SATISH REDDY VELAGALA

Nashville, TN | velagalasr1984@gmail.com | 615-609-1243

Portfolio: https://velagalasr.github.io/
LinkedIn: https://linkedin.com/in/satishvelagala
GitHub: https://github.com/velagalasr | **Hugging Face:** https://huggingface.co/velagalasr

Professional Summary

AI Engineer specializing in AI-native security and anomaly detection systems from 0→1 prototypes to production scale platforms. Expert in multi-agent architectures (LangGraph, LangChain, AutoGen) with deep GenAI/LLM expertise, RAG pipelines, and evaluation-driven development. Proven track record shipping enterprise AI solutions on Azure, AWS, and GCP serving 500+ users with measurable business impact. Strong background in threat detection, fraud prevention, and real-time anomaly identification with regulatory compliance expertise. Skilled at leading cross-functional initiatives, translating complex security workflows into AI-powered experiences, and driving clarity through ambiguity.

Technical Skills

- **Programming & Tools:** Python, Java, JavaScript, SQL
- **ML/AI Frameworks**: LangGraph, LangChain, AutoGen, CrewAI, LangSmith, TensorFlow, Keras, XGBoost, Scikit-Learn, FAISS, Spark
- **Cloud & MLOps**: Azure AI Foundry, Azure Machine Learning, Azure OpenAI Service, AWS Bedrock, AWS SageMaker, GCP Vertex AI, MLflow, Kubernetes, Docker, CI/CD, Blue-Green Deployment
- **Evaluation & Experimentation:** Custom Eval Harnesses, Precision/Recall/F1 Metrics, Groundedness Assessment, A/B Testing, Statistical Significance Testing, Model Quality Metrics, Telemetry Analysis
- **AI Concepts:** Generative AI, LLMs (GPT-4, Claude, LLaMA, Gemini), RAG, Multi-Agent Systems, Agent Orchestration, Prompt Engineering, Fine-tuning (LoRA, QLoRA), Vector Search, Content Safety & Guardrails, Hybrid AI (LLM + Rules based), Model Routing & Fallbacks, Zero-to-One (0→1) Prototyping, One-to-Many (1→N) Platform Integration
- **Security Operations:** Real-time anomaly identification, behavioral analysis, automated incident response, security event correlation, SOC automation, alert triage, threat hunting, account takeover detection, credential compromise identification, log analysis, threat intelligence integration, vulnerability assessment
- **Governance:** Model Validation, Bias Mitigation, Regulatory Compliance (GDPR, CCPA), Risk Assessment, Ethical AI, Audit Trail Management

Certifications

- AWS Certified Machine Learning – Specialty (Expected March 2026)
- Microsoft Certified: Azure AI Engineer Associate (Expected March 2026)
- Professional Machine Learning and Gen AI Engineer – Uplevel
- Certified Business Analysis Professional (CBAP) – IIBA
- Certified Agile Leader (CAL1) – Scrum Alliance

Professional Experience

Caterpillar Financial Insurance Services – AI Engineer
Nashville, TN | 11/2014 - Present

- Spearheaded 0→1 AI security initiatives from ambiguous business problems to production MVPs and led 1→N platform adoption across enterprise systems; championed evaluation-driven development culture with custom eval harnesses (precision/recall/groundedness metrics), A/B testing frameworks, and telemetry analysis requiring statistical significance before production rollouts, reducing deployment risks by 60%.
- Orchestrated cross-functional AI product delivery aligning data science, engineering, security, compliance, and business stakeholders while mentoring 12+ engineers; established AI governance frameworks with bias mitigation, content safety guardrails, model validation, and fairness assessments ensuring regulatory compliance (GDPR, CCPA, financial) with zero violations over three years.

DuPont Pioneer - Sr Software Developer
Des Moines, IA | 06/2014 - 11/2014

- Designed and prototyped UI screens, obtaining approval from stakeholders, resulting in a 15% increase in user experience and satisfaction.
- Led team in supporting UAT testing, facilitating smooth integration with backend systems, and reducing post-launch defects.

Wells Fargo & Company – Technical Lead
Des Moines, IA | 06/2012 - 05/2014

- Led cross-functional teams designing large-scale banking systems with security and compliance requirements, coordinating across engineering and risk management teams.
- Implemented document automation with API integration including security validation and audit logging, reducing processing time by 40% while maintaining compliance standards.

Tata Consultancy Services – IT Analyst & Assistant Systems Engineer
Nashville, TN | 06/2008 - 06/2012

- Delivered 8+ projects across insurance and financial services, improving process efficiency by 20%.
- Trained 30+ end-users, reducing support tickets by 40%.

ML and Generative AI Projects

Dealer Solution Bot – Multi-Agent Security & Support System

- Architected production-scale multi-agent system using LangChain and LangGraph serving 500+ dealers with 24/7 autonomous security operations; designed agentic AI architecture with specialized agents for intent classification, RAG-based knowledge retrieval, anomaly detection, and automated incident response coordinating through state management for threat identification and fraud detection.
- Built evaluation-obsessed development process with custom eval harness achieving 92% precision, 89% recall, 94% groundedness, and 85% first-contact resolution; deployed container-first architecture with blue-green deployment and hub-and-spoke model delivering 70% reduction in support tickets, 60% faster threat identification, and 40% reduction in warranty fraud.

Fraud Shield AI – Real-Time Threat Detection Platform

- Engineered production-grade hybrid ML + DL threat detection system combining XGBoost, Random Forest, and LSTMs for real-time fraudulent transaction and account takeover detection; implemented model-literate architecture with deterministic rules, feature-based detection, and sequential anomaly analysis optimized for SLO and cost targets.
- Built comprehensive evaluation framework with precision/recall/F1, ROC-AUC analysis, and rigorous A/B experimentation reducing false positives by 35% while improving recall by 25%; delivered $2M annual savings, 40% faster threat identification, 90% reduction in manual review workload, with SHAP explainability and regulatory compliance.

Security Prototype – AI-Powered SOC Analyst Assistant

- Developed rapid 0→1 proof-of-concept for AI security analyst assistant using LangChain with RAG architecture indexing 500+ security playbooks and threat intelligence feeds; integrated security tools APIs for automated threat investigation, incident summarization, and remediation recommendations.
- Established evaluation framework achieving 88% query relevance, 92% response groundedness, and 45% analyst time savings through iterative user testing with security operations stakeholders for real-time context retrieval and automated response actions.

Claims Assessment AI Support

- Developed and deployed production-scale Small Language Model (SLM) platform using fine-tuned Llama 2-7B processing insurance claims with 99.9% uptime, achieving $500K+ annual cost savings through serverless ML infrastructure and 85% reduction in operational costs.
- Engineered parameter-efficient fine-tuning pipeline using QLoRA on domain-specific insurance data, reducing claims processing time by 70% (2-3 days → 2-4 hours).

Auto mapping Engine

- Engineered AI-powered claims data mapping system using XGBoost and Python with hybrid classification architecture combining rule-based pattern matching (regex) and ML models, achieving 95%+ accuracy, 97% field detection across multiple formats, and processing 1000+ records per minute with 80% reduction in manual data entry time.
- Trained XGBoost classifier on 10K+ labeled claim records with custom feature engineering (25+ features including token patterns, positional encoding, character type ratios) achieving 93% F1-score and 15ms inference time; implemented three-tier confidence scoring system (pattern-based → XGBoost → LLM) with automated routing reducing false positives by 65% and improving model accuracy from 85% to 93%.

Education

Bachelor of Engineering, Computer Science | Anna University | 2005