KARAN DESAI

AI/ML Engineer (GenAI-Focus)

9970843464 @ karansd00@gmail.com Linkedin Porfolio Jaysingpur, Maharashtra

SUMMARY

With a robust background and successful internships in computer vision and AI/ML, a keen enthusiasm is demonstrated for contributing innovative AI solutions. Driven by passion for utilizing Python, TensorFlow, and AWS cloud infrastructure, the focus is on enhancing machine learning models and frameworks. Eagerness is shown to apply technical skills and project experience in building sophisticated ML models to advance organizational goals in the AI/ML domain.

EXPERIENCE

Resolute AI — AI Intern

June 2025 – Present

- Developed AI modules for traffic analysis using YOLOv8 and Supervision, including lane-wise vehicle detection, counting, and speed estimation.
- Built a face recognition pipeline using Face_recognition, capturing and encoding face data for video-based identification.
- Applied PaddleOCR with GPU to extract text from complex image data (license plates, voltmeters), enhancing OCR accuracy and speed.
- Engineered a video question-answering system using Whisper for transcription and Gemini API for GenAI-powered responses.

RacksonIT - Computer Vision Intern

12/2023 - 05/2024

Pune, Maharashtra

- Specialized in CNN-based image classification and object detection using Python, TensorFlow, and OpenCV.
- Improved model accuracy by 25% through advanced data augmentation techniques and optimized preprocessing workflows.
- Collaborated with the engineering team to deploy deep learning models via Flask-based REST APIs.
- Refactored the image processing pipeline, achieving a 30% increase in processing speed and better real-time performance.

EDUCATION

Bachelor of Technology — 02/2021 - 10/2024

Sharad Institute of Technology College of Engineering — CGPA 7.6 / 10

HSC

A.N.N Junior College — 01/2019 - 01/2020

CERTIFICARIONS

AWS Academy - Machine Learning Foundation:
Course covering foundational concepts in machine learning with AWS cloud.

AWS Academy - Cloud foundation:
Hands on experience in training and deploying autonomous model on AWS cloud.

Deep Learning using Ai-jetson Nano:
Learned deep learning application using ai-jetson nano.

PUBLICATIONS:

Phishing Link Detection Using Machine Learning

SKILLS

- **Programming Languages**: Python, C++
- **Machine Learning**: Scikit-learn, SciPy, NumPy, Pandas
- **Deep Learning**: TensorFlow, Keras, CNNs, Vision Transformers
- **Computer Vision**: YOLOv8, OpenCV, Supervision, ResNet, PaddleOCR, face_recognition,Vision Transformer
- **Natural Language Processing**: Hugging Face Transformers, DistilBERT, Spacy, NLTK
- **Generative AI**: Langchain,Phidata, Whisper,HuggingFace SentenceTransformers
- **Audio Processing**: Speech-to-Text (Whisper),
- **Web & Deployment**: Flask, Streamlit, FastAPI
- **Cloud & DevOps**: AWS EC2,Lambda
- **Database & Retrieval**: FAISS, Pinecone, ChromaDb,Mysql
- **Version Control** : Git, GitHub

PROJECTS

Document-Aware Chatbot using RAG and Gemini API

•Developed a Flask-based Retrieval-Augmented Generation (RAG) chatbot.

•Enabled document uploads (PDF, DOCX, TXT) for contextual Q&A.

•Utilized FAISS and SentenceTransformers for information retrieval, and generated responses using the Gemini API.

Built a question answering system using DistilBERT.

•Fine-tuned DistilBERT on the SQuAD dataset to build a QA system extracting precise answers from context.

•Used Hugging Face Transformers, achieving high F1 on validation.

Transformer Models for Image Classification on CIFAR-10

•Built a Vision Transformer (ViT) from scratch using TensorFlow/Keras to classify CIFAR-10 images, achieving competitive accuracy against CNNs.

•Implemented multi-head self-attention and patch embedding