

NIKHIL SHINGADIYA

DATA SCIENCE SOFTWARE ENGINEER

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

About Me

I have 1.5+ Year 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. Bringing forth a proven track record of analyzing complex data sets and serving as a strong advisor. Seeking challenging opportunities to apply analytical-machine learning skills in a collaborative environment.

EDUCATION

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

Ahemadabad, Gujarat

Bachelor of Engineering

May 2022

Major in Computer Engineering

Cumulative CGPA: 8.62/10.0

Relevant Coursework: Data Science, Machine Learning, Software Engineering, Algorithms, Artificial Intelligence

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:** To predict the MCP values for DAM and also predict confidence value on forecast mcp values.
- **DAM:** One day is divided into 96 time blocks and the users place bids one day in advance.
- Optimize the decision making process.
- **Indian Energy Exchange (IEX)** is a marketplace where electricity is traded. The major markets of focus are Day Ahead Market (DAM) and Real Time Market (RTM).
- A **MCP**(market-clearing) price is the price at which quantity of electricity supplied is equal to quantity demanded, also called the equilibrium price.
- Scraping(Using Selenium) the data from **IEX Market** , **Weather Data** and **G-Trends Data**.
- Combined the all sources of data and perform the data cleaning, transformation and validate data.
- Using **Random Forest** model we got **70**(+/- 6)% accuracy with 10% tolerance of actual mcp values.
- Also, **Design confidence metrics** ml model which helps to trader to decide on which block our forecasting model forecast right mcp values.
- Design the **monte-carlo simulation** to show to the client's profit using our model vs random trading streategy.
- Used **ml-flow** for storing and tracking experiments(around **192** ml model build for 96 block forecast).

Video Redering Wowsly :

- **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)** for scaling up system.
- **Discord bot project :**
 - **Goal:** in this project, I need to make a discord chatbot using **NLP** and **NLU** that can perform the below functionality(5-6 Functionality).
 - **Functionalities:**
 - Ability to Auto Respond with Predefined Answers for Predefined questions.
 - Ability to communicate with us as an email message to our support.
 - Ability to create groups for discussions.
 - Build a rasa machine learning model which handles user Q & A.
 - Rasa Framework gives us testing and validation functionality.
 - We got 92-95% Accuracy of the rasa model that handles Q & A.
- **Web Scraping Project :**
 - **Objective:** 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.
 - **Functionalities:**
 - First, collects all the job URLs from pagination (63,000 URLs)(fetching_url.py)
 - After collecting those(1 job) URLs, we need to start scraping data that clients want(collecting_data.py).
 - Main.py file is triggering alternatively the above 2 scripts. I made montior.json.

Lookman Electroplast Ltd(Secura)

Deep learning Intern

Ahmedabad, Gujarat

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).

UNIVERSITY PROJECTS

RECOMMENDATION SYSTEM

Dec 2020

- Built a recommendation system and it uses a heuristic approach.
- For filling the missing IMDb and Rotten tomatoes values, I used the regression model with 85 % accuracy.
- Enabled users to be recommended movies based on 350+ data points. It's based on content filters.

CUSTOMER CHURN PREDICTION

Jan 2021

- [Web Apps](#)
- Based on 300 Customer Feedback predict whether next time customer food order give us or not (end-to-end ml project).
- I have 55 features inside those features. I need to decide on the 10 most relevant features for prediction.
- Using Random Forest algorithms for prediction, achieved 88 (+/- 0.04) % accuracy, and tried the SVM and xgboost algorithms.

YOUTUBE DASHBOARD

April 2021

- Often Time Users are Confused about How to Choose the Best Channel For My Learning Or Entertainment? So this Dashboard helps those users. I am using 1000+ channel data for Visualization.
- In This Dashboard, I am including below some useful functionality(8-10):
- Top 10 trending videos by date likes/views:- **Bar Chart**.
- Channel-wise descriptive statistics:- **Table**.
- Channel-wise(2) comparison distribution views and likes:- **Boxplot**.
- Longitudinal analysis of video views:- **Timeline Chart**.

TECHNOLOGY

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

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

IDE: Visual Studio Code, Jupyter Notebook, Spyder, Google-Colab, And Pycharm. **Version Control :** Git & Github

Operating System: Linux (Ubuntu 20.04 LTS), Windows.

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

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).

ADDITIONAL ACTIVITIES

Extracurricular Activities

Ahemadabad, Gujarat

Volunteer

April 2019

- Volunteering at the Java Guru Event (Lakshay's talent LDCE) in this event 250+ people participated.
- In this 2 days event, Learned teamwork and how to manage people's time with effective management.

Kabaddi Player at District Level.

Ankolwadi(Gir),Junagadh

Role: All Rounder

March 2016

- District level defender, as well as a rider, and we, reached 3rd position in this tournament. In this event, I learned the **presence of mind, crisis management, understanding the opponent's strategy, and team management.**