Guruprasad Dahiphale

+91 7249382377   Portfolio   guruprasaddahiphale7@gmail.com   Linkedin   Github   Pune

Work Experience

BTD Service India

Pune
March 25 - August 25

Data Scientist
- Built a custom OCR model to extract text from scanned PDFs and image-based documents.
- Implemented a robust post-processing pipeline to clean, validate, and structure extracted text.
- Worked on a real-world Hotel Menu Card Analysis project:
    1. Extracted menu items and corresponding prices using the OCR pipeline.
    2. Structured and stored processed data in MongoDB for analytics and retrieval.

Projects

1 - SentimentGuru AI – YouTube Comment Sentiment Analysis

Developed a web app to analyze YouTube comments and classify them as Positive, Neutral, or Negative using RoBERTa-based NLP. The system extracts data via YouTube API, performs sentiment analysis with Hugging Face Transformers, and displays results on an interactive website with video preview, sentiment distribution, ratings, and categorized comments.

Key Skills : Python, NLP, Sentiment Analysis, Hugging Face Transformers, Flask, REST API Integration

Project Link : https://github.com/Guruvd07/SentimentGuru-AI.git

2 - GuruInsight AI – YouTube Video Keyword Insight Extractor

Developed an AI system that analyzes YouTube videos by extracting transcripts, detecting keyword-specific segments, and generating focused summaries of what the speaker said about the chosen topic/person. Built a robust 4-level transcript pipeline (VTT → TimedText → Transcript API → Whisper ASR), applied context-aware keyword extraction, and used FLAN-T5 for transformer-based summarization. Delivered through a lightweight Flask web app for real-time video insight generation.

Key Skills : Python, NLP, Hugging Face Transformers, Text Summarization, Whisper ASR, REST API Integration.

Project Link : https://github.com/Guruvd07/GuruInsight-AI.git

3 - CarGuru AI – Used Car Price Prediction Web App

Built an end-to-end project involving data engineering, data cleaning, data analysis, ML engineering, and deployment. Scraped real-world car listings using Requests and BeautifulSoup , transformed and analyzed data with Pandas, trained a RandomForest Regression model and deployed a Flask web application with for real-time predictions and interactive search.

Project Link : https://github.com/Guruvd07/car-price-predictor.git

TechnicalSkills

Strongest Areas
- **Data Engineering :** (ETL Pipeline) Data Extraction , Data Transformation , Data Loading
- **Data Analysis :** Exploratory Data Analysis , Data Visualization , Statistical Analysis
- **ML Engineering :** Model Building , Model Evaluation , Feature Engineering , Model Deployment
- **Deep Learning :** ANN, CNN, Transfer Learning (MobileNetV2), Transformers (BERT, T5, GPT), NLP (Sentiment, Summarization, Translation)
- **OCR :** Tesseract OCR for Image , PDF Text Extraction

Languages — Python , SQL

Database — MySQL

Developer Tools — GitHub, VS Code, PowerBI , Jupyter Notebook , Google Colab , MySQL Workbench , Docker

Education

MIT Academy of Engineering

Pune , Maharastra
Dec 2022 - Aug 2026

BTECH - Computer Engineering

Achievements

1. **Grand Finalist of Smart India Hackathon 2024**
2. **Winner**, **MIT E-SUMMIT Ideathon** 2024 , Secured **first prize** for innovative project proposal, showcasing problem-solving skills.
3. **Finalist** of **BMC** Company Hackathon 2025 BMC Software , VimanNagar - Pune
4. **Organizer** at **MAHATECH** Exhibition 2025 ShivajiNagar - Pune February - 2025
5. **Microsoft , Edunet foundation , TechSaksham** Regional Showcase Winner 2025
6. MITAOE **Datathon** 2025 **Runner-Up**
7. **Judge** – **PromptWars** 2025, ASSCET MIT AOE