

DEV KHANT

☎ +919054594757 ✉ devkhant46@gmail.com 🌐 devkhant 📱 Dev-Khant 📄 devkhant24

EDUCATION

Darshan University

B.Tech in Computer Science Engineering

July 2020 – May 2024

Rajkot, Gujarat, India

EXPERIENCE

Polymerize

Nov 2022 – Current

MLOps Engineer

Remote

- Integrated **RAG** for multiple languages and file formats(PDFs, CSVs, and excel). It's used for QA from documents, analyzing data, and creating user-specific plots. Also working around LLMs like GPT-4 to add better capabilities.
- Added new **Inverse Prediction** flow. This led to a **32% reduction** in the prediction time for material properties based on a set of ingredients.
- Developed **Explainable AI** for regression and classification models. Multiple charts are shown explaining the model's way of predicting.
- Reduced Machine Learning API response time by **20%**. Integrated **data sanity** flow before model training.

AiBorne | [link](#)

July 2022 – Oct 2022

Machine Learning Intern

Remote

- Implemented model on **Instance Segmentation and Object Detection** using Pytorch. The goal here was to detect the dents, scratches, scrapes etc. present on cars and estimate the repair cost.
- Integrated **MLflow** for model tracking and Data pipeline for Classification tasks. The classification model was trained to determine the view of the car with **30+ classes**.
- Developed a **custom Object Detection + OCR model** for reading Digital Odometer. The **mAP for the detection model was 96%**. Experimented on DINO which is an enhanced version of Detection Transformer(DeTR).

ResoluteAI | [link](#)

April 2022 – July 2022

Deep Learning Intern

Remote

- Contributed to ongoing projects using **YOLO** for Object Detection and **OpenCV** for tracking and counting objects. This was used in manufacturing industries for counting bottles, towels and boxes. And **mAP achieved here was 97.5%**. During the Inference test video ran at around **50 FPS**.
- Deployed **2 POCs using Streamlit and Flask**. One project was to detect defects in fabrics and calculate the area covered by them. Another project was extracting tables from bills or receipts using OCR. **Pytesseract** was used for OCR.
- Got hands-on experience of annotating real world industry data using **LabelMe** and **LabelImg** for the custom detection models.

PROJECTS

Analyze Github Code | [Github](#)

August 2023

- Developed a Chrome extension to get a detailed summary of a Github Repository using **LLM**.
- It uses **Langchain** to interact with **gpt-3.5-turbo** from OpenAI. **Flask** is used as a backend server to handle the requests.
- It traverses through each code file in the repository and stores the summary of it. At the end, the model summarizes all the stored summaries into one final response.

YouTube Video Explainer | [Github](#)

September 2023

- For a youtube video it provides an understanding of video's content through **audio transcription** and **image captioning**.
- Used **Huggingface** for OpenAI **Whisper model** for audio transcription and **swiftformer-xs** to get image embeddings, Google Lens API through **SerpAPI** for searching about image frames and **Chromadb** as vector database.
- UI is built using **Streamlit**. Video and Audio are processed **parallelly**.

TECHNICAL SKILLS

Languages: Python, C, C++, HTML/CSS

Libraries: Tensorflow, PyTorch, FastAI, Scikit-learn, Langchain, Pandas, Numpy, Matplotlib, NLTK, OpenCV

Databases: MongoDB, MySQL

Web Technologies: Flask, FastAPI

ACHIEVEMENTS

- I am a **Kaggle Expert** in Competitions, Notebooks, Datasets and Discussions category. [link](#)
- Achieved rank of **208 (top 8%) out of 2800+** participants in Ubiquant Market Prediction competition on Kaggle. I was awarded a **Bronze medal** for this competition. [link](#)
- I ranked **280th (top 13%) from more than 2300+** competitors in Jigsaw Rate Severity of Toxic Comments Competition on Kaggle. [link](#)