Uzair Naeem

Date of birth: 12/01/2001 | **Nationality:** Pakistani | **Gender:** Male | **Phone number:** (+92) 3114288714 (Mobile) | **Email address:** uzairnaeem303@gmail.com | **LinkedIn:** www.linkedin.com/in/uzairnaeem | **GitHub:** https://github.com/UzairNaeem3 | **Hugging Face:** https://huggingface.co/uzairnaeem | **Address:** Shad Bagh, 05475, Lahore, Pakistan (Home)

ABOUT ME

AI and ML Engineer with a robust background in Mechatronics Engineering.

EDUCATION AND TRAINING

01/10/2018 – 15/06/2022 Lahore, Pakistan
B.SC MECHATRONICS & CONTROL ENGINEERING University of Engineering & Technology, Lahore

01/02/2023 – 30/04/2023 Lahore, Pakistan
DATA SCIENCE BOOTCAMP Programmers Force (Pvt) Ltd

WORK EXPERIENCE

07/2024 – CURRENT Lahore, Pakistan
DATA SCIENTIST BIG DATA ANALYTICS

- Developed a **BERT-based classification model** to predict user profiles from tweets, achieving 90% accuracy.
- Developed the entire backend for a website offering AI services in the legal domain, using **Python** and **Flask,** set up **JWT authentication** for user security, and used **MySQL** for database management.
- Added **Stripe payment processing** to support free, monthly, and annual subscription plans (Freemium, Essential, and Professional).
- Used the **Llama3** language model to help the website answer user questions with AI-powered responses.
- Utilized **PySpark** and **Delta Lake** for efficient big data processing, improving performance over traditional SQL queries.
- Used **LangGraph** to create AI systems that work together effectively, making it easier to handle complex tasks with multiple agents.
- Currently building APIs using **Django** and **Django REST Framework**, collaborating with frontend developers to connect the front end to these APIs, and coordinating with the DevOps team to deploy the product.

07/2023 – 06/2024 Lahore, Pakistan
ASSOCIATE DATA SCIENTIST STECH AI

- Developed robust APIs for machine learning models using **FastAPI**, **Django**, and **Docker**.
- Leveraged **Llama Index**, **Langchain**, and **OpenAI** frameworks to enhance performance and user experience in RAG applications.
- Worked with various RAG types, including **Raptor RAG**, **Corrected RAG (CRAG)**, **Self RAG**, **Graph RAG**, and **Multimodal RAG**.
- Designed a user-friendly front end using **Streamlit** and **Gradio**, facilitating seamless interaction with the developed APIs.
- Developed an AI-powered recruitment assistant using **OpenAI** models and **HuggingFace** library to automate the candidate selection process, resulting in a more efficient and accurate evaluation of candidate fit for job openings.
- Utilized **NLP** techniques to analyze job descriptions and identify relevant skills in candidate resumes, increasing the accuracy of candidate ranking and reducing the time spent on manual screening.
- Utilized custom vector databases such as **Qdrant**, **Chroma**, **Pinecone**, and **AWS OpenSearch** to efficiently store and retrieve data.
- Implemented the **PEFT LoRA** and **QLoRA** methods to fine-tune various LLM models such as **Llama3**, **Falcon**, and **Bert**, enhancing their performance while reducing computational and storage costs.
- Created autonomous agents using **Langchain**, equipped with specialized tools for complex workflows. These multi-agent systems enhance the overall performance of LLM-based applications.

05/2023 – 06/2023 Lahore, Pakistan
TRAINEE ASSOCIATE DATA SCIENTIST PROGRAMMERS FORCE

- As a member of the AI team, I was responsible for developing innovative artificial intelligence solutions, with a specific focus on optimizing **Know Your Customer (KYC) application** tailored to diverse industries.
- I worked on this state-of-the-art AI system that leveraged deep learning and natural language processing to solve complex problems and make accurate predictions based on large-scale data.
- I identified and analyzed the areas within the KYC application where the performance was suboptimal and the results were inaccurate.

- I extracted and prepared relevant data from the database using SQL queries and data manipulation tools, and cleaned and processed the data for further analysis.
- Designed, built, and trained a new model using advanced techniques and tools such as TensorFlow, PyTorch, and Scikit-learn to automate and improve the deficient AI steps.

DIGITAL SKILLS

Machine learning (Tensorflow, Keras, Pytorch, Scikit-Learn, OpenCV) | NLP: NLTK, spaCy, Gensim, HuggingFace, Langchain, LlamaIndex, Transformers | Databases: MySQL, PostgreSQL, MongoDB, Vector Database (Qdrant, Chromadb, Pinecone), Graph db | Web Frameworks: Django, Flask, FastAPI | Linux | Git and GitHub | Docker | SQL | Postman | Programming Languages: Python, C++

PROJECTS

Image Segmentation with UNET: A Comprehensive Implementation of UNET Architecture for Precise and Efficient Medical Image Segmentation

Link https://github.com/UzairNaeem3/Deep_Learning/blob/main/UNET_Model.ipynb

Efficient LLM's Optimization: Falcon and Llama2 Fine-Tuning on Resource-Constrained Devices

Link https://github.com/UzairNaeem3/LLMs/tree/main/LLM's

Autonomous Agents: An Exploration of Langchain and Autogen-Based Multi-Agent Systems for Large Language Models

Link https://github.com/UzairNaeem3/LLMs/tree/main/Agents

Image Enhancement with Denoising Autoencoders: An Investigation into Image Denoising Techniques for Improved Visual Performance

Link https://github.com/UzairNaeem3/Deep_Learning/blob/main/Denoising_Autoencoder.ipynb

Forecasting Future Global Temperature Trends: A Time Series Analysis Approach

Link https://github.com/UzairNaeem3/Global_Climate_Temperature_Prediction

BERT-Based Tweet Classification: Leveraging BERT to Develop a High-Accuracy Model for Predicting User Identities from Social Media Tweets

Link https://github.com/UzairNaeem3/LLMs/blob/main/LLM's/BERT/BERT_UserTweets_Classification.ipynb

CERTIFICATIONS

Data Science Mastery - IBM

Links https://www.coursera.org/account/accomplishments/verify/XGK9KFWGVMQ9 | https://www.coursera.org/account/accomplishments/verify/VJY8T9CS9XTJ | https://www.coursera.org/account/accomplishments/verify/JDV3PTAAEXFW | https://www.coursera.org/account/accomplishments/verify/4Z2PB75FRT3R

Natural Language Processing - DeepLearning.AI

Link https://www.coursera.org/account/accomplishments/verify/MYEHWY4M6BHN

Machine Learning with Python - IBM

Link https://www.coursera.org/account/accomplishments/verify/5DBA4CX9XXYD

LANGUAGE SKILLS

Mother tongue(s): **URDU**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	C1	C2	B2	B2	B2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user