ANUM SALEEM

DATA SCIENTIST | Data ENGINEER | ML ENGINEER

Haveli Lakha - Okara - Pakistan • +92 311 6580619 • anum.saleem2004@gmail.com

SUMMARY

Passionate and highly motivated individual with a keen interest in coding and a strong drive for continuous learning. Currently focused on acquiring skills and knowledge in the fields of data science, machine learning, and deep learning. Possesses a solid foundation in programming and a commitment to applying technical expertise to solve real-world problems

SKILLS

TECHNICAL SKILLS : C++ | Python | Numpy | Pandas | Matplotlib | Image Augmentation | SQL - MYSQL | Django | OpenCV | Selenium | HTML | CSS | Javascript | Bootstrap | Supervised Learning || Data scientist | C# GUI | OOP | Data Structure |

Tools: VS Code | Pycharm | Microsoft Visual Studio | Figma | Canva| Git | Word | PowerPoint | Excel | Linux | Jupyter notebook | Mendeley |

EDUCATION

Bachelor of software Engineering 2021 - 2025

Bachelor in Software Engineer / Current University of Management and Technology (UMT) - Lahore - Pakistan
Currently enrolled in 7th semester with CGPA 3.69

Intermediate F.SC (Pre - Medical) 2019- 2021

ILM Group of college - Haveli Lakha - Pakistan

CERTIFICATIONS

- Supervised Machine learning **- DeepLearning.ai**
- Introduction to Data Science **- IBM**
- Python for everybody **- University of Michigan**
- Tools of Data Science **- IBM**

PROJECTS

Hotel Management GUI Project
Established a robust connection between the database and C# application.
Designed and implemented user-friendly forms for Login, Signup, Customer, and Staff management.

Image Augmentation Dataset

Created a dataset with 1000 images, each containing three different breeds of dogs.

Each breed was represented by only three original pictures, emphasizing the use of augmentation techniques to expand the dataset.

Online Attendance System with Face Recognition

Created a student attendance system using OpenCV for face recognition.

Trained a model on a dataset of student faces, enabling automatic attendance system

Automated Selenium Script for Picture Downloads

Developed a Selenium-based automation script to download pictures from the Unsplash website.

Incorporated comprehensive Selenium concepts, such as web scraping, handling dynamic content, and downloading images.

Election Prediction - 2024

Scraped election-related data from multiple websites and performed Exploratory Data Analysis (EDA) to understand patterns and trends.

Prepared a robust dataset, trained a machine learning model, and made predictions about election outcomes.

Dynamic website with Django Backend

Developed a dynamic and fully responsive website using Django backend, integrating HTML, CSS, JavaScript, and Bootstrap.

Added functionalities for email configuration and password reset authentication to enhance user experience and security.

Netflix Clone Project

Developed a Netflix clone front-end project using HTML, CSS, avaScript, and Bootstrap.

Ensured the project is fully responsive, providing an optimal viewing experience across different devices.

Included panel navigation for easy and intuitive user interaction.

Image Classification Project

Developed and implemented a binary classification model using a pre-trained VGG16 architecture in TensorFlow. Optimized the model for high accuracy in classifying images, employing callback functions and visualizing training results with performance graphs.

Image Steganography with django website

Developed a Django-based web application for image steganography, enabling users to encode and decode hidden text in images using PIL and Stegano.

Implemented features like image upload, text encoding/decoding, and data management with SQLite, with a responsive UI using Bootstrap

Video-to-Text Conversion Using IBM Watson API

Implemented a video-to-text conversion application using IBM Watson's Speech to Text API, enabling automated transcription of video files.

Optimized the workflow for accurate and efficient extraction of text from audio content.

Sentiment Analysis for IMDb Movie Reviews

Built a sentiment analysis tool that scrapes and analyzes IMDb movie reviews using selenium, PyTorch and transformers, classifying reviews into 'Happy,' 'Neutral,' and 'Bad' categories.

This automated system processes large datasets efficiently to extract valuable insights from user feedback.

Retrieval-Augmented Generation (RAG) System

Designed and implemented a Django-based RAG system enabling users to upload documents and query them effectively.

Integrated the Llama 3 model via the ollama library to process documents and generate dynamic, context-aware responses.

Optimized the system for accurate information retrieval and seamless user experience.

Educational AI Platform using RAG system

Designed a Django-based AI-driven educational platform capable of delivering dynamic, context-aware responses in formats such as paragraphs, lists, and tables, performing OCR, and downloadable response in the form of doc.

Leveraged GROQ for adaptive response generation, ensuring flexibility without modifying the system architecture.

Implemented session-based chat history to enhance user engagement and maintain conversational context.

DataCamp - Data Science Track Projects

- **Investigating Netflix Movies**
  By using simple python techniques
- **Exploring NYC Public School Test Result Scores**
  Finding Top School by using pandas Techniques