



MUHAMMAD AHMED

SR. AI ENGINEER

PERSONAL PROFILE

Experienced AI Engineer with proficiency in developing complex AI solutions, architectures, and scalable deployments on Cloud services. Skilled in building infrastructure, optimizing models, boosting inference speed, and finetuning models. Also had experience in teaching web development and developing Android application.

WORK EXPERIENCE

SR. AI Engineer

Vyro AI | March 2024 - Present

Imagine Art

Completely changed their model infrastructure from single pod service to load all models on CPU and on request switch to GPU so that no need to run a lot of pods for each service on a single pod.

Musicly

Built a complete voice cloning singing product for the company with the best accuracy possible. Owner of the product with complete architecture till the deployment.

SR. AI Engineer/Team Lead

EKKEL AI | July 2023 - March 2024

Rizz

Standalone build Custom Queue based scalable multi-user stable diffusion finetuning and images generation product deployed on HuggingFace inference endpoint.

Meta1st

A cybersecurity training courses generation project which creates courses based on security threat based on the spam email they received.

Metavox

Speech filtering project which can analyze and select the desired content in the audio based on the original transcript for voice over artists.

CareerLi

Worked with 20 years+ experienced career coach to build this resume building application with complex prompting techniques.

Lunchbreak

Build the solution and led the AI team to create AI detection and bypassing tool that has the best accuracy from its competitors.

SellerIQ

Built RAG based real-time streaming sales associate chatbot that can answer queries of clients during zoom call.

AI Engineer

EKKEL AI | May 2022 - present


Kodezi (Grammarly for programmers)

Currently serving as an ML Engineer/Technical Lead at Kodezi and leading the AI team. Developed end-to-end tools for Kodezi and utilized Node.js to implement efficient pipelines for VS Code extension integration.

CONTACT ME AT

 Lahore, Pakistan, 54000

 i.am.a.pakistani.programmer@gmail.com

 <https://www.linkedin.com/in/muhammad-ahmed-1a8a7719b/>

 <https://github.com/muhammad-ahmed-ghani>

 https://www.producthunt.com/@muhammad_ahmed_ghani

 <https://huggingface.co/ahmedghani>

 +923014902083

SKILLS

- Python, C++, C, Javascript, Kotlin
- Flask, FastAPI, Socketio, HTML, CSS, Bootstrap
- AWS, GCP, Runpod, Vast.ai, HuggingFace
- Github, Docker
- MySQL, MongoDB, Firebase, Cloud Firestore, Room DB
- Leadership, Team Management, Solution Architecture Building

EXPERIENCED BUILDING PROJECTS WITH

- Stable Diffusion v1.4/v1.5/v2/v2.1/XL
- GPT 3/3.5/4/4v/4t/4o, Gemini 1.5
- DALL.E-2/DALL.E-3
- Open Source LLM's (Llama2/3, Phi3 etc.)
- Nvidia Nemo/Riva Realtime (TTS, STT)
- Comfy UI (Workflows)

loopsie

Developed and implemented a project with autoscaling on a cloud platform. The application uses generative AI to create avatars and high res images.

profilephotosai

Developed and implemented a project with autoscaling on a cloud platform. The website is able to train an AI model with 20 images and generate over 100 diverse images with the best prompts created from multiple AI models and variations.

writingo

Served as Technical Lead in the development of an AI-based social media post generation application, utilizing GPT-3 prompt engineering.

Snappin (Text based image editing)

Developed a scalable image inpainting project for Ekkel based on stable diffusion. Optimized and traced the model for improved inference, and deployed it on a scalable platform to serve as an API.

● Associate AI Developer

EKKEL AI | Sept 2021 - April 2022

Arvify

Created End to End multi-user scalable containerized solution in NodeJS for Unreal Engine Pixel Streaming and deployed it on aws

Multi-User Scalable Real-Time ASR System (using Nvidia Riva)

An AI model that will transcribe text with beam search and has the lowest latency and higher accuracy.

AV Society

Finetuned the StyleGAN3-Nada model to generate avatars from images in High resolution.

Online Speaker Diarization

An application to detect who speak when. The speaker classification using Spectral Clustering.

● F Y P

Multiple Voice Separation for Unknown Number of speakers

An AI model to separate multiple voices from a mixed signal and trained on LibriMix dataset.

● PROFESSIONAL CERTIFICATIONS

AI Specialization - Samsung Innovation Campus

DeepLearning Specialization - DeepLearning.AI

TensorFlow Developer Specialization - DeepLearning.AI

Intro to AR/VR/MR/XR - University of Michigin

Web Development - EwigLife

OpenCV for Computer Vision - LinkedIn

● EDUCATION

BS Computer Science

3.52 CGPA

University Of Central Punjab

2018-2022