Tara Chand

📞 +91 7895293563 | ✉ tarac.1998es@gmail.com | 🔗 Linkedin | 🌐 Github | 📍 Noida,IN

EDUCATION

Indian Institute of Technology Guwahati   07/2021 - 06/2023
Master in Robotics and Artificial Intelligence   CGPA - 8.69

PROFESSIONAL EXPERIENCE

Axtria Pvt. Ltd.   07/2023 - 04/2025
Data Analyst   Noida, IN

AutoEDA with GenAI

- led the development of an automated Exploratory Data Analysis (AutoEDA) tool in **PySpark** that allows multiple datasets to be processed concurrently with a single YAML configuration, optimising workflows and enhancing analysis effectiveness with insights from GenAI.
- Integrated dynamic HTML report generation with individual plot summaries and dataset overviews using **GenAI** to deliver actionable, data-driven insights for enhanced decision-making.   '
- Designed and developed an interactive chatbot powered by **model parameters tuned LLM** and advanced **prompt engineering**, integrated with a **React-based** UI and **Databricks** backend, enabling users to explore datasets with accurate, timely, and context-aware insights.
- Delivered 40% cost savings for a major Axtria client by reducing the resources and time required to perform EDA on new datasets.

Shippable Code Accelerator

- Enhanced **InsightMax** by integrating AutoDL capabilities using advanced deep learning models such as NAM and TabNet to predict the next best action in the Life Sciences domain.
- Designed and implemented an end-to-end deep learning pipeline supporting both **regression** and **classification** tasks, with backend integration to fetch APLD data from **AWS S3** via a dedicated feature store.
- Connected the pipeline to **MLflow** for model registry, evaluation, and experiment tracking, and incorporated optimization (**Optuna**) and interpretability (**SHAP, LIME, Permutation Importance**) techniques to facilitate transparent and effective model analysis.

PROJECTS

Flight Fare Prediction Using Multiple Linear Regression

- Performed extensive data preprocessing and exploratory data analysis to identify patterns and outliers, and applied data visualization for feature exploration.
- Implemented **feature selection** techniques including **ExtraTreesRegressor, AIC/BIC**, and **hypothesis testing** to enhance model performance.
- Trained and evaluated multiple regression models **(Linear, Ridge, Lasso)**, leveraging **RandomizedSearchCV** for **hyperparameter tuning** and optimization.

Face Recognition Based Attendance Marking System

- Built an automated attendance system using face recognition that logs user presence into a MySQL database upon detection in live camera feed.
- Employed a pre-trained VGGFace model combined with HOG and facial landmark estimation for robust recognition.
- Integrated a back-end database to store facial embeddings and attendance logs, enabling real-time identification and tracking.

TECHNICAL SKILLS

DS/AI Courses : Machine Learning, Deep Learning, Statistics, Data Structures and Algorithms,Artificial Intelligence and Probability

Technologies : OpenAI, Langchain, Gemini, Hugging Face, RAG, Prompt Engineering, LoRA, qLoRA, GPT, BERT **SQL**, **Python**, Random Forest **(Bagging)**, Decision Trees, Linear Regression, Logistic Regression, SVMs, GBM-**(XgBoost, CatBoost, AdaBoost)**, Clustering methods **(KMeans)**, PySpark, Text processing, Vector Database (ChromaDB), LLMOps, MLOps

Libraries : Scikit-Learn, Pandas, Matplotlib, Numpy, Seaborn, DeepEval, Tensorflow, PyTorch, Keras

Development Tools : **GitHub**, Bitbucket, Visual Studio Code, Databricks, **AWS SageMaker and S3**, **Dataiku**, Streamlit, React, Jupyter, MLFlow

ACHIEVEMENTS

Tab Hack 2.0 Secured the seventh rank among 82 teams, An ML Hackathon conducted by IITG.ai

Beyond Data Analysis Secured the seventh rank in the Data Science Hackathon conducted by Techniche among 111 teams