# Athar Parvez

Dhahran , KSA | +966548217161 | atharparvezce@gmail.com | LinkedIn | GitHub

**AI/ML Engineer & Research Assistant(4+ years).** Deploying ML models and LLM solutions to production. Cut churn for Drahim a Saudi fintech with predictive modeling using ML; delivered RAG chatbots and fine-tuned LLMs. Expertise in Python, PyTorch/TensorFlow, LangChain, and AWS (Docker, MLflow); fluent in NLP/CV, end-to-end pipelines, and MLOps best practices.

## SKILLS

**Machine Learning & Deep Learning :** Python, Scikit-Learn, TensorFlow, PyTorch, Keras, OpenCV, Pandas, NumPy, NLTK, spaCy, GANs, Transformers, Feature Engineering, Hyperparameter Tuning(Optuna), Model Evaluation (F1-score, ROC-AUC, Accuracy) and MLOPs using Next_.

**Generative AI:** LangChain, LangGraph, OpenAI, Gemini, HuggingFace, Retrieval-Augmented Generation (RAG), Large Language Models (LLMs), Fine-tuning (including LoRA, QLoRA, and model quantization techniques), Transfer Learning, Prompt Engineering.

**AI Automation & Orchestration:** n8n, Zapier, MLflow and LangFlow.

**Cloud and Databases:** AWS, MongoDB, MySQL and Oracle.

**Development & Version Control:** Git and Docker.

**Web Development:** HTML, CSS, FastAPI (Postman, Thunder Client, Insomnia), Streamlit and Flask.

**Data Visualization & Microsoft Tools**: Seaborn, Matplotlib, PySpark, Tableau, Power BI, Excel, PowerPoint and Word.

## WORK EXPERIENCE

**Research Assistant** (Jan, 2025 – present)
**King Fahd University of Petroleum and Minerals(KFUPM)**

- Conducted advanced research in **Machine Learning**, contributing to the development of innovative, data-driven solutions aligned with project objectives.
- Designed and executed experiments; collected, pre-processed, and analysed large-scale datasets using machine learning and deep learning frameworks including ***Seaborn, Matplotlib, PySpark, Pandas, NumPy, Scikit-learn, PyTorch, TensorFlow, NLTK, spaCy,*** **and** ***OpenCV.***
- Developed comprehensive technical documentation, research reports, and professional presentations to effectively communicate findings to academic and industry stakeholders.
- Mentored undergraduate students by supervising their research tasks, providing technical guidance, and promoting collaborative teamwork to enhance their academic and practical skills.

**Machine Learning Intern** (Sep, 2025 – Sep 2025)
**SkillCraft Technology(SCT)**

- Collected, cleaned, and structured raw datasets to ensure high-quality model input.
- Built baseline **machine learning models**, performed hyperparameter tuning, and evaluated performance.
- Learned and applied **advanced ML techniques** in supervised and unsupervised learning.
- Gained hands-on exposure to **end-to-end ML workflow**, including data gathering, preprocessing, model creation, and evaluation.

**Teaching Assistant** (Jan, 2024 – Jan 2025)
**King Fahd University of Petroleum and Minerals(KFUPM)**

- Assisted in teaching fundamental programming concepts using **C and Python** to approximately 60 undergraduate students per semester.
- **Led weekly lab sessions**, providing hands-on guidance in debugging, code structure, and best programming practices.
- Designed, administered, and graded lab tests and programming assignments.

**AI/ML Engineer (3 Years)** (April, 2021 – Jan, 2024)
**Tata Consultancy Services(TCS)**

- Developed and deployed **Machine Learning and Deep Learning models** to support financial applications, including tasks such as **classification, prediction, and anomaly detection**.
- Fine-tuned **pre-trained models and LLMs** using techniques like **LoRA, QLoRA, and transfer learning**, adapting them to domain-specific datasets.
- Built **end-to-end ML pipelines** from data preprocessing, feature engineering, and training to deployment, ensuring models met **business and performance requirements**.

## PUBLICATIONS

**Generalizable Diabetes Risk Stratification via Hybrid Machine Learning Models**
**Submitted to *Journal of Diabetes and Its Complications Q1)*** ***(Confidential)***

- Contributed as **a first author to a research study** on diabetes risk prediction.
- Developed hybrid machine learning models for accurate risk stratification, combining traditional statistical methods with advanced ML techniques.
- Collaborated with researchers to design experiments, preprocess clinical datasets, and evaluate model generalizability.

**Usability Requirements Catalog for Modeling Dynamic User Interactions in Mobile Applications**

- Proposes a graph-based usability requirements catalog for dynamic mobile interactions (touch, gestures)
- Structures relationships among key factors (e.g., Efficiency, Learnability, Errors) and maps them to design/evaluation.
- Validated with experts to ensure practical applicability for developers

## PROJECTS

**Churn Prediction for Fintech Retention Strategies** ***(Confidential)***

- Developed a churn prediction model for *Drahim*, a Saudi-based fintech company, using machine learning techniques to identify at-risk customers and support retention strategies.

**Purchase Order Item Categorization.** ***(GitHub Link)***

- Processed multilingual purchase order text data and built an ML pipeline with embeddings and clustering to classify items, enhancing categorization accuracy and spend analysis.

**Duplicate Question Detection on Quora Dataset** ***(GitHub Link)***

- Built an NLP pipeline with ML models to classify duplicate questions using token, fuzzy, and semantic features on Quora Dataset.
- Deployed an interactive Streamlit app for real-time duplicate detection on Heroku.

**iArabUI: A Large Dataset for Arabic-Supported iOS Mobile Applications** ***(GitHub Link)***

- Creating iArabUI, the first comprehensive dataset of Arabic-language iOS app interfaces, leveraging LLMs for automated UI labeling to advance accessibility, HCI, and intelligent UI design.

**YouTube RAG: A Chatbot for YouTube Video Transcript Understanding** ***(GitHub Link)***

- Built a Retrieval-Augmented Generation (RAG) chatbot that answers questions about YouTube videos by extracting transcripts via YouTubeLoader, converting them into embeddings using FAISS/Chroma, and generating responses using OpenAI/HuggingFace LLMs based on semantically retrieved content.

**TCS BaNCS: AI-Driven Predictive Analytics Platform for Financial Services** ***(GitHub Link)***

- Developed and fine-tuned ML/DL models and LLMs for predictive analytics, integrating SQL data pipelines and deploying AI-driven microservices with Docker, Kubernetes, and AWS.

## EDUCATION

- **King Fahd University of Petroleum and Minerals (67 QS World Wide in 2025)**
  Master's in Computer Science (GPA 3.4/4) *(Jan 2024 - Present)*

- **Maulana Azad National Urdu University**
  Bachelors in Computer Science **(**GPA 3.5/4) **(** *Aug 2016 - May 2020)*

## CERTIFICATIONS

- **Python Basics** – University of Michigan, Jan 2024 | Credential ID: CZ8HXFHOXK8N
- **LangChain for LLM Application Development** – DeepLearning.AI, Aug 2025
- **Finetuning Large Language Models** – DeepLearning.AI, Feb 2025
- **Generative AI for Everyone** – DeepLearning.AI, June 2025
- **AI Agents in LangGraph** – DeepLearning.AI, April 2025
- **n8n Course: No Code AI Agent Builder** – Simplilearn, Aug 2025

## REFERENCES


- **Dr Omar Jamal Hammad**
  Assistant Professor, Information & Computer Science Department, KFUPM (omarjh@kfupm.edu.sa)

- **Dr Adel Fadhl Ahmed Noor**
  Assistant Professor, Information & Computer Science Department, KFUPM (adelahmed@kfupm.edu.sa)

- **Dr Jameel Ahmed**
  Assistant Professor, Department of Computer Science, University of Bisha (jhussain@ub.edu.sa.sa)