Shivam Chaurasia

Nerul , Navi Mumbai
+919326147036 chaurasiashivam721@gmail.com LinkedIn | Leetcode | Github

Technical Skills

Languages : Python, C/C++ , SQL
Data Science: Machine Learning, Deep Learning, , Time-Series Analysis ,Data Manipulation, EDA , Data Warehouse
Tools Frameworks: Pandas, Numpy, Matplotlib, Plotly, Scikit-Learn, FastAPI, Celery ,PowerBI , Tensorflow
Technologies Tools: DSA,OOPS , EC2 , S3 , Docker, mlflow , rabbitmq, Linux, Git ,DBMS , MongoDB

Work Experience

Data Scientist (internship) **Aug 2023 – Mar 2024**
FACLON LABS *Andheri, Marol*

- Developed a No-Code ML tool for data analysis, enabling model training for regression, classification, and time series forecasting. Reduced manual effort by 40% with automated scheduling features, streamlining workflows for 100+ users across the organization. Successfully commercialized and sold the solution to JSW Steel.
- Developed data pipelines for Tata Steel and Grasim, optimizing steam trap efficiency by calculating steam loss and saving 34% of the steam for the organization.
- Enhanced forecasting accuracy by 20% for Mankind Pharma by applying advanced smoothing techniques to real-time sensor data and publishing high-quality streams via the MQTT protocol.
- Developed a data pipeline for UltraTech's solar power plant, aggregating and processing real-time IoT sensor data to optimize solar irradiance forecasting using LSTM model, resulting in a 44% reduction in energy costs.
- **Key Skills**: Python , Machine learning algorithms , Docker , Rabbitmq , celery , gcp , linux , streamlit , LSTM, MQTT,Rapid API , Airflow , Mlflow , Bentoml

Education

Shah and Anchor Kutchhi Engineering College **2020 - 2024**
Bachelor of Engineering in Information Technology (Honours-Data Science) (CGPA - 8.6/10) *Mumbai, MH*

Wilson College **2020**
12th— MSBSHSE— (Percentage: 67.08 / 100) *Mumbai, MH*

Projects

AI-Workbench | *Python , LSTM , Docker , Streamlit , mlflow , bentoml* **July 2023 – ongoing**

- Developed AI-Workbench, a no-code AI platform using Streamlit that automates data preprocessing, feature engineering, and machine learning modeling, including advanced time series forecasting with ARIMA and LSTM models. By integrating Python-based technologies and eliminating the need for coding, the platform empowers non-technical users to perform end-to-end AI tasks effortlessly.
- This innovation democratizes AI, enhancing productivity and enabling faster, data-driven decisions for businesses.

Multi-Disease Diagnostic Platform | *Python, TensorFlow, Scikit-learn, Pandas, NumPy* **Feb 2024 – April 2024**

- Developed a machine learning platform to predict multiple diseases, including Heart Disease, Diabetes, Pneumonia, Brain Tumors, Breast Cancer, and COVID-19, using Logistic Regression, SVM, and CNN models.
- Enhanced diagnostic accuracy and speed, aiding healthcare professionals in early detection and treatment, improving patient outcomes.

Climate Change Analysis of Mumbai | *Pandas, scikit-learn, ML algorithms, Power Bi* **Feb 2023-Mar 2023**

- Conducted an in-depth climate change analysis of Mumbai (Santacruz area) using NOAA data (2000-2023). Applied linear regression and ridge regression, achieving a MAE of 3.97 and R-squared value of 0.73. Developed a comprehensive Power BI dashboard for monthly, weekly, and yearly climate trend visualizations, providing insightful data-driven conclusions.

Energy Consumption Data Analysis | *pandas , sklearn , matplotlib* **Sep 2023 – Sep 2023**

- Performed Data analysis for predicting energy consumption by an IoT device employing exploratory data analysis techniques to uncover trends and patterns
- Applied linear regression, achieving a Mean Absolute Percentage Error (MAPE) of 0.24 and an R-squared score of 0.58 percent.
- Enhanced prediction accuracy using advanced regression models, including DecisionTreeRegressor, AdaBoostRegressor, and Random Forest.