Woiny T Teka
512-590-1418
woinyt2014@gmail.com| Github
https://www.linkedin.com/in/woinyteka27/

Professional Summary:
Machine Learning and AI Engineer with expertise in designing, developing, and deploying enterprise-scale Generative AI, Agentic AI, and Retrieval-Augmented Generation (RAG) solutions. Experienced building AI agents and multi-agent workflows using Google Agent Development Kit (ADK), LangGraph, LangChain, CrewAI, and LlamaIndex. Strong hands-on experience with Vertex AI, Gemini models, model evaluation, prompt optimization, and LLM performance benchmarking. Skilled in LLMOps, MLOps, cloud deployment, vector databases, enterprise data platforms, and production-grade AI systems on Google Cloud Platform (GCP) and AWS.

Technical Skills
Programming Languages: Python, Java, JavaScript, Go, SQL
AI/ML & GenAI: LLMs, RAG Systems, LangChain, AI Agents, NLP, Generative AI, Prompt Engineering, Vector Embeddings, Fine-Tuning
Machine Learning Frameworks: PyTorch, TensorFlow, Scikit-learn, SageMaker
Backend Development: FastAPI, Spring Boot, REST APIs, Microservices Architecture
Cloud & DevOps: AWS (Bedrock, Lambda, SageMaker, EKS, S3, IAM, CloudWatch), GCP, Docker, Kubernetes, CI/CD, GitHub Actions, Jenkins
Databases & Search: MongoDB, Redis, Elasticsearch, OpenSearch, Pinecone, FAISS
MLOps & Infrastructure: Terraform, CloudFormation, Monitoring, Logging, Model Deployment, ML Pipelines
Security & Networking: JWT Authentication, Network Security, Risk Detection, Secure API Design
Software Engineering: Object-Oriented Programming, Distributed Systems, System Design, Testing (pytest, unittest)

Certification:
- IBM certification in Risk Management and Fraud Detection using Generative AI (Issued Apr 2026)
- Harvard University certification in Data Science: Building Machine Learning Models (Issued Jan 2026)
- AWS certification in Software Development on AWS (Issued Aug 2025)
- AWS Certified Cloud Practitioner (Issued Jun 2025)
- AWS certification in Generative AI with Diffusion Models (Issued Aug 2025)
- Google Cloud Foundations certification with hands-on experience in Google Cloud Platform (Issued Jul 2025)
- Google AI Essentials certification focused on Artificial Intelligence fundamentals and applications (Issued Jan 2025)
- Google certification in Automating Cybersecurity Tasks with Python (Issued May 2024)
- Google certification in Sound the Alarm: Detection and Response (Issued Apr 2024)
- Google certification in Connect and Protect: Networks and Network Security (Issued Jan 2024)
- Google certification in Play It Safe: Manage Security Risks (Issued Jan 2024)

Education:
- Maharishi Maharishi International University, MBA, Software Applications, Fairfield, IA
- Mekelle University, - Bachelor of Science, Computer Engineering, Mekelle, Ethio Licenses & Certifications

Professional Experiences
Uber AI Solutions San Francisco, CA **Oct 2025 – Present**
Machine Learning Engineer, AI/ML & GenAI
- Designed and deployed Agentic AI solutions using LangGraph, LangChain, CrewAI, and LlamaIndex for enterprise automation and intelligent workflow orchestration.
- Built production-grade RAG applications leveraging Vertex AI, Gemini Flash, vector databases, and semantic retrieval pipelines to improve response accuracy and business insights.
- Developed scalable AI microservices using Python, FastAPI, and Pydantic for cloud-native machine learning applications.
- Integrated BigQuery, Databricks, and enterprise data platforms to support large-scale AI workloads and advanced analytics.
- Implemented LLMOps frameworks including model evaluation, governance, observability, monitoring, and production deployment practices.
- Collaborated with business stakeholders to identify use cases, translate business requirements into AI solutions, and deliver measurable operational improvements.
- Containerized AI services using Docker and deployed workloads on Kubernetes and cloud environments across GCP.

Google, San Francisco, CA **Oct 2025 – Dec 2025**
Machine Learning Researcher

- Built AI agents and multi-agent workflows using LangGraph, LangChain, and Google's Agent Development Kit (ADK) to support complex reasoning, tool orchestration, and enterprise automation use cases.
- Developed AI-powered applications and research prototypes using Python and JavaScript, integrating Vertex AI, Gemini models, APIs, and enterprise data sources to support intelligent automation and knowledge retrieval.
- Built agent-based workflows, evaluation tools, and internal utilities using Python and JavaScript to streamline experimentation, model validation, and AI application development.
- Developed and optimized Generative AI applications using Vertex AI and Gemini models, including prompt engineering, retrieval augmentation, and agent-based architectures.
- Designed and implemented Retrieval-Augmented Generation (RAG) pipelines leveraging vector databases, semantic search, embeddings, and knowledge retrieval frameworks to improve factual accuracy and contextual relevance.
- Evaluated and benchmarked Large Language Models (LLMs) using quantitative and qualitative metrics, driving improvements in model performance, response quality, reliability, and production readiness.
- Applied LLMOps and MLOps best practices for experimentation, deployment, monitoring, observability, governance, and continuous optimization of enterprise AI systems on Google Cloud Platform.

HireArt New York, NYC **Jun 2024 – Oct 2025**
Senior AI Software Engineer

- Delivered AI/ML solutions using Python, FastAPI, AWS, and cloud-native microservices.
- Built scalable RESTful APIs and integrated AWS services including Lambda, S3, and API Gateway.
- Developed AI-driven automation systems and optimized ML pipelines for improved operational efficiency.
- Designed and developed scalable AI applications, RESTful APIs, and automation services using Python, JavaScript, FastAPI, and AWS cloud services.
- Built full-stack solutions leveraging Python backend services and JavaScript-based web integrations to deliver AI-driven business workflows and enterprise automation capabilities.
- Developed and maintained API integrations, data pipelines, and cloud-native microservices using Python and JavaScript to support production AI and machine learning systems.
- Implemented distributed data architectures using MongoDB, Redis, and Elasticsearch.
- Enhanced production deployment reliability using CI/CD pipelines, Git workflows, and automated testing frameworks.

Maharishi International University, Fairfield, Iowa **Jan 2023 – Aug 2023**
Research Assistant

- Conducted research on machine learning models and AI-driven applications for predictive analytics and intelligent automation.
- Assisted in data preprocessing, statistical analysis, and model experimentation using Python and machine learning frameworks.
- Collaborated with faculty and graduate researchers on software engineering and AI research projects.
- Built data pipelines and automated workflows to improve research efficiency and reproducibility.
- Contributed to documentation, technical reporting, and presentation of AI research findings.

ITEX Iowa USA/Worldwide, Full Time **Sep 2019 – Aug 2022**
Full Stack Developer

- Developed secure software solutions for scalable data-driven applications.
- Developed full-stack enterprise applications using JavaScript, Python, Java, Spring Boot, HTML, CSS, and REST APIs to deliver scalable and secure software solutions.
- Built interactive web applications and backend services using JavaScript and Python, improving system performance, user experience, and operational efficiency.
- Collaborated with cross-functional teams to design, develop, test, and deploy software features using modern JavaScript frameworks and Python-based backend technologies.
- Built RESTful APIs and dynamic front-end systems improving user experience and engagement.
- Collaborated with cross-functional teams to implement new features and troubleshoot production issues.
- Conducted code reviews, performance optimization, and software maintenance across distributed systems.
- Supported backend development using Java, Spring Boot, JavaScript, and cloud-native technologies.