Sahil Birje

+91 8779616204 | sahilbirje24@gmail.com | Kothrud, Pune
https://sahilbirje24.github.io/MyPortfolio/ | https://www.linkedin.com/in/sahil-birje

SUMMARY

Motivated data science student seeking internship or entry-level position to apply theoretical knowledge and hands on experience in machine learning, deep learning, and data analytics. Committed to contributing analytical skills and programming proficiency to solve real-world challenges in an evolving and collaborative environment and stay current with emerging data science trends.

EDUCATION

MSc Data Science & Big Data Analytics, 8.79 CGPA August 2022 – June 2024
MIT WPU, Pune

BCA, Smt.P.D.Hinduja Trust's Institute of Management Studies, 8.63 CGPA August 2019 – August 2022
Mumbai

PROJECTS

ROAD LANE LINE RECOGNITION — AUGUST 2023 – OCTOBER 2023

- Created a dataset of 8750 images by extracting images from video recordings of the car simulation game Forza Horizon 4.
- Labeled 3495 images using the ROBOFLOW platform for the right, left, and middle lane lines and implemented augmentation techniques to generate 8750 images.
- Trained the model using YOLOv8 instance segmentation, achieving an average mAP50 of 0.881 and mAP50-95 of 0.719 for object detection. For mask detection, obtained a mAP50 of 0.867 and mAP50-95 of 0.576.

SALES ANALYSIS/DASHBOARD FOR DISTRIBUTOR — MARCH 2023 - JULY 2023

- Leveraged Python to combine data spread across folders in Excel format into a single true source, comprising 2.1 million records over 3 years.
- Employed Python to clean and transform the data entailing 5 schemas to 1 final schema.
- Analyzed the data and utilized power bi to create a dashboard with three reports, offering insights on revenue generated, product sales, and supplier-retailer performance.

ENERGY LOAD TYPE PREDICTION FOR STEEL INDUSTRIES — FEBRUARY 2023 - APRIL 2023

- Employed machine learning algorithms, including Light Gradient Boosting, XG Boost, and Random Forest Classifier, to construct a model for predicting energy load types.
- Collected data from Kaggle, comprising 35,040 observations from the Korean Steel Industry.
- Emphasized recall for the High Load type class, with the goal of assisting the industry in resource management for optimized energy consumption, achieved an accuracy of 93.42% with the LGBM Classifier.

PEACE SUBJECT ANALYSIS/DASHBOARD — MARCH 2023 - APRIL 2023

- Consolidated 300 survey responses for Peace Subject Faculty to comprehend students' perceptions of the subject.
- Cleaned and processed the 300 records using Python, addressing null values, and mapping numerical values to categories.
- Leveraged Power BI to create a dashboard encompassing 3 reports on students' views about teaching methods, course content, faculty members, interest in learning more about the subject.

SKILLS

- Python (NumPy, Pandas)
- Machine Learning (Scikit-Learn, Seaborn, Matplotlib)
- Deep Learning (TensorFlow, Keras, PyTorch, ANN, CNN, YoloV8)
- Data Analytics
- Data Engineering
- Data Modeling
- Power BI
- MySQL
- Cloud: Azure, AWS

CERTIFICATIONS

- Neural Networks and Deep Learning
  DeepLearning.AI
- MySQL for Data Analytics and Business Intelligence
  Udemy