

# VINIT KUMAR

AI Engineer & Full Stack Developer — Building RAG Systems, Fine-Tuned LLMs & Production Web Apps

Melbourne, Australia | 0417 872 419 | kumarvinit0510@gmail.com | github.com/VinitKumar0510-sudo

485 Visa — Full Work Rights (3 Years)

## SUMMARY

AI Engineer and full stack developer with a Master of IT (Artificial Intelligence) from La Trobe University, Melbourne (GPA: 6.0/7.0) and a portfolio of production AI systems including RAG pipelines, LLM fine-tuning with LoRA, local LLM automation pipelines, and computer vision applications. Delivered an industry capstone project for Albury City Council — building the rules engine, frontend, and test suite within a team of 5 — and currently leading IT operations and QA at Aaditya Rentals. Brings hands-on depth across the full AI development stack, from data ingestion and model training through to deployment and automation. Based in Melbourne with full work rights.

## EDUCATION

**Master of Information Technology** | La Trobe University · Melbourne 2024 – 2026  
Specialisation: Artificial Intelligence | GPA: 6.0/7.0

**Bachelor of Computer Science** | Lovely Professional University · Punjab, India 2019 – 2023  
Major: Data Science

## EXPERIENCE

**IT Support & Operations Lead** | Aaditya Rentals · Melbourne 2023 – Present  
Part-time during Masters; Full-time from December 2025

- Lead a team of 4 staff, overseeing daily task allocation and ensuring smooth front-line operations
- Conduct ongoing software testing, bug reporting, and QA of company website and internal systems — writing structured defect reports and coordinating fixes through to resolution
- Built internal automation tools integrating Xero and Zoho Mail APIs to streamline invoicing and debtor management workflows
- Manage business systems including customer database, fleet records, and administrative reporting

## PROJECTS

**Exempt Development Compliance App** | Albury City Council · Industry Capstone 2025  
React · Node.js · Express · Google Maps API · TailwindCSS · Jest · AWS

Final semester industry project delivered for a real government client — team of 5

- Designed and built the rules engine evaluating development proposals against NSW SEPP Part 2 legislation, returning structured legal decisions with clause references
- Developed the React frontend and Google Maps integration, connecting a 25,305-record property database with real-time address search and zoning auto-classification
- Led testing across the application, writing and executing the Jest unit and integration test suite to validate functionality against client acceptance criteria

### Vehicle DIY Guide — AI Mechanic

Python · TypeScript · RAG · LLM · Docker · Vector Store

- Built an AI-powered car repair guide platform ingesting repair manuals, YouTube transcripts, and Reddit community knowledge to generate step-by-step guides tailored to a specific vehicle
- Implemented an in-session AI mechanic using RAG architecture allowing users to ask real-time follow-up questions grounded in retrieved document context

### RAG PDF Chat — Local AI Document Q&A

Python · RAG · Vector Store · Local LLM · LangChain

- Built a fully local RAG application that allows users to chat with any PDF using a local AI model — no API keys required, runs entirely on-device

- Implemented vector store retrieval and embedding pipeline to answer natural language questions from document content with high accuracy

## LoRA Doctor — Medical LLM Fine-Tuning

[Python](#) · [LoRA](#) · [LLM Fine-Tuning](#) · [HuggingFace](#) · [Medical NLP](#)

- Fine-tuned a large language model on medical Q&A data using LoRA (Low-Rank Adaptation) — training a lightweight adapter on top of a base model without retraining full weights
- Achieved measurable improvement in domain-specific medical Q&A accuracy over the base model through parameter-efficient adapter training on a curated clinical dataset

## AI News Digest — Automated Daily Briefing

[Python](#) · [Local LLM](#) · [Automation](#) · [macOS Launch Agent](#) · [Email Pipeline](#)

- Built an automated daily news digest that summarises headlines using a local LLM and delivers them via scheduled email every morning through a macOS launch agent
- Designed end-to-end automation pipeline covering news ingestion, LLM summarisation, and scheduled delivery with zero manual intervention

## Mood-Based Music Recommender

[Python](#) · [OpenCV](#) · [TensorFlow](#) · [CNN](#) · [Spotify API](#)

- Built a real-time facial expression recognition pipeline using a TensorFlow CNN, achieving ~88% classification accuracy across 7 emotional states from live webcam input
- Integrated Spotify API to auto-play mood-matched playlists based on detected emotion in real time

*Earlier projects: Invoice Automation & Debtor Dashboard (Xero API, Zoho Mail, JavaScript) · Gym Log Android App (Java, SQLite, Android Studio) — available on [GitHub](#).*

## CERTIFICATIONS

---

### Supervised Machine Learning: Regression and Classification

[DeepLearning.AI](#) & [Stanford University](#) · [Coursera](#)

### Google Data Analytics Professional Certificate

[Google](#) · [Coursera](#) · 8 Courses

## TECHNICAL SKILLS

---

|                           |                                                                                                   |
|---------------------------|---------------------------------------------------------------------------------------------------|
| <b>Languages</b>          | Python, JavaScript, TypeScript, SQL, Java                                                         |
| <b>AI / ML</b>            | TensorFlow, RAG, LangChain, LoRA Fine-Tuning, HuggingFace, OpenCV, CNN, Local LLMs, Claude AI API |
| <b>Web &amp; Frontend</b> | React, Next.js, Node.js, Express, HTML5, CSS3, TailwindCSS                                        |
| <b>Mobile &amp; DB</b>    | Android Studio, SQLite, MySQL, Prisma, REST APIs, JSON                                            |
| <b>Cloud &amp; Tools</b>  | AWS (EC2, S3, Lambda), Docker, Git, GitHub, Jest, Playwright, Postman                             |