Transfer Learning, Reinforcement Learning, Ensemble Methods, Hyperparameter Optimization
Deep Learning: CNN, LSTM, GRU, ResNet, DenseNet, EfficientNet, VGG, Vision Transformers, DQN, PPO, Distributed Data Parallel, Multi-GPU Training, OpenCV, torchvision, scikit-image, Grad-CAM, Image Segmentation, Object Detection
Natural Language Processing: BERT, RoBERTa, Transformers (Hugging Face), NLTK, spaCy, CodeT5, VADER, TextBlob, NER
Data Science & Analytics: NumPy, Pandas, SciPy, statsmodels, Statistical Analysis, A/B Testing, Hypothesis Testing, Experimental Design, Time Series Forecasting, Feature Engineering, PCA, t-SNE, UMAP, Causal Inference, Monte Carlo Simulations
Data Engineering & ETL: Apache Spark, PySpark, Apache Kafka, Apache Airflow, dbt, Hadoop, Dask, ETL/ELT Pipelines, Stream Processing, Batch Processing, Data Orchestration, Great Expectations, Pydantic, Medallion Architecture
Databases & Data Storage: PostgreSQL, MySQL, MongoDB, Neo4j, SQLite, Snowflake, BigQuery, Redis, SQLAlchemy, FAISS, Database Design, Window Functions, CTEs, Indexing
Cloud Platforms: AWS (EC2, S3, Lambda), GCP (BigQuery, GCS, Cloud Pub/Sub)
MLOps & Production: MLflow, Docker, Kubernetes, CI/CD, GitHub Actions, Model Monitoring/ Versioning, Prometheus, Grafana, Terraform
Software Engineering: FastAPI, Flask, REST APIs, Microservices, Git, Unit/Integration Testing, Async/Await, Agile/Scrum

PROFESSIONAL EXPERIENCES

Research Assistant | Amal Lab, Northeastern University | Boston, MA   *June 2024 - August 2025*
- Engineered Med-SAM medical image segmentation system for multi-modal datasets (MRI, CT, histopathology) with automated preprocessing pipelines, achieving superior boundary detection through fine-tuned transformer architectures on clinical datasets
- Architected scalable big data infrastructure processing 18TB+ TCGA multi-omics datasets using distributed computing frameworks (Apache Spark, Dask), implementing unsupervised clustering and dimensionality reduction for biomarker discovery across 33+ cancer types
- Developed production-ready deep learning classification system achieving 97.28% accuracy on skin cancer detection through comparative analysis of state-of-the-art architectures (ResNet-34, EfficientNet-B1, Vision Transformers) with ensemble learning and external validation

Research Trainee | Brigham and Women's Hospital | Boston, MA   *August 2024 - December 2024*
- Architected meta-analysis framework processing ClinicalTrials.gov and PubMed databases using machine learning algorithms (random forests, gradient boosting) with automated data extraction pipelines, implementing statistical modeling and maintaining HIPAA compliance
- Engineered clinical trial news analytics pipeline analyzing 10,000+ pharmaceutical reports using transformer-based NLP models (BERT, RoBERTa) to identify correlations between media coverage sentiment and regulatory failure rates

TECHNICAL PAPERS

Analysis of Explainable AI Methods on Medical Image Classification | Link   *May 2023*
2023 3rd International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies
- Examined explainable AI (XAI) methods for medical image classification using CNN models (VGG-16, ResNet-50), evaluating computational efficiency and interpretability tradeoffs
- Benchmarked performance of Grad-CAM, Integrated Gradient, and LIME approaches through execution time analysis and memory profiling, demonstrating implications for production AI systems

Adversarial Attacks and Defenses for Skin Cancer Classification | Link   *April 2023*
2023 2nd International Conference for Advancement in Technology
- Investigated adversarial attacks (PGD, FGSM) on CNN-based skin cancer classification, demonstrating model vulnerabilities to adversarial examples through systematic attack evaluation and accuracy degradation analysis
- Applied PGD-based adversarial training to enhance model robustness, achieving 27.73% point improvement against adversarial attacks

PROJECTS

Healthcare Data Pipeline: Real-Time Analytics Platform | Link   *June 2025 - October 2025*
- Architected production-grade data pipeline processing 10M+ patient records with real-time streaming via Apache Kafka and Spark Structured Streaming (500+ events/minute) with exactly-once semantics and stateful aggregations, plus batch ETL achieving 70% processing time reduction and 85% materialization optimization, orchestrated via Apache Airflow with retry logic, SLA monitoring, and custom operators
- Implemented Spark-based data quality and observability infrastructure with validation jobs across 6 tables with Prometheus metrics and Grafana dashboards tracking pipeline throughput, processing latency (avg 3.2s), system resources, maintaining 99%+ pass rates with automated alerting
- Designed cloud-native deployment with Docker containerization (12-service stack), Kubernetes orchestration, Terraform infrastructure-as-code provisioning GCP resources (BigQuery, GCS, Cloud Pub/Sub) with role-based access control, and comprehensive testing suite

Personal Finance API: Microservices Architecture & Infrastructure | Link   *September 2025 - October 2025*
- Architected production-grade microservices system with 4 independent FastAPI services implementing RESTful architecture with JWT-based authentication, achieving 92% test coverage across 80+ pytest unit/integration tests and automated CI/CD pipeline with security scanning and coverage enforcement on every GitHub Actions deployment
- Engineered high-performance backend infrastructure with PostgreSQL cluster (indexed queries achieving sub-50ms response times), Redis caching reducing API latency by 30% with 85%+ cache hit rates, Nginx load balancing, and connection pooling handling 500+ users in testing
- Designed scalable deployment with Docker containerization (health checks), Kubernetes orchestration (auto-scaling, resource limits), observability stack (Prometheus, Grafana), and infrastructure-as-code demonstrating production DevOps practices

Skin Cancer Classification using High Parallel Machine Learning | Link   *March 2025 - April 2025*
- Implemented deep learning pipeline for dermatological disease classification using EfficientNet-B3 architecture, achieving 91% multi-class accuracy across 35 skin conditions on 245,000-image dataset, supporting automated medical diagnosis applications in clinical dermatology
- Accelerated model training by 3.3× through distributed computing infrastructure utilizing PyTorch Distributed Data Parallel framework across 4× NVIDIA A100 GPUs with 84% parallel efficiency; conducted comprehensive performance profiling and scalability analysis to identify computational bottlenecks in training workflows and data loading processes, systematically optimizing mixed-precision inference, memory usage, and resource utilization across GPU infrastructure