SATISH REDDY VELAGALA

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

LinkedIn: linkedin.com/in/satishvelagala | **GitHub**: github.com/velagalasr | **Hugging Face:** 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

Security & Domain: Threat Detection, Anomaly Identification, Real-time Monitoring, SIEM/SOAR Workflows, Identity & Access Management, Fraud Prevention, Security Operations, Incident Response Automation, Compliance & Governance.

Programming & Tools: Python, Java, JavaScript, C# (familiar), SQL, Rust

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

- Microsoft Certified: Azure AI Engineer Associate (In Progress - Expected Dec 2025)
- Professional Machine Learning and Gen AI Engineer – Uplevel
- Certified Business Analysis Professional (CBAP) – IIBA
- Certified Agile Leader (CAL1) – Scrum Alliance

AI & GENERATIVE AI PROJECTS

Dealer Solution Bot – Multi-Agent Security & Support System (2022-2025)

- Architected and shipped production-scale multi-agent system using LangChain and LangGraph, serving 500+ dealers with 24/7 autonomous threat identification, query resolution, and automated incident response across product security, compliance violations, and warranty fraud detection.
- Designed 0→1 agentic AI architecture with specialized security agents: intent classification (threat vs. support), RAG-based knowledge retrieval, solution synthesis, escalation, and anomaly detection coordinating through state management and autonomous decision-making to identify security incidents and fraudulent patterns.
- Built evaluation-obsessed development process: custom eval harness measuring precision (92%), recall (89%), groundedness (94%), first-contact resolution (85%), false positive rate (<5%); rigorous A/B testing with statistical significance analysis before production rollout.
- Led 1→N platform integration: container-first architecture with hub-and-spoke model across dev/staging/prod, blue-green deployment for zero-downtime updates established durable API contracts and SDKs for cross-team integration.
- **Impact**: 70% reduction in support tickets, 85% first-contact resolution, 60% faster threat identification, 40% reduction in warranty fraud; SLOs for latency and cost optimization through model routing/caching

Fraud Shield AI – Real-Time Threat Detection Platform

- Engineered production-grade hybrid ML + DL threat detection system (XGBoost, Random Forest, LSTMs) for real-time identification of fraudulent transactions, account takeovers, and anomalous behavior.
- Designed model-literate architecture: evaluated LLM vs. classical ML approaches, implemented hybrid system leveraging deterministic rules for known threat patterns, XGBoost for feature-based detection, LSTMs for sequential anomalies; optimized routing to meet SLOs and cost targets.
- Built comprehensive evaluation framework: precision/recall/F1, ROC-AUC analysis, confusion matrix monitoring; reduced false positives by 35% while improving recall by 25% through iterative refinement and threshold optimization.
- Conducted rigorous A/B experimentation: multi-arm experiments comparing Random Forest baseline vs. hybrid XGBoost + LSTM; demonstrated statistically significant gains in threat detection accuracy (25% improvement) through telemetry analysis.
- **Impact**: $2M annual savings from fraud prevention, 40% faster threat identification, 90% reduction in manual review workload; responsible AI deployment with bias mitigation, fairness metrics, 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 Azure OpenAI Service and LangChain, demonstrating automated threat investigation, incident summarization, and remediation recommendations.
- Implemented RAG architecture with vector database indexing 500+ security playbooks and threat intelligence feeds; integrated security tools APIs for real-time context retrieval and automated response actions.
- Established evaluation framework: query relevance (88%), response groundedness (92%), analyst time savings (45%); conducted user testing with security operations stakeholders for iterative improvement.

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).

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, then led 1→N platform adoption across enterprise systems; established AI governance frameworks ensuring responsible deployment with bias mitigation, content safety guardrails, and regulatory compliance.
- Championed evaluation-driven development culture built custom eval harnesses with precision/ recall/ groundedness metrics, A/B testing frameworks, telemetry analysis; insisted on measurable improvements with statistical significance before production rollouts, reducing deployment risks by 60%. Designed and implemented ML infrastructure for high-throughput, multi-tenant AI serving: orchestrated model deployment on Kubernetes with auto-scaling, established monitoring pipelines with custom alerting on performance/security/cost metrics, ensured 99.9% uptime SLOs
- Orchestrated cross-functional AI product delivery: aligned diverse stakeholders across data science, engineering, security, compliance, and business teams; created technical architecture documents, API contracts, and SDKs enabling seamless integration; mentored 12+ engineers through code reviews and technical workshops
- Drove responsible AI practices: conducted model validation, bias audits, fairness assessments; ensured adherence to regulatory requirements (GDPR, CCPA, financial compliance) through regular audits, resulting in zero compliance violations over three years

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
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%.

Education

Bachelor of Engineering in Computer Science – Anna University (2005)