

KAMRAN SHAIKH

kamran.aiml.engineer@gmail.com | sk2787434@gmail.com – +91 8591246016 – Mumbai, India

EXPERIENCE

Biz Technologies - (Decentrwood)

AI/ML Engineer

September 2025 – Present

Mumbai

- Leveraging the Gemini API for image, video, and music generation to explore multimodal AI capabilities.
- Developing and training machine learning models for cryptocurrency price prediction, utilizing time-series forecasting and deep learning techniques.

HealthIndia Insurance TPA Services Pvt. Ltd

Machine Learning Engineer

April 2024 – September 2025

Mumbai

- Developing AI/ML solutions for automating healthcare document processing, including claims forms, bills, and medical reports.
- Building end-to-end NLP and computer vision pipelines using YOLO, Azure OCR, and LLMs (LLaMA, GPT) for data extraction and classification.
- Designing scalable tools for automating NEFT extraction, CCN generation, and invoice digitization using Python, FastAPI, and OpenAI.
- Collaborating with cross-functional teams to deploy models and tools into production environments, improving operational efficiency and accuracy.

DG Market

Python Developer

Apr 2023 – Apr 2024

Mumbai

- Built Python-based automation tools and APIs for data extraction, transformation, and reporting.
- Developed robust pipelines for scraping, processing, and storing large datasets used in global tender analytics.

Ardentisys-Tebillion

Associate Engineer

October 2022 – March 2023

Mumbai

- Contributed to the development of business automation solutions, focusing on backend logic, API integrations, and process optimization using Python and automation tools.
- Supported product teams in building scalable software modules for client-driven enterprise solutions.

SKILLS

- LLMs & AI Agents: Fine-tuning, Retrieval-Augmented Generation (RAG), Tool-using Agents, Chatbots, LLM Prompts, GPT-4, LLaMA, Gemma, Bloom, Ollama, LangChain, LangGraph
- Machine Learning & Deep Learning: Supervised/unsupervised ML, neural networks, CNNs, LSTMs, GANs, Libraries: Scikit-learn, TensorFlow, PyTorch, Keras
- NLP & Document AI: Named Entity Recognition (NER), Text Classification, OCR, Document Parsing, Tools: SpaCy, BERT, Azure OCR, Tesseract, OpenCV, YOLO, Stable Diffusion
- Programming & Automation: Languages: Python, SQL, Tools: Pandas, NumPy, Selenium, Regex, FastAPI

- **Deployment & Infrastructure:** Experience with cloud and on-prem environments, AWS EC2 & S3, Linux servers, ChromaDB
- **Data Engineering:** Web scraping, data pipelines, SQL-based data storage and integration, SQL, ER diagrams, data validation checks

PROJECTS

Document Classification System YOLO, CNN, LSTM, GRU, Python

Automated classification of insurance claim documents using a multi-stage pipeline combining object detection, image classification, and text classification.

Auto NEFT Extraction YOLOv8s, Tesseract OCR, Python

Extracted NEFT data from scanned cheques using object detection, OCR, and regex-based postprocessing for structured output.

Auto CCN Generation YOLO, Azure OCR, LLM (LLaMA/Gemma), Python

Developed an AI pipeline to extract data from healthcare claim forms and generate Control Claim Numbers (CCNs) automatically.

Auto Bill Entry System LLaMA 3.1, Python

Built a system to extract structured data from invoices and generate JSON outputs for seamless database integration.

Aadhaar Card Masking Tool YOLOv8s, OpenCV, Python

Designed a privacy tool to detect and redact Aadhaar numbers from scanned documents using real-time object detection.

Medical NER System SpaCy, BERT, Python

Trained custom NER models to extract patient names, diagnoses, treatments, and doctor information from medical records.

ITVx – Web Scraper & Data Pipeline Selenium, Python, AWS S3, SQL

Built an automated web scraper and data pipeline to extract, clean, and store structured data.

Chatbot with Memory GPT-API, LangChain, OpenAI, Python

Created an intelligent chatbot using LangChain agents capable of memory, context retention, and tool usage.

PDF Chatbot LangChain, ChromaDB, OpenAI, PyMuPDF

Built a document Q&A chatbot that can ingest and answer questions from PDF files.

Sentiment Analysis CountVectorizer, Scikit-learn, Python

Implemented a machine learning model to classify customer sentiments in product reviews.

Face Recognition Attendance System OpenCV, Dlib, Python

A real-time facial recognition system for automated employee attendance tracking.

Stock Market Prediction Python, Pandas, Matplotlib, LSTM

Used time series forecasting techniques to predict future stock prices for cryptocurrency (Bitcoin).

EDUCATION

M.H. Saboo Siddik College of Engineering

B.E. in Electronics Engineering

2015 – 2021

Mumbai

CERTIFICATIONS

Master's Certification in Data Science

Covered machine learning, deep learning, NLP, Python, and real-world data science projects.