

NIKHIL SHINGADIYA

MACHINE LEARNING ENGINEER

Ahmedabad,Gujarat,India| P: +91 6354071470 | shingadiyanikhil1634@gmail.com | [LinkedIn](#) | [Github](#)

About Me ([Portfolio](#))

I am a Machine Learning Engineer with over 3+ years of experience designing and deploying innovative AI systems for next-generation storytelling. My work is driven by a robust foundation in data science and a passion for practical AI applications. I excel in crafting and implementing machine learning systems that empower data-driven decision-making and continuously push the boundaries of algorithm development.

TECHNOLOGY

Technical Skills:

Statistics and Probability, Python, Data Visualization, Exploratory Data Analysis, Machine Learning, Web scraping,Pyspark, ANN, CNN,RNN,Transformers, LLM,HuggingFace,Generative AI,Docker, Docker-compose, MySQL, MongoDB,Object-Oriented programming, FastAPI,Celery, Redis Familiar with M-flow,AWS(EC2,S3,ECR,Lambda).

Libraries:

BeautifulSoup, Selenium, Numpy, Pandas, Matplotlib, Seaborn, SciPy, Scikit-learn, Statsmodels, Dash Plotly, Tensorflow, Keras, OpenCV, Langchain,Llamaindex.

IDE: Visual Studio Code, Jupyter Notebook, Spyder, Google-Colab, And Pycharm.

Version Control : Git & Github **Operating System:** Linux (Ubuntu 20.04 LTS), Windows, Mac Os.

WORK EXPERIENCE

Unlimited WP - (100-150 employees)

Ahmedabad, Gujarat

AI/ML LEAD DEVELOPER

Present – Dec 2023

★ Project 1: [WeamAI](#)

● Objective:

- Develop an **AI-powered platform** that enables companies to register, onboard employees, and engage in collaborative communication with integrated company documents—streamlining business operations through advanced AI.

● Tech Stack:

- FastAPI, Celery, Redis, Docker-compose, LangChain,Llmaindex, Ray framework, Pinecone, Qdrant, Generative AI, LLM models,OpenAI,HuggingFace,Anthropic, Streamlit,EC2 Instance.

● Key Features & Achievements:

- Engineered robust containerization with **Docker** to ensure scalable and reliable deployments.
- A **Layered Architecture** was designed that enhanced maintainability and scalability.
- Developed a comprehensive **CI/CD** pipeline to automate testing, deployment, and monitoring, reducing release cycles.
- Built a custom **Retrieval-Augmented Generation (RAG)** pipeline, significantly improving the accuracy and relevance of Q&A responses.

- Integrated **LangChain** agents to facilitate intelligent document interactions and guide users through sequential actions.
- Employed **Pinecone** and **Qdrant** for efficient storage and retrieval of embedding vectors, optimizing AI model performance.
- Utilized the Ray framework to maximize **GPU utilization**, ensuring efficient training and inference of custom LLMs.
- Integrated models from **OpenAI, HuggingFace, Gemini and Anthropic** to enhance the platform's AI capabilities.
- Successfully deployed the entire AI backend on **Amazon EC2** for robust performance.
- Developed custom metrics for scaling **Celery workers** via **CloudWatch**, improving resource efficiency.
- Created a custom module to resolve **LangChain & FastAPI** bugs, bolstering overall system stability.

★ Project 2: GARUDA AI

● Objective:

- Develop an AI-driven blog writing engine designed to boost website blog rankings by generating **SEO-Optimized Content**.

● Tech Stack:

- Google Keyword Planner, Ahrefs, Generative AI, LLM models.

● Key Features & Achievements:

- Utilize **Google Keyword Planner** to conduct in-depth keyword search volume analysis and identify high-impact keywords.
- Scrape data from targeted websites using **Ahrefs** to gather comprehensive and relevant information.
- Automatically generate five optimized blog titles, complete with keyword tags, to enhance **SEO performance**.
- **Produce high-quality**, SEO-friendly content based on the generated titles using advanced generative AI and **LLM models**.

F(x) Data Labs PVT LTD. (40-50 employees)

MACHINE LEARNING ENGINEER

Ahmedabad, Gujarat

Dec 2023 – Jan 2022

- Built a strong foundation in **statistics and probability**, specializing in hypothesis testing, A/B testing, regression analysis, and Bayesian testing.
- Expert in translating business requirements into precise **mathematical metrics** for robust analysis.
- Proficient in data **cleaning** and **transformation** to ensure high **data quality** and **reliability**.
- Proficient in a wide range of machine learning techniques, such as **classification, clustering, and Genetic algorithm**.
- Knowledge in state-of-the-art **generative image models**, such as **diffusion models** and controlnet.
- Skilled in constructing Large Language Models (**LLMs**) and fine-tuning them for precise applications, specifically in generating stories and **summarizing** complex narratives.
- Proficiency in deep learning models, including **CNNs, RNNs, and Attention Networks(Transformers)**.
- Familiarity with **MLflow** and **DVC** (Data Version Control) for efficient **data versioning** and model tracking.
- Experience with **Docker** and **Docker Compose** for containerization and deployment.
- Familiar with cloud platforms such as **Azure** and **AWS** for model deployment and scalability.

★ **Project 1 : The GMR Group - MCP Prediction(Time Series Forecasting)**

● **Objective:**

- Developed a predictive model for forecasting the **Market Clearing Price (MCP)** in the **Day-Ahead Market (DAM)** using a Random Forest algorithm, achieving an accuracy of **70% (±6%)**.

● **Key Features & Achievements:**

- Integrated data from diverse sources—including **IEX Market**, weather data, and Google Trends—using Selenium for web scraping, followed by comprehensive data cleaning and transformation.
- Designed a **confidence metrics** model to help traders assess the reliability of MCP forecasts for each time block.
- Employed **Monte Carlo simulation** to demonstrate that our model-based trading strategy achieved an **11% improvement** in profitability compared to random trading strategies.

★ **Project 2: Video Redering Wowsly (Computer Vision)**

● **Objective :**

- Develop an API using **FastAPI** that allows multiple users to upload a **Video** and a corresponding CSV file. Based on the CSV, the system automatically edits the video by overlaying text on specified frames for set durations, stores the edited videos in an **S3 bucket**, and generates a tracking token for users to monitor task progress.

● **Key Features & Achievements:**

- Edited videos by overlaying text on video clips at predetermined positions and durations based on CSV inputs.
- Utilized **Docker & Docker Compose** to manage microservices including FastAPI, Celery, Redis, and Flower.
- Deploy this whole system on **AWS Lambda (AWS,EC2,Lambda)** for scaling up system

Lookman Electroplast Ltd (Secura)

Ahmedabad, Gujarat

DEEP LEARNING INTERN

Aug 2022 – Nov 2022

- Worked on pedetstrain re-identification (Re-ID) task. Person re-identification task using Yolo-v3/Tiny algorithms for object detection and for multiple object tracking we were using the Kalman filter+Deep Sort algorithm.
- **Accuracy is around 80%(+/- 0.04) and It takes around 170-220 ms for processing time per frame(NVIDIA Jetson).**

EDUCATION

L.D.College Of Engineering (Gujarat Technological University)

Ahemadabad, Gujarat

Bachelor of Engineering

May 2022 - July 2018

Major in Computer Engineering ,Cumulative CGPA: 8.62/10

Certifications & Training: Online Course in [Statistics with Python by Michigan University \(Coursera\)](#), [Mathematics for Machine Learning by Imperial College London](#), Computational Thinking and Data Science(MIT).

Languages: Fluent in Gujarati, and Hindi; Conversational Proficiency in English