Akarsh Upadhyay

uakarsh | Akarsh Upadhyay | akarshupadhyayabc@gmail.com

Education

2019 - 2023 B.Tech In Electrical Engineering at **IIT Jodhpur** (GPA: 8.43/10.0)

Work Experience

Applied Scientist, Microsoft September 2024 - Present

- **Intent-Based Retrieval for Ad Selection**:
  - Led development of **next-generation encoder models** for ad retrieval based on user intent, establishing **foundational data quality practices** adopted across **MSAN (Microsoft Audience Network) teams globally**.
  - **Data Quality Innovation**: Identified **systematic biases** in quality model scoring and positive/negative pair definitions. Designed improved curation methodology addressing label noise, creating high-quality dataset foundation for encoder and generative retrieval models.
  - **Modeling Impact**: Achieved **+14.32 absolute precision points (P@100)** over production baseline and **+1.03 points over second-best model** using **30× less training data**, demonstrating multiplicative impact of principled data quality on model performance.
  - **Production Results**: A/B testing in NA region showed **+0.61% CTR and +0.51% revenue lift**. Conducted **metric gap analysis** providing actionable roadmap for next-generation data strategy across MSAN.
  - **Cross-Team Adoption**: Data quality methodology and error analysis patterns **adopted as standard practice** by **US and India MSAN teams**.
- **Unified Evaluation Framework for MSAN Selection Models**:
  - Designed **unified evaluation framework** serving as **single source of truth** across **IDC, STCA, and US MSAN teams**, supporting encoder-based, Single/Multi-Intent, Generative Retrieval, and ANN variants.
  - Introduced **GPT-based evaluation** to reduce bias from quality models. Implemented advanced metrics with **multi-dimensional slicing** for improved diagnostic depth.
  - Optimized for compute constraints through **caching and operation fusion**, enabling **multi-team concurrent usage**. Framework outputs query-document-score files enabling error pattern analysis and investigation of precision-distance relationships.
  - Enabled **data-driven A/B testing** by surfacing query-ad distance vs relevance relationships, replacing heuristic-based distance threshold tuning with empirical analysis.
  - Established as **required standard** for offline studies and A/B testing. Drove adoption across regions through **user-centered design** and iterative validation with model owners.
- **High-Impact Index for Product Ads**:
  - Designed **ML-driven retrieval system** identifying high-impact product offers based on historical engagement patterns, optimizing ad monetization across MSAN.
  - Led **end-to-end development** from defining impact criteria through production deployment and replaced legacy system, delivering **2.77% revenue lift**

Machine Learning Engineer, Zomato January 2024 - August 2024

- **Image Quality Score**:
  - **Project Overview**: Led the development of a **Image Quality Score** system that evaluates and classifies the quality of food images uploaded by restaurants, determining whether an image should be **accepted or rejected** based on its suitability for user display.
  - **Impact**: This model is now a critical component used in evaluating nearly every food image displayed to users, from **homepage banners** to **catalog images**, significantly improving the visual quality and consistency of content shown to customers.
  - **Data Expertise**: Curated a diverse, high-quality **dataset** of Indian food images, using comprehensive **data preparation guidelines**. Supervised the **annotation team**, established a **data quality pipeline** to ensure accuracy, and also utilized **GPT-4** for data preparation.
  - **Modeling Success**: Fine-tuned a **CNN-based ResNet-50 model** to classify image quality, achieving an impressive **F1 Score of 90%** for distinguishing between 'good' and 'bad' quality images.

- **Ads Creation**:
  - **Project Overview**: Led the development of an automated **ad creation system** designed to generate high-quality marketing ads tailored to specific brand styles, food items, and promotional content.
  - **Innovative Approach**: Coordinated a multi-step process involving **generative models** to create visually compelling backgrounds, optimal placement of visual elements, and seamless integration of brand-specific fonts and styles. The process included leveraging AI techniques to ensure consistency with brand identity while maximizing visual appeal.
  - **Proof of Concept**: Successfully demoed the proof of concept, demonstrating the system's ability to autonomously generate ads that meet high-quality standards and adhere to brand guidelines, setting the stage for further development and potential deployment.

- **Photo Cake**:
  - **Real-Time Image Overlay**: Developed the **real-time image overlay** system using **heuristic methods** and **OpenCV functions**, allowing user-uploaded images to be seamlessly placed onto cake templates.
  - **Photo Cake Ops Pipeline**: Devised and implemented a robust **photo cake operations pipeline**, enabling merchants to upload and validate images for quality and correctness. This validation process ensures only suitable images are used for photo cakes.
  - **Impact**: Successfully launched the feature on **Mother's Day**, resulting in the sale of over **3000 photo cakes** across India, significantly enhancing user engagement and driving sales.

ML Research Intern, Enterpret June 2023 - December 2023

- **Primary Responsibilities and Achievements**:
  - Led a project to assess **text similarity** between sentences based on their potential **business value**. Conducted extensive research into **text similarity methods**, devised **experimental setups**, and resolved **40% of the challenges**. Scaled the solution to handle over **a million texts**.
  - Prioritized **data preparation** in later project phases, creating comprehensive **guidelines spanning over 20 pages** to streamline data processing. Supervised the **annotation team** and set up a **data quality pipeline** to ensure accurate data preparation.
  - Played a pivotal role in devising a sophisticated **modeling approach**, achieving an impressive **F1 Score of 85%** for identifying similarities within distinctly related texts.
- **Skills Used**:
  - Explored diverse **modeling strategies**, including **prompt engineering**, fine-tuning **transformers**, working on **novel loss function**, and leveraging **Large Language Models (LLMs)** for optimal results.

Publications

Trivedi, Akkshita et al. (2024). "GDP: Generic Document Pretraining to Improve Document Understanding". In: *18th International Conference on Document Analysis and Recognition (ICDAR 2024)*.

Projects

Document Layout Understanding — GitHub Link July 2021 – May 2022

- Worked under the guidance of Dr. Santanu Chaudhary on the topic of **"Document Understanding"**
- The work was under the topic Transformers for Document Layout Understanding, and it included the usage of **DocFormer** for Document Understanding in English as well as other multilingual languages.
- About DocFormer: DocFormer is a **multi-modal transformer-based architecture** for the task of **Visual Document Understanding (VDU)**

Medical Visual Question Answering — GitHub Link Feb 2021 – May 2021

- Resolved a problem, which was: Given an image of an **X-ray scan or MRI Scan (of the brain, kidney, lungs)**, and a set of questions, provide the answer to the question, taking the attention from the input image
- **Approach**: Applied the concepts of Natural Language Processing and Computer Vision, and utilized the Tensorflow framework as a tool to solve the problem
- **Results**: The model was able to answer satisfactorily on the training data set, and the model was able to predict about 90 % of the time correctly on the test data set

Position of Responsibility

- **Core Member of Robotics Club**: Served as a core member of the Robotics Club, which involves solving problems related to the domains of artificial intelligence and many more fields, and participating in multiple competitions and hackathons