Arunim Malviya

+91 7440460090 arunim.malviya.13@gmail.com Arunim Malviya Arunim313

Education

Master of Technology(Dual Degree) in Data Science and Artificial Intelligence **2020 - 2025**
School of Data Science and Forecasting, DAVV *Indore, India*

Class XII (84.60 %) **2020**

Professional Experience

Machine Learning Engineer — Internship **Feb 2024 – Present**
NYX *Bangalore Urban, Remote*

- **Developed and deployed a production-level Text-to-Image model** using the Flux API, prompting the model to generate images, analyzing visual features impacting CTR, and iterating on creatives based on recommendations to enhance engagement and performance.
- Built recommendation models to optimize **Cost Per Click (CPC)** and **Click-Through Rate (CTR)**, resulting in a **16% reduction in CPC** and a **12% increase in CTR**.
- Engineered an ETL pipeline to extract, transform, and load data from **Google Ads**, **Meta Ads**, and **LinkedIn Ads** into a **PostgreSQL** database, which helped in improving accuracy for data analysis and reporting.
- Conducted **QA testing** on over **6 systems** and guided **2 PowerBI** and **3+ Metabase dashboards** to track key performance indicators (KPIs) and support data-driven decision-making.
- Migrated **6TB** of data from **Amazon S3** to **Google Cloud Storage (GCS)**, ensuring **100% data integrity** and improving data accessibility for scalable project development.
- **Optimized** Python scripts through multithreading, achieving a **30% reduction in response time**. Implemented web scraping scripts that extracted over **30 million** images and videos, significantly enhancing data collection and processing workflows.

Technical Skills

Programming and Scripting: Python, C++, R, SQL, Linux, Selenium, Finetuning LLMs

Data Science: Machine Learning, Deep Learning, NLP, CNN, Transformers, Tensorflow

Data Visualization: PowerBI, Tableau, Matplotlib, Advanced Excel

Databases: MySQL, PostgreSQL, ChromaDB, MongoDB

Cloud Platforms: Basic understanding of GCP, Azure

Code Management: Git, GitHub, Docker

Core Subjects: Object-Oriented Programming, Data Structures

Projects

AutoSQL: The Intelligent SQL Query Agent

- Built AI agent to transform natural language queries into SQL/CSV operations with seamless query execution.
- Integrated **LLaMa 3 (70B)**, **ChromaDB**, and **PostgreSQL** for SQL operations, ensuring high-performance information retrieval and query handling..
- Built a scalable backend using Python and Flask, offering a RESTful API for easy integration with existing systems.

LLM App for PDF Insights Extraction | Link

- Utilized **Ollama framework to integrate Mistral and Zephyr** language models for PDF-to-knowledge graph conversion.
- Implemented **FastAPI for the backend** and **Gradio UI for the user-friendly frontend**, deployed application on Vercel.
- Enabled seamless conversion of input PDFs into knowledge graphs, facilitating hidden insight discovery.

Image Profiling Python Library | Link

- a Python library for in-depth image analysis, including texture analysis, fractal dimension, and sharpness measurement, enhancing image dataset profiling for data scientists.
- Implemented advanced image processing techniques such as Local Binary Pattern Variance (LBPV) and object detection models, enabling comprehensive analysis and object recognition in image datasets.

Achievements

- Attained a **5-star rating** in SQL, Python, and C++ on Hackerrank.
- Served as the team lead during the eSummit hackathon at MANIT, securing a **top 10 final position**.

Certifications

- DeepLearning.AI: Machine Learning Specialization
- Databases and SQL for Data Science with Python
- Microsoft Certified: Azure AI Fundamentals AI-900
- SQL Programming
- Advance Usage of Excel