KAKARA DURGAPRASAD

9493744785 | durgaprasadkakara29@gmail.com | Hyderabad | Linkedin | Portfolio | Github |

Professional Summary

Detail-oriented Data Science professional with 1+ year of hands-on experience in data handling, machine learning, and model deployment. Started as a Graduate Engineer Trainee in EHS, where working with data sparked a shift into data science. Skilled in Python, ML, NLP, time series forecasting, ANN, and tools like Scikit-learn, TensorFlow, FastAPI, and MLflow. Completed a data science internship and currently leading freelance projects end-to-end. Seeking a full-time role to apply and grow in real-world, data-driven problem solving.

Work Experience

Freelance Data Science Projects | Remote
Self-driven | 04/2025 - Present

- Designed and implemented **end-to-end machine learning projects**, from data preprocessing to deployment using **FastAPI**
- Built and trained **deep learning models (ANN)** and applied advanced **time series forecasting techniques** for real-world prediction tasks
- Integrated **MLflow** into projects for model tracking, version control, and performance monitoring
- Worked on real-world use cases in **NLP**, recommendation systems, and forecasting, while actively deepening knowledge of core data science topics

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

- Cleaned and preprocessed real-world datasets using **Python (Pandas, NumPy)** for machine learning tasks
- Performed **exploratory data analysis (EDA)** and visualized insights using **matplotlib**, **seaborn**
- Built and evaluated ML models for **classification and regression** using **Scikit-learn** and **XGBoost**
- Applied **natural language processing (NLP)** techniques and deployed models via **Streamlit** and **Flask** for demo applications

Renew Private Limited | Jaisalmer
Graduate Engineer Trainee - EHS [Data Focussed] | 06/2023 - 07/2024

- Collected, cleaned, and analyzed EHS data including incident logs, safety inspections, and audit reports using **Python (Pandas)**
- Generated **monthly safety reports** and visualized patterns in incident types, root causes, and near misses using **matplotlib** and **seaborn**
- Contributed to internal presentations by preparing **data summaries and visual reports** for safety performance reviews
- Supported EHS audits by tracking **environmental KPIs** such as emissions, waste disposal, and energy usage, contributing to compliance reviews

Technical Skills

Python, SQL, Machine Learning, Deep Learning, NLP, Forecasting, Supervised and Unsupervised ML Algorithms, Python Libraries, Data Cleaning, EDA, Data Visualization, Model Evaluation, Fastapi, Flask, Streamlit, MLflow, Pytorch, TensorFlow, Hugging Face Transformers, Fine-Tuning Pre-trained Models, LLMs, Bert & Gpt architectures, OpenCv, Computer Vision, Object detection, Pandas, Numpy, Seaborn, Matplotlib

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. Agent Auto-Assignment Engine for Ticket Routing

- **Objective:** Developed a machine learning engine to automatically assign incoming support tickets to the most suitable agents based on contextual and performance-based features.
- **Approach:** Built and evaluated multiple classification models using historical ticket and agent performance data. Features included ticket priority, sentiment, agent load index, skills, and resolution history. The best-performing model, **XGBoost Classifier**, was deployed via **FastAPI** and tracked using **MLflow**.
- **Outcome:** Reduced manual ticket routing, improved response time, and enhanced overall agent productivity through intelligent, data-driven assignment.

Tools & Technologies: Python, Scikit-learn, XGBoost, MLflow, FastAPI, Pandas

3. Sentiment Analysis Using Transformer Models

- **Objective:** Built a sentiment analysis model to classify text (reviews, feedback, and general content) as positive or negative, supporting better understanding of customer opinions.
- **Approach:** Fine-tuned the **DistilBERT** model on the IMDB dataset using Hugging Face Transformers, with data preprocessing, tokenization, and model evaluation. Developed a **Streamlit** web app for real-time sentiment classification across various text types.
- **Outcome:** Achieved high accuracy in distinguishing sentiment in diverse texts, including reviews and social media posts. Deployed for real-time predictions via Streamlit.

Tools & Technologies: Python, Hugging Face Transformers, DistilBERT, NLP, Streamlit,Pytorch.

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