SHREESH DWIVEDI

Ayodhya, U.P.| dwivedishreesh8@gmail.com | +91 9415039944 | LinkedIn | GitHub

EDUCATION

VELLORE INSTITUTE OF TECHNOLOGY   **Bhopal, Madhya Pradesh**
Bachelor of Technology Computer Science (Specialization in Cloud Computing & Automation)   *October 2022 – Present*
CGPA: 7.52
Coursework: Data Structures & Algorithms, Machine Learning, Database Management, Cloud Computing, Operating Systems, Computer Networks

TECHNICAL SKILLS

- ❖ **Programming & Analysis:** Python (Pandas, NumPy), SQL, JavaScript/TypeScript
- ❖ **Data Science & ML:** OpenCV, Data Visualization, Predictive Modeling, Computer Vision, Statistical Analysis
- ❖ **Analytics & BI Tools:** Power BI, Tableau, MS Excel (Advanced Formulas, VLOOKUP/XLOOKUP, Pivot Tables, Macros)
- ❖ **AI & NLP Tools:** Generative AI, RAG Systems, Embeddings, Vector Databases, Hugging Face
- ❖ **Databases & Cloud:** PostgreSQL, MySQL, MongoDB, AWS, Google Cloud, Firebase
- ❖ **Web Development (MERN Stack):** React, Node.js, Express.js, MongoDB, REST APIs, HTML/CSS
- ❖ **Tools & Methodologies:** Git, Docker, CI/CD, Jupyter, Agile, FastAPI

WORK EXPERIENCE

Quant AI Limited, Auckland, New Zealand   May 2025 – August 2025
Remote Intern – Data Scientist
- ❖ Built a YOLOv8 + FastAPI + React video analysis system that automated detections and improved processing efficiency by 83%.
- ❖ Designed a RAG-based semantic search pipeline using embeddings and vector stores, improving retrieval accuracy by 60% and reducing lookup time.
- ❖ Spearheaded deployment of 3+ CI/CD pipelines, reducing release time by 70% while mentoring peers on automation workflows.
- ❖ Automated PDF data extraction workflows with Python, saving 3+ hours/week and increasing validation accuracy by 8%.
- ❖ Collaborated with global agile teams to integrate secure cloud infrastructure, enhancing backend reliability and scalability.
- ❖ Led feature demos and documentation efforts, aligning ML outputs with product requirements and improving cross-team communication.

ACADEMIC PROJECTS

CivicSpot – Social Awareness Web App | (MERN Stack, AI, Google Cloud)
- ❖ Developed a full-stack MERN platform that processed 200+ user-reported civic issues, visualized on an interactive Google Maps interface.
- ❖ Engineered low-latency Node.js APIs (<200ms), increasing user engagement by 25% and enabling real-time submissions.
- ❖ Integrated Google Vision API + custom ML classifier (58% accuracy) for automated content validation and categorization.

AI Powered Video Object & Activity Detection | (YOLOv8, FastAPI, React)
- ❖ Architected a video analytics engine capable of processing 50+ hours of content with 95% detection accuracy.
- ❖ Built a real-time error monitoring pipeline that autonomously resolved 95% of processing failures.
- ❖ Achieved 30 FPS real-time inference using optimized OpenCV pipelines and GPU-accelerated preprocessing.

Local Llama 3 8B Q&A System | (Llama 3, Python, Transformers)
- ❖ Devised a local Llama 3 8B inference setup for efficient offline Q&A and summarization using quantization techniques.
- ❖ Implemented a retrieval + generation workflow using Transformers to improve semantic response accuracy.
- ❖ Optimized prompts and decoding strategies, reducing hallucinations and improving answer consistency.
- ❖ Accelerated inference speed by 20% while achieving more consistent and reliable model outputs.
- ❖ Tested model quality using a structured prompt set, assessing accuracy, coherence, and hallucination rate.

CERTIFICATIONS

- ❖ Data Science Professional | Oracle (2025)
- ❖ Generative AI Professional | Oracle (2025)
- ❖ MERN Full-Stack Certification | Ethnus (2025)
- ❖ Solutions Architect Associate Training | AWS (2025)