Fiza Gafoor M

fizagafoorm@gmail.com | Phone: +974 39920488
LinkedIn | GitHub
Nationality: Indian | Location: Doha, Qatar | Date of Birth: 07 July 2000

Professional Summary

AI Engineer with 1+ years of experience in designing, deploying, and optimizing AI/ML models for real-world applications. Expertise in NLP, computer vision, OCR, and cloud-based AI solutions. Skilled in Python, TensorFlow, PyTorch, and AI/ML. Passionate about automating workflows, improving model accuracy, and integrating deep learning models into production environments.

Education

- **M.Sc. in Computer Science with Specialization in Data Analytics**
  Digital University Kerala 2022 – 2024
- **B.Sc. in Statistics**
  Farook College (Autonomous), Calicut 2019 – 2022

Work Experience

AI Engineer – Innovation Incubator Advisory Aug 2024 – Present

- Developing a chatbot using AWS Bedrock, currently in its initial phase to automate workflows and enhance user interaction.
- Optimized the Appraisal Digitization Tool (ADT) pipeline to resolve production challenges.
- Conducted research on state-of-the-art deep learning models for document processing.
- Deployed models for document classification and information extraction.
- Configured TensorFlow Serving and FastAPI for model deployment.

AI Engineer Intern – Innovation Incubator Advisory Feb 2024 – Aug 2024

- Built Named Entity Recognition (NER) models using SpaCy and fine-tuned OpenAI GPT-4.
- Worked on OCR pipelines and document classification for insurance data processing.
- Conducted sentiment analysis on disaster tweets using NLP techniques.

Data Science Virtual Intern – OASIS-INFOBYTE Sep 2023 – Oct 2023

- Developed ML models for car price prediction, email spam detection, and Iris flower classification.

Research Intern – UNNATHI (IIM-Kozhikode) Mar 2022 – Mar 2023

- Assisted in data collection, analysis, career mentoring, and event coordination.

Technical Skills

- **Programming:** Python, SQL
- **ML/DL Frameworks:** TensorFlow, PyTorch, Scikit-learn, Keras
- **NLP Tools:** SpaCy, OpenAI GPT-4, BERT, Gemini
- **OCR Tools:** Tesseract, PaddleOCR, AWS Textract

- **Cloud Platforms:** AWS (Bedrock)
- **Model Deployment:** TensorFlow Serving, FastAPI, Docker, Kubernetes
- **Tools:** Streamlit, Jenkins, RabbitMQ, MongoDB, NoSQL Booster
- **Data Analysis:** Pandas, NumPy, Excel, SPSS
- **Data Visualization:** Matplotlib, Seaborn
- **Version Control:** Git, Bitbucket

Projects

- **Medical Image Classification for Malaria Detection** (09/2023–01/2024)
  Developed a CNN model to classify blood sample images, achieving high diagnostic accuracy.
- **EmoDiarize: Emotion-Aware Speaker Diarization** (04/2023–09/2023)
  Built a deep learning system for real-time speech emotion recognition and speaker diarization.
- **Multi-Modal Deep Learning Classifier**
  Created a Streamlit-based app for tumor detection and sentiment classification.
- **Proof of Concept: GraphReader**
  Developed a POC based on "GraphReader: Building Graph-based Agent to Enhance Long-Context Abilities of Large Language Models" (arXiv). Used Neo4j for graph management and GPT models for prompting, analysis, and answer generation.
- **Automated Invoice Data Extraction using Gemini API**
  Designed an AI pipeline to extract key entities from publicly available invoice data by prompting Gemini API.

Publications

- **[2024]** Sibinraj V. M., Fiza Gafoor, et al. Automated Malaria Cell Image Classification Using CNN. *TechRxiv Preprint*. https://doi.org/10.36227/techrxiv.171744390.01569902/v1
- **[2023]** Hamza, H., Gafoor, F., et al. EmoDiarize: Emotion Identification from Speech Signals Using CNNs. *arXiv Preprint*. https://arxiv.org/abs/2310.12851

Certifications

- **Introduction to Data Analysis using Excel** (Coursera)
- **Statistical Analysis with SPSS** (Farook College)
- **Data Analytics Virtual Internship** (KPMG via Forage)
- **SQL for Data Science (Ongoing)** (Coursera)

Languages

- **English:** Advanced (C1)
- **Malayalam:** Native (C2)
- **Hindi:** Basic Speaking and Reading
- **Arabic:** Reading