Shivasankaran Vanaja Pandi

631-949-0169 | shivasankaranvanajapandi@gmail.com | linkedin.com/in/shivasankaran-vp | github.com/Shiva-sankaran Education

Stony Brook University, SUNY

Master of Science in Computer Science: GPA: 3.97

Thesis: Robust Particle Detection for Cryogenic Electron Microscopy.

• Coursework: Computer Vision, Machine Learning, Natural Language Processing, Distributed Systems.

Indian Institute of Technology (IIT)

Bachelor of Technology with honors in Computer Science; GPA: 3.8

EXPERIENCE

Machine Learning Engineer | Healthcare Data Integration, Data Quality, Python Sep 2024 – May 2025 Deep Forest Sciences Remote

- Supported **integration of biomedical datasets** from multiple institutions, validating quality and completeness through profiling that resolved **94% of inconsistencies** before deployment.
- Maintained **business rules for data transformation** using Python and SQL, ensuring accurate mapping of clinical terminology and achieving 99.2% data validation accuracy across diverse healthcare formats.
- Collaborated with **interdisciplinary technical teams** to gather requirements for data structure analysis, providing clear documentation and process guidelines that reduced onboarding time for new datasets by 45%.

Software Engineer (Open-Source) | Data Analysis, Documentation, Collaborative May 2024 – Aug 2024 Google Summer of Code – DeepChem Remote

- Assisted in validating scientific datasets for machine learning applications, conducting systematic data profiling and quality assessments that ensured comprehensive data completeness across 12 different databases.
- Contributed to clear documentation of analytical processes and best practices, creating detailed guides and workflow documentation that improved team knowledge sharing and reduced project setup time by 38%.
- Demonstrated **methodical problem-solving approach** while working with remote teams across multiple time zones, maintaining excellent communication and delivering consistent results in collaborative environment.

Projects

Real-Time Event Processing System | Python, C++, Event-Driven Architecture, Systems Jan 2025 – Apr 2025

- Architected an event-driven financial data processing system using Python asyncio and C++ extensions, handling market data streams with **millisecond processing latency** and supporting concurrent processing.
- Implemented distributed message queuing with Redis and custom serialization protocols, enabling reliable event delivery across multiple processing nodes with automatic retry mechanisms and dead letter queue handling.
- Applied **test-driven development methodology** throughout the project lifecycle, writing comprehensive unit tests and integration tests that validated system behavior under various market conditions and stress scenarios.

Distributed Portfolio Analytics Platform | Python, C++, Microservices, Software Testing Sep 2024 – Dec 2024

- Built a scalable microservices architecture for portfolio risk analysis using Python Flask and C++ computation engines, supporting **real-time analytics** for 1000+ securities with horizontal scaling capabilities.
- Designed event-driven communication patterns between services using message brokers and RESTful APIs, ensuring loose coupling and enabling independent deployment cycles while maintaining data consistency.
- Implemented **comprehensive testing strategies** including unit testing, integration testing, and performance testing, achieving 98% code coverage and establishing automated quality gates for continuous deployment.

High-Performance Market Data Simulator | C++, Python, Software Design, Quantitative May 2024 – Jul 2024

- Developed a realistic market simulation engine using C++ for computational core and Python for configuration management, generating synthetic market data with statistical properties.
- Applied advanced software design principles including SOLID principles and design patterns, creating a modular architecture that supports multiple market models and enables easy extension for new asset classes and trading scenarios.
- Utilized quantitative analysis techniques to validate simulation accuracy, achieving 95% correlation with historical market statistics and enabling robust backtesting capabilities for trading strategy development.

Technical Skills

Languages: Python, C++, SQL, JavaScript, Bash

Software Development: Test-Driven Development, Refactoring, Design Patterns, Code Review, Version Control Architecture & Systems: Event-Driven Architecture, Distributed Systems, Microservices, RESTful APIs Tools & Frameworks: Docker, Flask, FastAPI, Redis, PostgreSQL, Git, CI/CD, Linux

Quantitative Skills: Statistical Analysis, Mathematical Modeling, Algorithm Design, Performance Optimization

Achievements

- Top 100 in Joint Engineering Examination (Math + Physics) among 1.2 million students
- Published first-author research papers at top-tier conferences including AAAI, EMNLP, WACV, and ISBI
- Google Research Travel Award

Stony Brook, NY Aug 2023 - May 2025

Gandhinagar, India Aug 2019 - May 2023