Yashwant K

Machine Learning Engineer

✉ yashwantk0305@gmail.com   📞 +91 9342877060   📍 Coimabatore,India

○ github   in LinkedIn

Professional Summary

Aspiring Software Engineer with a strong foundation in **AI/ML and System Design**. Experienced in developing full-stack AI applications using **Streamlit, Python, and Vector DBs**. A consistent academic performer and active contributor to the tech community through event coordination and hackathon participation. Committed to driving innovation through impactful technical solutions.

Education

B.E. Computer Science Engineering,   2023 – 2027
Sri Shakthi Institute of Engineering and Technology

HSC (Computer Science - Maths) | 84.3%, *Amrita Vidyalayam Nallampalaym*   2022 – 2023

Core Technical Skills

Generative AI & LLM Frameworks
- **Agentic Systems:** Multi-Agent Systems, Google ADK (Agent Development Kit).
- **Orchestration & Retrieval:** LangChain, RAG (Retrieval-Augmented Generation).

Databases & Storage
- **Vector Database:** Pinecone.
- **Relational & NoSQL:** SQL, SQLite, MongoDB.

Machine Learning & Data Science
- **Libraries:** Scikit-learn, Pandas, NumPy.
- **Specialization:** YOLO (Computer Vision), NLP (Natural Language Processing).

Programming, Web & Cloud
- **Languages:** Python.
- **Web & Async:** AsyncIO, Web Crawling (specifically Crawl4AI).
- **Cloud Platform:** Google Cloud Platform (GCP).

Hackathon

SIH-2025, *Problem Statement AI-Driven Unified Data Platform for Oceanographic, Fisheries, and Molecular Biodiversity Insights* ⊡
Collaborated on the development of an AI-Driven Unified Data Platform to deliver actionable insights across oceanographic, fisheries, and molecular biodiversity. This platform strategically assists fishermen in identifying optimal harvesting zones, aiming to reduce reliance on trawler boats and mitigate damage to vital coral reef ecosystems.

Agents Intensive, *Capstone Project* ⊡
As a team, we successfully completed and developed a model to tackle issues arising from the updating or deprecation of package versions.

Vercel AI Gateway Hackathon Project
Developed "Among AI," a game-like interface based on the mechanics of Among Us, designed to help identify less performant AI models.

Google Tunix Hackathon, *Training LLMs using Google's JAX-native library*
Google Tunix Hackathon 2025, Problem Statement: Enhancing LLM Transparency and Reasoning Capabilities via JAX-Native Post-Training Collaborated on fine-tuning Google's open-weight Gemma models using the Tunix library to generate explicit reasoning traces. Engineered a reinforcement learning pipeline utilizing Group Relative Policy Optimization (GRPO) on Cloud TPUs to incentivize the model to logically deduce answers step-by-step, resulting in measurable improvements in interpretability and accuracy on complex reasoning benchmarks.

Certifications

HashGraph Developer Course - Hedera
(Nov 2024)

Computer Vision using Azure - Coursera
(Sep 2025)

Intro to Deep Learning - Coursera
(Aug 2024)

Intro to Machine Learning -
(Sep 2024)

Project Portfolio

AI Powered Package Conflict Support Agent,   09/2025 – 11/2025
Concise Intelligent Dependency Management & Self-Healing Code Infrastructure ⬀

- **Advanced Speech Processing Pipeline**: Designed and implemented a robust ASR (Automatic Speech Recognition) and diarization pipeline to accurately transcribe multi-speaker meeting audio and segment it by speaker. This formed the foundational structured speech data necessary for subsequent linguistic analysis.
- **LLM-Powered Semantic Analysis Engine**: Built a custom Large Language Model (LLM)-based semantic analysis engine specifically trained to identify subtle yet critical communication issues. The engine was configured to systematically detect and flag instances of technical jargon, ambiguity (e.g., vague pronouns, undefined terms), and general clarity issues (e.g., overly long sentences, complex syntax).
- **Granular Communication Scoring**: Developed and integrated sophisticated algorithms for topic segmentation to break down meetings into logical sections. Furthermore, a phrase-level scoring mechanism was implemented to provide targeted, quantitative feedback on the effectiveness and clarity of specific utterances.
- **Interactive Visual Clarity Dashboard**: Developed a user-friendly and intuitive interface using Streamlit to present the AI-generated clarity insights. This visual dashboard allowed users to quickly pinpoint problematic sections, review clarity scores, and understand the specific language contributing to potential miscommunication.
- **Quantifiable Efficiency Improvement**: The deployment of the Meeting AI Agent led to a substantial and measurable increase in post-meeting review efficiency, resulting in a 70% reduction in the manual workload previously required for reviewing meeting transcripts and assessing communication effectiveness.

Hydroflow, *Hydraulic Assistant is a comprehensive Flutter application that provides an AI-powered expert system for hydraulic hose pressure and systems.* ⊡

01/2025 – 05/2025

- **Cross-Platform Architecture:** Engineered a robust, scalable mobile interface using **Flutter (Dart)**, ensuring a seamless and responsive user experience across both Android and iOS platforms. Implemented advanced state management solutions (Provider/Bloc) to handle complex application states, ensuring data consistency and UI responsiveness without performance degradation.
- **Real-Time Data Visualization Engine:** Developed a dynamic dashboard capable of rendering high-frequency sensor data into intuitive visual formats. Utilized custom painting and charting libraries to create interactive graphs for water level monitoring, flow rates, and historical usage patterns, allowing users to interpret complex hydrological data at a glance.
- **Intelligent Leakage Detection Algorithms:** Integrated logic to process incoming telemetry data against historical usage baselines. The system featured automated anomaly detection that could identify irregular flow patterns indicative of pipe leakages or sensor malfunctions, triggering instant notifications to prevent resource wastage.
- **Cloud-Native Backend Integration:** Orchestrated a serverless backend architecture using **Firebase** for real-time database synchronization and user authentication. This setup ensured millisecond-latency updates between the hardware sensors and the mobile application, facilitating immediate remote control capabilities and reliable data storage for long-term consumption analysis.

Currency Detection App, *A multi-faceted mobile application focused on highly accurate Indian currency detection and accessibility features* ⊡

02/2024 – 05/2024

- **State-of-the-Art Object Detection**: Trained and optimized a YOLO (You Only Look Once)-based object detection model specifically on a large dataset of Indian currency notes under various lighting and angle conditions. This model provided real-time, bounding-box detection of the notes in the camera feed.
- **Enhanced Denomination Accuracy**: Integrated a separate, finely-tuned Convolutional Neural Network (CNN) classification model to analyze the detected note regions. This two-stage approach significantly improved the accuracy of denomination identification compared to a single-stage object detection system.
- **Visually Impaired Accessibility**: Crucially, the application incorporated a speech synthesis module (Text-to-Speech) to audibly announce the detected currency denomination and total amount, ensuring full accessibility and utility for visually impaired users.
- **Backend Inventory Management**: Expanded the application's utility by developing and integrating rudimentary merchant and pharmacy inventory systems. This feature allowed small business users to track cash transactions and automatically update stock levels based on currency counted.
- **Performance Metrics**: Through algorithmic optimization and model refinement, the application achieved a 30% improvement in currency counting speed while simultaneously exhibiting a significantly reduced error rate in both detection and classification tasks.

AI-Powered Autonomous Retail & Smart Checkout Solution, *Automated Store*   12/2024 – 02/2025
involves lot of Processing and automation and it usually cost high , Why can't we create something that integrates with our phone ? Well the Idea Came from there and We started ⬈

- **High-Performance Computer Vision Pipeline**: Engineered a robust real-time object detection system using YOLOv8 and PyTorch, specifically trained on a custom dataset of retail products (e.g., electronics, groceries). Achieved an impressive mean Average Precision (mAP@0.5 ≈ 90%), ensuring accurate product recognition even in varied lighting conditions.
- **Scalable REST API Architecture:** Developed a lightweight, production-ready **Flask API** to serve the detection model, creating a seamless bridge between the computer vision backend and mobile client applications. The system was designed for low-latency inference, enabling instant "scan-and-bill" functionality directly from a user's device.
- **Immersive AR & Voice Integration:** Architected a multi-modal user experience by integrating **Text-to-Speech (TTS)** synthesis for accessible product feedback and **AR-based indoor navigation** (via Unity/Flutter). This allowed users to not only auto-scan items but also receive real-time visual guidance to specific products within the store.
- **Edge-Optimized Mobile Deployment:** Optimized the deep learning model for mobile deployment, ensuring high-speed inference on mid-range devices without compromising accuracy. The solution effectively eliminates the need for expensive hardware sensors, replacing them with a cost-effective, smartphone-centric automated checkout workflow.

Professional Competencies

- Applies advanced Agentic AI engineering practices including multi-agent orchestration, RAG pipeline optimization, and automated context management

- Approaches complex system design with a focus on asynchronous efficiency (AsyncIO) and robust error handling in distributed agent workflows

- Familiar with deploying interactive AI applications using Streamlit and integrating Generative AI models via Google Vertex AI/Gemini APIs.

- Comfortable handling real-world data issues such as noise, imbalance, and incomplete datasets

- Comfortable processing unstructured data sources, managing PDF parsing pipelines, and automating web-based data extraction for real-time analysis.

Career Direction

- Aim to grow in roles related to Generative AI, Large Language Models (LLMs), and Autonomous Agent Systems

- Continue advancing skills toward senior-level AI Engineering and Full-Stack Machine Learning roles

- Contribute to teams developing scalable Multi-Agent architectures and high-performance software solutions

- Work on AI applications combining natural language processing, semantic retrieval (RAG), and intelligent process automation

- Build reliable, agentic workflows that deliver measurable real-world efficiency and problem-solving capabilities