

# Udit Sharma

uditsharma0529@gmail.com • [LinkedIn](#) • [GitHub](#) • (585) 831 4497

## PROFESSIONAL EXPERIENCE

### Machine Learning Engineer

Feb 2022- Present

*Viant Technology, Irvine, CA*

- Developed and deployed a deep learning model using **Tensorflow** that improved **Cost Per Action (CPA)** by **20%**, serving over **10 million impressions** daily. Model trained on **Vertex AI**.
- Proposed, designed, and implemented a **recommendation system** to identify high-value users in **PyTorch**, applying **transformer**-based attention on time-snapshot historical user graphs, achieving a **2x improvement in CPA**.
- Leveraged **OpenAI LLM** models (*4o-search & text-embedding-3*) API to summarize advertiser URLs and generate embeddings for effective targeting.
- Researched and engineered innovative features using **Large Language Models (LLM)** enhanced by **Retrieval-Augmented Generation (RAG)**, which were successfully adopted across **multiple models**.
- Contributed to production model deployment and integration using **GoLang**.
- Built robust **ETL pipelines** using **Python+SQL** for dataset generation, offline evaluation, and Looker Studio model monitoring to support experimentation and analysis.

### Graduate Research Assistant

Sep 2019 – Dec 2021

*Rochester Institute of Technology, Rochester, NY*

- Bosch**: Designed a lightweight, multi-stage neural network pipeline for edge devices using **PyTorch** to detect glassbreak events with an exceptionally low false positive rate of **0.0003**.
- Teaching Assistant: Enhanced Deep Learning/Math courses & tutored 35+ students/class, boosting learning outcomes.

### Software Engineer

May 2017 - June 2019

*Accenture Solutions Pvt. Ltd., India*

- Worked as a client-facing SAP MM and ABAP consultant specializing in the Accounts Payable process.
- Automated** creation of manual payment forms based on past creation history, reducing new incidents by 75%.
- Reduced** manual effort by **40%** in applying **Machine Learning** and automation for GR/IR reconciliation in SAP.

## SKILLS

**Programming Languages:** Python, SQL, R, GoLang

**LLMs & Generative AI:** Open AI API, Gemma, Gemini, LangGraph, LangChain, RAG, Prompt Engineering, Agentic AI, Vector DB

**Machine Learning Tools:** PyTorch, Tensorflow, Keras, scikit-learn

**Enterprise:** Git, Docker, Flask, Streamlit, SAP

**Google Cloud Platform:** BigQuery, GCS, Vertex AI, Cloud Functions, Cloud Scheduler, Looker Studio, Colab, Kubernetes

## AI PROJECTS

### PDF insights extractor with multi-modal RAG

- Built a standalone RAG application using **Gemma-3**, **Langchain**, and **Streamlit** for chat-based, context-aware information retrieval from PDFs (text, tables, images), showcasing unstructured data processing and user-facing AI tool development.

### SQLite to Google Slides AI Agent

- Developed an AI agent with **Gemini-2.0-Flash** to autonomously analyze **SQLite** data, synthesize insights, and generate Google Slides presentations via API, demonstrating agentic AI capabilities for automation.

## PUBLICATIONS

### GourmetNet: Food Segmentation Using Multi-Scale Waterfall Features and Dual Attention [\[Paper\]](#) [\[Github\]](#)

- Developed a novel deep learning architecture in **PyTorch** for semantic segmentation of food images, advancing the **state of the art** and enabling more accurate food calorie and volume estimation.
- Deployed the model via a **Flask API** for real-time inference and usability in research and health applications.
- The paper has received **21+** citations to date.

## EDUCATION

**Master of Science, Computer Engineering**, Rochester Institute of Technology, Rochester, NY

**Dec 2021**

**Bachelor of Engineering, Information Technology**, Devi Ahilya University, India

**May 2017**