

DHARMENDAR RAO

Technical Manager – AI, Computer Vision & Generative AI

12 Years of experience | AI Systems, Full stack & Cloud Architecture

Location: Jaipur, Rajasthan, India

+91 9413748809 | dharmendar.rao8@gmail.com | [LinkedIn](#)

Professional Summary

Technical Manager and hands-on engineer with 12 years of software development experience and 4+ years leading engineering teams.

Specialized in delivering end-to-end AI solutions with a strong focus on Computer Vision, Generative AI, and scalable production systems. Comfortable owning projects independently from requirements to deployment, integrating LLMs with CV workflows, building RAG pipelines, and shipping cloud-native systems with minimal oversight. Adept at translating business problems into reliable, measurable AI-driven solutions.

Core Skills: Python, Computer Vision (OpenCV, YOLO, PyTorch, TensorFlow), Generative AI & LLMs (OpenAI, Gemini, RAG, LangChain), FastAPI, Node.js, React, Next.js, TypeScript, MongoDB, PostgreSQL, MySQL, Milvus, Docker, CI/CD, AWS, Azure, GCP, REST APIs, System & Database Architecture, Technical Leadership

EDUCATION

RAJASTHAN INSTITUTE FOR ENGINEERING AND TECHNOLOGY | JAIPUR, RAJ

Master of Computer Application – M.C.A 2010 to 2013

NATIONAL COLLEGE FOR RESEARCH & TECHNOLOGY | JAIPUR, RAJ

Bachelor of Computer Application 2007 to 2010

SKILLS

- **Programming Languages & Frameworks:** Python (AI/ML, FastAPI), MERN Stack, Mean Stack, JavaScript, ReactJs, NodeJs, TypeScript, ExpressJs, NextJs
- **Computer Vision (CV):** Object Detection & Recognition (YOLO, Faster R-CNN), Image & Video Processing, Tracking, OCR, OpenCV, PyTorch, TensorFlow, Video analytics & real-time inference, Model optimization with GPU / edge deployment exposure
- **Generative AI & LLMs:** Large Language Model (LLM) integration (OpenAI, Gemini), Retrieval-Augmented Generation (RAG) pipelines using Milvus Vector DB, LangChain, Chat Bot Development
- **Databases:** MySQL, PostgreSQL, MongoDB, Milvus
- **Version Control & CI/CD (Continuous Integration/Continuous Deployment):** Git, Docker
- **Cloud Platforms:** AWS, Azure, GCP (Google Cloud Platform)
- **Design & Architecture:** Architecture Design, Database Design

EXPERIENCE

Technical Manager – AI & Engineering | Biz4Group (Returned) May 2025 – July 2026

- Own technical delivery of AI-driven products, balancing hands-on engineering with people and delivery management.
- Lead and mentor cross-functional teams of 8-10 engineers, ensuring accountability, technical quality, and timely execution.
- Drive end-to-end AI solution delivery including requirement analysis, architecture design, development, deployment, and monitoring.
- Provide architectural guidance for Computer Vision and Generative AI pipelines (data ingestion → inference → post-processing → serving).
- Oversee integration of Generative AI solutions (LLMs, RAG pipelines) with existing platforms.
- Establish coding standards, review best practices, and guide MLOps workflows (CI/CD, Docker, cloud deployments).
- Collaborate with product managers, stakeholders, and leadership to translate business needs into measurable technical outcomes.

Independent Software Consultant - Self employed August 2024 to May 2025

Provided independent software development and technical troubleshooting services, working directly with clients to scope requirements, set timelines, give technical direction, and deliver production-ready fixes and features with minimal supervision.

- Served as Technical Lead on a client engagement secured through Upwork, with scope expanding into ongoing development and server maintenance.
- Consulted for Cosco India Limited — resolved bug fixes, testing, feature troubleshooting, and code cleanup, and diagnosed a website issue affecting their Google Ads campaign, including performance optimization.
- Provided technical guidance, code reviews, and architecture input for client teams.

Biz4Group as Tech Lead June 2020 to June 2024

With over a decade of software development experience, including 3+ years as a technical lead, I managed project requirements, provided technical guidance to developers, and handled complex tasks to optimize team efficiency.

I specialized in both frontend and backend technologies, using React for web applications and Ionic for mobile development. On the backend, I primarily worked with Node.js and occasionally with Python. My responsibilities included high-level architecture and database design, custom chatbot development using OpenAI and Gemini APIs, and integrating MilvusDB with Retrieval-Augmented Generation (RAG) techniques. For deployment, I leveraged cloud platforms such as GCP, AWS, and Azure.

Earlier Roles – Web & Application Development (2014 – 2020) Built a strong foundation in backend development, APIs, scalable systems, and software engineering best practices, enabling progression into technical leadership and AI-focused roles.

Webware Info Solutions Pvt. Ltd as Senior Web and App Developer April 2018 to May 2020

All Cad Services Pvt. Ltd as Web Developer January 2017 to March 2018

Logicspice Consultancy Pvt. Ltd as Web Developer September 2015 to December 2016

Webgenesis as Web Developer July 2014 to August 2015

Major Projects

Project Name: Kalix QC – AI Cannabis Quality Grading (kalixqc.com)

Tech Stack: Python, OpenCV, PyTorch, YOLO, TensorFlow, XGBoost, FastAPI, Docker, CI/CD, AWS (EC2, S3, SQS, CloudFront, VPC)

Description: Role: Technical Manager & Technical Lead (Biz4Group LLC) — also contributed hands-on as a developer. An AI-powered mobile app that grades cannabis quality from a phone-camera video, scoring 7 measurable factors and returning a data-backed market price range in near real time, reaching approximately 80% grading accuracy in early evaluation. The product website is live (kalixqc.com); the app is in beta via TestFlight, with the Android release pending.

Built the Computer Vision pipeline using Python, OpenCV, YOLO, XGBoost, PyTorch, and TensorFlow for object detection, quality-factor scoring, and structured extraction from video frames, with real-time inference served via FastAPI.

Engineered for production with real-time inference, API-based model serving via FastAPI, and performance optimization for scalable deployment, producing standardized quality scores, factor-level breakdowns, and a data-backed price range from each scan. Designed and implemented the AWS infrastructure — Dockerized services on EC2 within a custom VPC, secured S3 storage, SQS-based asynchronous processing, CloudFront delivery, and CI/CD pipelines for automated builds and deployments.

Built end-to-end, covering CV model development, backend services, and cloud-native deployment, and currently moving from beta toward public release.

Project name: Valinor - Website and ChatBot

Tech Stack: ReactJs, NextJS, Python, FastAPI, AWS, OpenAI, Postgresql, Chat bot development, Stripe

Description: Role: Technical Lead (Biz4Group LLC) — also contributed hands-on as a developer. An AI-enabled platform that let people stay connected with relatives after they had passed away. Users saved stories, images, and answers into an AI database, and the chatbot generated responses on the owner's behalf. Owners could invite relatives to converse with the bot, and the platform supported subscription plans and transfer of chatbot ownership to relatives.

Project name: Lazify - ChatBot

Tech Stack: ReactJs, NextJS, Python, Django, AWS, OpenAI, Postgresql, Milvus, Dgraph, Stripe

Description: Role: Technical Lead (Biz4Group LLC) — also contributed hands-on as a developer. A multi-tenant chatbot-as-a-service platform where organizations could sign up, purchase plans, train bots, and add live agents. The bot handed off to human agents when needed, and agents could further train it mid-conversation to improve future responses. The chatbot was embeddable via a script tag and could be trained to perform custom actions such as processing refunds and returning order information, making it well suited for e-commerce. Tech stack included a Milvus vector DB and Dgraph for retrieval and knowledge handling.

Project name: Subsciety - Ecommerce (Live: subsciety.com)

Tech Stack: ReactJs, NodeJs, Azure, MySQL, Stripe, Zoom

Description: Role: Technical Lead (Biz4Group LLC) — also contributed hands-on as a developer. A marketplace e-commerce platform offering physical and digital products, services, and professional appointments. Users can book time slots with professionals (including Zoom-based sessions) and place orders seamlessly. Built with a React/Node.js stack on Azure, with Stripe payments and Zoom integration.

Project name: DR-HR - AI-enabled HRM Platform: Web, ChatBot & Mobile App (drhr.com)

Tech Stack: ReactJs, NextJS, NodeJs, Google Cloud Platform, MongoDB, Ionic

Description: Role: Technical Lead (Biz4Group LLC) — also contributed hands-on as a developer. A live AI-enabled HRM platform spanning web and mobile (published on the App Store). Organizations register and manage employees, roles, departments, and permissions, with AI-assisted suggestions for job postings and review ratings. Postings are auto-shared to job boards such as Indeed, and applications are collected and managed in-platform. Built with a React/Next.js and Node.js stack on Google Cloud, with MongoDB, and Ionic for the cross-platform mobile app.