

Umair Yousaf

umair.yousaf@colorado.edu | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of Colorado Boulder

MS Computer Science

Starting Fall 2025

Boulder, CO

Lahore University of Management Sciences (LUMS)

BS Computer Science – CGPA 3.6

2019 – 2023

Lahore, Pakistan

TEACHING EXPERIENCE

Teaching Assistant – LUMS

TA for the course Data Structures (CS 202)

Jan 2023 – May 2023

Lahore, Pakistan

- Managed a class of 100+ students for a C++ based Data Structures course.
- Created homework assignments on hashing, heaps, sorting, and algorithm selection
- Developed quizzes and programming assignments on tree data structures.
- Graded assignments and quizzes. Ran MOSS plagiarism checks, and invigilated assessments.
- Held weekly tutorials and office hours to assist with C++ coding, course concepts, and debugging.

WORK EXPERIENCE

Reteta, a Visionet Subsidiary

Junior Consultant - AI Team

Nov 2024 – Jul 2025

Cranbury, NJ · Remote

- Designed and implemented methods to generate accurate and structured SOAP notes from encrypted patient-doctor interactions via **LLMs**.
- Achieved model performance **competitive with state-of-the-art (SOTA) systems** in the medical NLP domain, enhancing the product's standing in a competitive healthcare AI market.
- Deployed and maintained the complete ML backend on Azure using **Azure Machine Learning and Azure Container Instances**, enabling scalable and secure model inference.
- Developed **NER-based techniques** to anonymize Personally Identifiable Information (PII) in patient-doctor conversations, ensuring compliance with HIPAA.
- Built and managed **CI/CD pipelines on Azure DevOps** for automated testing, deployment, and monitoring of models.

Visionet-Systems Limited

Junior Consultant - AI Department

Jul 2024 – Jul 2025

Lahore · On Site

- Developed a comprehensive **PDF parsing pipeline** capable of handling regular text, headings, subheadings, paragraphs, tables, images, and formulas.
- Integrated the pipeline with existing **RAG** architectures, resulting in significant enhancements in information retrieval and document querying.
- Developed a clothing recommendation system utilizing **transformer models** and re-ranking techniques.
- Conducted **multiple live client demonstrations of our solutions**.

Visionet-Systems Limited

Associate Consultant - AI Department

Jul 2023 – Jul 2024

Lahore · On Site

- Created a suite of **Generative AI** solutions for the insurance industry, streamlining operations at every phase of the insurance process. Developed the frontend in **React**, backend in **Flask**, and all **ML & AI pipelines**.
- Used **multimodal LLMs** for **VQA** and image-based information extraction, surpassing in-house OCR methods.
- Developed multiple end-to-end **RAG** pipelines using LangChain and LangGraph.
- Developed a reverse image search module for clothing articles using **Siamese Networks** and contrastive loss.
- Conducted research on SOA methods, hardware/software debugging, and created local LLM inference pipelines.

- Multi-class vehicle detection, number plate detection, person detection and facial recognition using **YOLOv5**.
- Dataset creation, validation, experimentation and model deployment.
- Created a dashboard using **React.js** and **Material UI** for real-time monitoring.

PROJECTS

LEGAL-UQA | *Ongoing Publication*

- Developed the first Urdu legal question-answering dataset with 619 question-answer pairs derived from Pakistan's constitution and corresponding legal article contexts.
- Implemented OCR extraction, manual refinement, and **GPT-4-assisted** translation to create a **parallel Urdu-English dataset** for domain-specific NLP.
- Evaluated state-of-the-art models, achieving **99.19% human-evaluated accuracy** with Claude-3.5-Sonnet and fine-tuned **mt5-large-UQA-1.0** for adaptation to legal domains.

Early Warning Detection System for Forest Fires | *Final Year Project*

- Employed **Siamese Networks** to classify images. Achieved best accuracy of **83%** and F1 score of **0.827**
- Generated synthetic dataset using video games and **Unreal Engine 4** for use in training other models
- Experimented using a variety of state of the art models such as **Deformable DETR**, **YOLO** etc.

Sentiment Classification using BERT on GoEmotions | *NLP Project*

- Multi-label sentiment classification on GoEmotions dataset using HuggingFace framework

Lane Change Warning System | *Computer Vision/Machine Learning Project*

- Developed a pipeline to read a dashcam video's frames, detect vehicles using **YOLOv7** and lanes using **LaneNet**
- Employed **auto dynamic homography** to generate a top-down view and detect any vehicles changing lanes

Language Detection on Audio Recordings | *Machine Learning Project*

- Used **KNN**, **SVM** and **NN** models on audio to detect if the language was English, Urdu or a mix of both.

PetNet | *Android Application*

- Developed an appointment system to connect Pet Owners and Vet Clinics
- Used the **MERN Stack** (with React Native), and deployed the backend on **Heroku**

DI-Log | *Web Application*

- Developed an online social media forum where users could create, comment on and like posts, customize their profiles and follow other users
- Used **Express.js** and **Node.js** for the backend, **MySQL** as the database and **EJS** for the frontend

TOOLS & LANGUAGES

Languages: Python, JavaScript, C++, C, Solidity, SQL, Haskell

Tools: LangChain, LangGraph, PyTorch, TF, HuggingFace, Scikit-Learn, OpenCV, Pandas, NumPy, PyTest

Software Engineering Frameworks: MERN Stack, MySQL, Flask, Django, Azure Suite, Docker

COURSES TAKEN AT LUMS

Deep Learning | Machine Learning | Computer Vision | Natural Language Processing | Software Engineering | Data Mining | Data Science | Statistics and Data Analysis | Applied Probability | Human Computer Interaction

HONORS & AWARDS

Dean's Honor List in Freshman, Sophomore and Junior year on maintaining a CGPA above 3.6
100% Merit scholarship in A Levels (both years)