

NIKHIL SHINGADIYA

DATA SCIENCE SOFTWARE ENGINEER

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

About Me ([Portfolio](#))

I have 2 years experience in the data science field. Accomplished data science software engineer with a passion for delivering valuable data through analytical functions and data retrieval methods. committed to helping companies advance by helping them to develop strategic plans based on predictive modeling and findings. Seeking challenging opportunities to apply analytical-machine learning skills in a collaborative environment.

TECHNOLOGY

Technical Skills: Statistics and Probability, Python, Data Visualization, Exploratory Data Analysis, Machine Learning, Web scraping, Learning, ANN, CNN, Docker, Docker-compose, MySQL, Object-Oriented programming, FastAPI, Familiar with ml-flow, AWS(EC2,S3,ECR,Lambda) and Trading Patterns of the Stock Market.

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

IDE: Visual Studio Code, Jupyter Notebook, Spyder, Google-Colab, And Pycharm. **Version Control :** Git & Github
Operating System: Linux (Ubuntu 20.04 LTS), Windows.

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

WORK EXPERIENCE

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

Ahmedabad, Gujarat

DATA SCIENCE SOFTWARE ENGINEER

Present – Jan 2022

- Advanced knowledge in **statistics**, probability, linear algebra, and calculus.
- Expertise in data visualization, exploratory data analysis (**EDA**), and understanding business requirements.
- Proven ability to convert business requirements into **mathematical metrics** for analysis.
- Skilled in **web scraping** techniques for collecting and processing large-scale data.
- Proficient in data **cleaning** and **transformation** to ensure high **data quality** and **reliability**.
- Strong understanding of advanced **statistical techniques**, hypothesis testing, and regression analysis.
- Experience in **feature extraction**, **feature engineering**, and **feature selection** for model development.
- Familiarity with dimensionality reduction techniques, including Principal Component Analysis (**PCA**) and Linear Discriminant Analysis (**LDA**).
- Proficient in a wide range of machine learning techniques, such as **classification**, **clustering**, and **Genetic algorithm**.
- Knowledgeable in image classification using deep learning models, particularly convolutional neural networks (**CNNs**).
- 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.
- Skilled in querying databases using **SQL**, particularly MySQL, for data retrieval and analysis.

★ **GMR :**

- **The GMR Group - MCP Prediction(Time Series Forecasting)**
- **Goal:** Developed a predictive model for Market Clearing Price (MCP) forecasting in the Day-Ahead Market (DAM) with an accuracy of 70% ($\pm 6\%$) using Random Forest.
- Integrated data from diverse sources, including IEX Market, Weather Data, and G-Trends, utilizing Selenium for scraping and performed comprehensive data cleaning and transformation.
- Designed a confidence metrics ML model to aid traders in assessing the reliability of MCP forecasts for each time block.
- Utilized Monte Carlo simulation to showcase the superior profitability(**11%**) of our model-based trading strategy compared to random trading strategies.

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

- **Goal :** Build the API(**Fast API**) where multiple users can upload one video and one csv file and based on the csv file generate the edited videos and store them to the **s3 bucket** and generate the one token for the user which helps them to track their task progress(In celery,redis and flower).
- Edited videos: based on csv file we put the text on **video clips** (frame at fixed position) for particular duration.
- Used **Docker & Docker-compose** for maintaining micro services (fast api, celery, redis and flower).
- Deploy this whole system on **Kubernets (AWS,EC2,Lambda)** for scaling up system

★ **Discord Bot Project(NLP) :**

- **Goal:** Develop a Discord chatbot with NLP and NLU capabilities to perform the following functions:
- Automatic response generation using predefined answers for predefined questions.
- Enable communication with the support team through email messages.
- Implement group creation functionality for discussions.
- Utilize a Rasa machine learning model to handle user Q&A with a high accuracy rate of 92-95%.

★ **Web Scraping Project(Web Scraping) :**

- **Goal:** In this project, we need to scrap data using selenium and BeautifulSoup like job title, contact number, job description, and location. Which the client gave us.
- First, collects all the job URLs from pagination (63,000 URLs)(fetching_url)
- After collecting those(1 job) URLs, we need to start scraping data that clients want(collecting_data).
- Main.py file is triggering alternatively the above 2 scripts. I made montior.json for handing above two script.

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).**

Personal Project:

[sumquerry-pdf-aisummarizer](#) :Sumquerry is your go-to tool for fast, clear summaries of long documents. It saves you time and makes research interactive. You can even ask specific questions and get smart answers. No more drowning in information – Sumquerry makes learning a breeze!(**Tech Stack** : Langchain, OpenAI,HuggingFace, Streamlit)

EDUCATION

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

Ahemadabad, Gujarat

Bachelor of Engineering

May 2022

Major in Computer Engineering ,Cumulative CGPA: 8.62/10