Awais Tahseen

Rawalpindi, Pakistan | awaistahseenaccoun@gmail.com | +923315108736 | awaistahseen.com

linkedin.com/in/awais-tahseen-57359a267/ | github.com/awaistahseen009

About Me

A results-driven person, skilled in data science, machine learning, and end-to-end ML solutions. Proficient in Python, ML algorithms, and Deep Learning, with strong analytical and problem-solving skills. Passionate about Computer Vision , Prompt Engineering, LLMs and Agentic Solutions

Education

Pakistan Institute of Engineering and Applied Sciences (PIEAS), BS in Computer and Information Science
Sept 2021 – July 2025

- GPA: 3.92/4.0
- **Coursework:** Computer Fundamentals and Programming, OOP , OOAD , Calculus I , Linear Algebra, Statistics and Probability , Operating System , Computer Security , Natural Language Processing , Data Mining , Artificial Intelligence , Computer Vision, Databases, Data Mining , Signal and Image Processing

Fazaia Inter College Nurkhan Base , Chakala, FSc Pre-Engineering
July 2019 – June 2021

- Grade: 1078/1100
- **Coursework:** Physics , Chemistry , Mathematics

Experience

AI and ML Engineer, Tkrupt – Lahore (Remote), PK
Dec 2024 – April 2025

- Developed and deployed ML models, chatbots, and AI solutions using Langchain, OpenAI, DeepSeek, Huggingface, vector databases, AI agents, and integrated low-code/no-code platforms like Zapier and n8n, with PostgreSQL and Airtable for data management.

Machine Learning Intern, NESCOM, PK
Sep 2024 – Nov 2024

- Built Transformer-based models by using Pytorch and leveraged advanced techniques like Wave2Vec2.0 for nuanced audio feature extraction instead of mel-filter banks.
- Optimized classification with ResNet-1D, Transformer Encoders, and BiLSTM layers, while creating custom datasets with fake audio segments to improve model robustness and minimize error rates.

Machine Learning Intern, TruID NSTP – NUST H12 Islamabad, PK
July 2024 – Sep 2024

- Annotated images using CVAT.ai to create high-quality labeled datasets for training YOLO models to extract features like flags, chips, and holograms
- Prepared and managed datasets while addressing challenges like data imbalance and also learnt about Perspective Warping and homography
- Trained and optimized YOLO models to detect features of the card and trained classifier in tensorflow improving metrics like accuracy, precision, and recall for distinguishing real vs. fake chip-based cards.

Consumer Insights Data Analytics Externship with Beats and Extern , Remote
July 2024 – Aug 2024

- Scraped and analyzed Amazon reviews for Beats and 5 competitors, extracting actionable insights.
- Conducted data cleaning, feature engineering, EDA, and sentiment analysis to uncover key trends, leveraging Gemini for AI-driven consumer insights and proposing product improvements.

Management and leadership skills

GDGoC Machine Learning Lead ,Working as Machine Learning Lead in GDGoC(GDSC) PIEAS
2024 – 2025

IEEE Web Master ,Handling all Web Management issue of IEEE Student Branch PIEAS
2024 – 2025

Technical Lead of ML Lead Security in PIEAS Cyber Security Club
2024 – 2025

Projects

Weather Prediction using LSTMs, FB Forecast, Airflow, and Langgraph Project Link

- Project consists of two parts:
    1. Implemented a DAG/workflow in Airflow that fetches historical temperatures from an API, cleans the data, performs forecasting, saves results (JSON) in MinIO (AWS S3 replacement), and generates an HTML report using Langgraph with clothing suggestions, restaurant recommendations, and public places to visit.
    2. Dockerized and deployed as an endpoint on DigitalOcean, with the container managed via Kubernetes (kops).

Advanced Self-RAG using Langgraph, OpenAI, FastAPI, and Kubernetes Project Link

- Project consists of two parts:
    1. Implemented an advanced Self-RAG (Retrieval-Augmented Generation) system using Langgraph and OpenAI, enabling efficient information retrieval and generation.
    2. Dockerized and deployed as an endpoint using FastAPI on DigitalOcean, with the container managed via Kubernetes (kops), and set up a Jenkins pipeline for automated dockerization and deployment.

Full Stack Multimodal Sentiment Analysis (Video, text , audio) Project Link

- Implemented the Full Stack Multimodal Sentiment Analysis (Video, text , audio) using Nextjs, tailwind, Postgresql;Prisma for frontend and Lightning.ai, Pytorch, Huggingface for backend and used Modalai for deploying the endpoint for GPU(T4) usage.
- Tools/Terms: Pytorch, modal, Transformers, Huggingface, SQL.

Concatenative Audio Spoofing Detection by Waveform Boundary Detection Method Project Link

- Implemented the "Waveform Boundary Detection for Partially Spoofed Audio" research paper in Pytorch to detect the concatenative spoofed audio.Obtained an EER of 10%. It involved using 1d CNNs (1d Resnet), Transformers Encoder and BiLSTMs. Created Flask app and deployed it on huggingface spaces
- Tools/Terms: Pytorch, Flask, Transformers, Huggingface

SQL GPT Project Link

- Finetuned Gemma-2b-it model of Google on SQL queries dataset. You have to just give the table information (context) and the question. It will generate the SQL query for that question. Created Flask application and deployed it on huggingface spaces.
- Tools/Terms: LLMs, FineTunning, SQL, Transformers, Huggingface, Lora/QLora/PeFt.

Document Authentication Project Project Link

- Distinguished real vs. fake chip-based cards using YOLO models, training models to extract features like flags, chips, and holograms. I handled the imbalanced dataset and optimized metrics such as F1-Score and ROC-AUC, worked with Python, OpenCV, YOLO, and TensorFlow, and addressed data imbalance and used techniques like Perspective Warping. Obtained F1-Score of 95% and ROC-AUC 97%
- **Tools/Term**s: Tensorflow, Cvat.ai, Annotation, YOLO.

Languages

Mother Tongue: Urdu.

English: Listening (C2) Reading (B2) Writing (C1) Spoken Production (B2) Spoken Interaction (C1) .

Technologies/Skills

Languages: C++, C, Python, JavaScript, SQL.

Technologies: Machine Learning, Data Science, Data Analytics, Data Visualization, Linux, Python, TensorFlow, Keras, PyTorch, scikit-learn, Deep Neural Networks (CNNs, GANs, VAEs), AI, ML, NLP, C++, Prompt Engineering, SQL, GitHub, Git, LLMs, FineTunning , Automation(Zapier , Make.com), Langchain, Langflow , LangSmith , Docker , Kubernetes , Jenkins and AWS.

Volunteering

Library Management System for PMC College Nilore   Nov 2022
Collaborated with a peer to develop an LMS solution, addressing complex challenges in system design and implementation. Recognized with an appreciation letter from the Head of Department for delivering an effective institutional resource management system.

CERTIFICATE AND CERTIFICATIONS

Associate Data Scientist | **DataCamp**   Validate Aug, 2024

Generative Adversarial Networks (GANs) | **deeplearning.ai**   Validate Jul, 2023

Prompt Engineering for ChatGPT | **Vanderbilt University**   Validate Sep, 2023

EXTRA-CURRICULAR & INTERESTS

Playing Cricket and Table Tennis, Programming , Exploring AI World , Watching Seasons , Reading Research Papers.