Kheem Parkash Dharmani

Kheem-Dh | Kheem | dharmanikheem@gmail.com | +92 336 318 2779 | Kheem's Portfolio

Summary

I am a Machine Learning Engineer with a Master's in Data Science and a strong foundation in AI for healthcare, particularly in natural language processing, multi-modal learning, and clinical applications of deep learning. My research focuses on explainable, personalized AI systems for medical diagnostics. Through my Master's thesis on radiology report generation using Vision Transformers and LLMs, and my current role building large-scale RAG systems at Clinical Pearl, I have demonstrated the ability to design and implement AI systems that align with the goals of predictive, preventive, and personalized medicine. I now seek a PhD opportunity where I can contribute to advancing robust, human-centered AI methodologies—particularly in biomedical imaging, clinical language models, and explainability frameworks that improve trust and usability in real-world healthcare deployments.

Work Experience

Lead AI Engineer at **Clinical Pearl, Princeton (REMOTE)** Apr 2025 – Present

- Architect and oversee production LLM pipelines (GPT-4o, LLAMA 3, Mistral-8×7B) powering clinical-trial evidence synthesis and medical-chat features for 50k monthly users.
- Built a multi-tenant, multimodal RAG service (text + imaging + tabular) that cut physician search time by **40%** and increased answer accuracy (BLEU +23, ROUGE-L +18).
- Introduced continuous fine-tuning (LoRA, QLoRA) and evaluation harness—reducing model-update cycle from 2 weeks to 2 days while keeping PII-safe via on-prem embeddings.
- Mentored a team of 3 engineers; instituted AI development practices (GitHub Actions + API Endpoints).

Machine Learning Engineer at **Cplus Soft, Islāmābād** Dec 2023 - Mar 2024

- Developed generative AI solutions focused on question answering (QA) and text summarization, optimizing task performance by 10–30% in various use cases.
- Integrated LangChain for streamlined AI applications, improving the modularity and scalability of workflows by 25%.
- Conducted research on emerging trends in generative AI to drive innovation and improve model performance.

Machine Learning Trainer at **AI Lounge, Gilgit-Baltistan** Jul 2023 - Nov 2023

- Designed and delivered a comprehensive Data Science and AI curriculum, equipping participants with practical skills in machine learning, deep learning, and data visualization.
- Conducted interactive training sessions with a focus on practical skills and real-world applications.
- Enhanced the skill set of the participants, contributing to the advancement of the regional tech industry.

NLP Teaching Assistant at **FAST-NUCES, Islamabad** Jan 2023 - Jun 2023

- Enhanced student understanding by simplifying complex NLP concepts, achieving a 95% course satisfaction rate.
- Provided mentorship and guidance, fostering a supportive and collaborative learning environment.

- Developed and implemented algorithms for NLP tasks, solving practical challenges and driving innovation in the classroom.

Projects

Medical Knowledge Retrieval and Question-Answering System

- Built a FastAPI-based application for medical queries using a Retrieval-Augmented Generation (RAG) approach, enhancing information retrieval efficiency by 40%.
- Developed a vector store from structured medical data via OpenAI and LangChain, improving retrieval accuracy for clinical use cases.

Master's Thesis: Radiology Report Generation

- Engineered a novel solution combining Vision Transformers and custom-trained Large Language Models for X-ray interpretation, achieving state-of-the-art diagnostic accuracy.
- Demonstrated the potential of AI-driven personalized diagnostics, bridging healthcare and machine learning.

ETL Pipeline for Churn Prediction using MS Azure

- Designed an end-to-end ETL pipeline integrating OLTP streams into a structured SQL database hosted on MS Azure.
- Leveraged Azure ADF and Streamlit to build a churn prediction system, enabling proactive customer retention strategies with an accuracy boost of 18

Skills

Programming Languages	Python, R, SQL
Machine-Learning Frameworks	PyTorch, TensorFlow, Keras, Scikit-Learn
Deep Learning	CNNs, RNNs, LSTMs, Autoencoders
NLP & Large-Language Models	Transformers, NLTK, SpaCy, BERT, GPT(-2/3/4o)
Cloud Platforms	AWS Solutions Arch., Azure (ADF, Cognitive Services)
Data Visualization	Tableau, Matplotlib, Seaborn
Version Control	Git, GitHub
Automation & Scraping	Selenium, BeautifulSoup, Playwright
RAG & Model Fine-Tuning	Mono-/multimodal RAG pipelines, LoRA/QLoRA, custom LLM fine-tuning

Certifications

Machine Learning Specialization	*DeepLearning.AI, Dec 2023*
Advanced Learning Algorithms	*DeepLearning.AI, Oct 2023*
Python for Data Science, AI & Development	*IBM, Mar 2023*
Supervised Machine Learning: Regression and Classification	*Stanford Online, Mar 2023*
Data Science For Everyone	*DataCamp, Apr 2022*
Machine Learning For Everyone	*DataCamp, Apr 2022*

Education

2022 - 2024 **MS** (Data Science) at **FAST NUCES, Islamabad** (CGPA: 3.2/4.0)
Research: "A Hybrid Approach for Radiology Report Generation using Vision Transformers & Language Models." Enhanced automated lung cancer diagnosis via deep learning, leveraging radiology literature and the MIMIC CXR dataset. The research innovatively combined vision transformers with large language models, optimizing report generation and summarization, and providing critical insights for improving patient care outcomes.

2017 - 2021 **B.E** (Computer Systems) at **Mehran, UET, Jamshoro** (CGPA: 3.7/4.0)

Courses

DS5007 - Natural Language Processing
DS5002 - Data Science Tools and Techniques
SAA-C03 - AWS Solution Architecture
DS5006 - Deep Learning
DS5004 - Machine Learning
Data Mining

Achievements

IEEE eLearning Course Trial Award Sep 2023
IEEE Xplore Challenge for Researchers in Pakistan 2023
Honored with the IEEE eLearning Course Trial award, presented by IEEE Xplore, in recognition of academic excellence and dedication. This significant achievement, endorsed by FAST-NUCES, Islamabad Campus, culminated in winning an Apple iPad.

Publications

Optimizing Sentiment Integration in Image Captioning Using Transformer-Based Fusion Strategies

CMC–Computers, Materials & Continua, Jul 2025

K. R. Narejo, H. Zan, **K. P. Dharmani** *et al.*—Proposes the Sentiment-Driven Caption Generator (SDCG) that fuses transformer-based visual and textual representations, achieving 94.5 % sentiment accuracy and more human-like captions.

EEBERT: An Emoji-Enhanced BERT Fine-Tuning on Amazon Product Reviews for Text Sentiment Classification

IEEE Access, Sep 2024

K. R. Narejo, H. Zan, **K. P. Dharmani***, L. Zhou, T. J. Alahmadi, N. Sehito, Y. Y. Ghadi, M. Assam—Introduces the Sentiment Adjustment Factor (SAF) for emoji tokenization, boosting accuracy to 99.2 % on a 26 k-review subset.

Enhancing Emoji-Based Sentiment Classification in Urdu Tweets: Fusion Strategies With Multilingual BERT and Emoji Embeddings

IEEE Access, Aug 2024

K. R. Narejo, H. Zan, D. Oralbekova, **K. P. Dharmani***, O. Mamyrbayev, K. Mukhsina—Demonstrates that multilingual BERT plus emoji embeddings significantly improves sentiment analysis for low-resource Urdu social-media data.