KAKARA DURGAPRASAD

9493744785 | durgaprasadkakara29@gmail.com | Hyderabad | www.linkedin.com/in/durgaprasadkakara-datascientist | https://github.com/Durgabigdata/DataScience-Excelr-Works.git | https://durgabigdata.github.io/Portfolio/ |

Professional Summary

Aspiring Data Scientist with strong skills in Python, SQL, R, and Tableau, specializing in machine learning, NLP, and deep learning frameworks(ANN,CNN,RNN). Experienced in data preprocessing, EDA, and model deployment, with a track record of delivering impactful, data-driven solutions. Quick learner with a passion for solving business challenges through analytics and innovation.

Work Experience

AI Variant | Hyderabad
Data Science Intern | 07/2024 - 01/2025

- Analyzed large datasets using **Python** and **Pandas** to derive actionable insights.
- Applied **NLP techniques** like TF-IDF to process unstructured text data.
- Built machine learning models (KNN, SVM, XGBoost, LightGBM) for classification tasks.
- Developed and deployed interactive applications with Flask for real-time insights.
- Implemented **item based Collaborative filtering** for a personalized book recommendation system.
- Collaborated with teams to ensure alignment with business objectives and improve model performance.
- Created **data visualizations** and reports to support decision-making.

Renew Private Limited | Jaisalmer
Graduate Engineer Trainee | 06/2023 - 07/2024

- Analyzed large datasets to identify safety risks and trends, providing actionable insights to improve safety measures.
- Collaborated with cross-functional teams to implement solutions based on data analysis, ensuring process optimization and compliance.
- Conducted exploratory data analysis (EDA) on historical safety data to detect patterns and anomalies.

Technical Skills

Python, SQL, Machine Learning, NLP, Forecasting, Deep Learning, Supervised and Unsupervised ML Algorithms, Python Libraries, Data Visualization, Data Cleaning, EDA, Model Evaluation, Streamlit, Flask, Computer Vision

Soft Skills

Problem solving, Analytical thinking,TeamCollaboration, Attention to detail.

Projects

1. Patient's Condition Classification Using Drug Reviews

- **Objective:** Created a machine learning model to analyze patient reviews and classify medical conditions and to recommend the most effective drugs.
- **Approach:** Applied **Natural Language Processing (NLP)** techniques, including **TF-IDF**, to preprocess and vectorize the text data. Built several classification models such as **Random Forest,KNN,XGBM,LGBM** and **SVM** to predict patient conditions from the reviews.
- **Outcome:** Deployed the final model using **Flask**, providing a user-friendly interface for real-time condition classification and drug recommendations based on patient feedback.
  Tools & Technologies: Python, Pandas, Scikit-learn, TF-IDF, NLP, Flask

2. Book Recommendation System

- **Objective:** Developed a personalized **book recommendation system** to suggest books based on user preferences and ratings, enhancing user experience.
- **Approach:** Implemented **item based collaborative filtering** to recommend books by analyzing both user ratings and book features. The system optimized recommendations based on explicit ratings and implicit user behavior .

- **Outcome:** Improved user engagement by providing accurate and relevant book suggestions, increasing satisfaction and user retention.
  Tools & Technologies: Python, Pandas, Collaborative Filtering, User-to-Item Filtering, Scikit-learn

3. Time Series Forecasting of CO2 Emissions
   - **Objective:** Developed a time series forecasting model to predict **CO2 emissions**, assisting organizations in meeting environmental regulations and minimizing their carbon footprint.
   - **Approach:** Performed **Exploratory Data Analysis (EDA)** to understand patterns, trends, and seasonality in emission data. Built and fine-tuned multiple forecasting models, including **ARIMA**, **SARIMA**, and **Exponential Smoothing**, with **ARIMA** emerging as the best-performing model for short-term forecasts.
   - **Outcome:** Deployed the model via **Streamlit** for real-time forecasting, enabling the organization to monitor emissions and ensure compliance with environmental standards.
     Tools & Technologies: Python, Pandas, ARIMA, SARIMA, Matplotlib, Streamlit

Certificates

Data Science Certification - Excelr 2024

Education

Rajiv Gandhi University of Knowledge and Technologies | Srikakulam
Mechanical Engineering | 05/2023
Achieved academic excellence with a stellar **92.9%** in graduation.

Languages

English, Hindi, Telugu