Yash Makwana

OBJECTIVE

Certified Data Scientist experienced in transforming business requirements into analytical models, designing algorithms, and building predictive models. Skilled in managing data pipelines, optimizing machine learning models, and improving system efficiency. Seeking to leverage my expertise in a dynamic team environment to drive impactful data-driven solutions

+91 7805046767

yashmakwana40@gmail.com

Indiranagar, Bangalore

https://yashmakwana.netlify.app/

SKILLS

PYTHON NUMPY PANDAS MYSQL STATISTICS MACHINE LEARING REGRESSION CLUSTERING SKLEARN CLASSIFICATION PREDICTIVE MODELING DEEP LEARNING ANN RNN CNN KERAS TENSORFLOW GENSIM SPACY BERT FAISS RAG STREAMLIT

EDUCATION

Master of Technology

National Institute of Technology, Durgapur,West Bengal

09/2020 - 05/2022, **CGPA- 8.25**

Bachelor Of Engineering

Sagar Institute of science and technology, Bhopal, Madhya Pradesh

08/2015 - 06/2019, **CGPA- 7.30**

EXPERIENCE

ANALYST

ALSTOM

2022 - Present

Design to Cost (D2C)

Goal: Implement a cost-efficient predictive model pipeline

- Managed the D2C project using Dataiku, focusing on the end-to-end data pipeline and predictive modeling.
- Processed and cleaned data received from the front-end, applying statistical methods like ANOVA for thorough analysis.
- Developed and fine-tuned machine learning models by experimenting with various algorithms and optimizing hyperparameters.
- Developed a prediction model that achieved high accuracy, tailored to meet the specific forecasting needs of the project.

Smart Trouble Shooting (STS)

Goal: The STS project aimed to create an automated system for querying and retrieving solutions from historical data to improve problem resolution efficiency.

- Worked on the STS project, creating a sophisticated data retrieval system using NLP techniques to analyse and match user queries with past issues, thereby improving the accuracy and efficiency of problem resolution.
- Optimized code to enhance time and space efficiency, resulting in significant improvements in performance and resource management.
- Designed and implemented a prediction pipeline to enable seamless data processing and accurate model predictions.
- Adapted the output format to align with business needs and presented the model for final deployment approval, ensuring compliance with all business and technical requirements.

CERTIFICATES

- OCI 2024 Generative AI certified Professional.
- Dataiku: Core Designer Certificate.
- Neural Network and Deep Learning,, Coursera..
- Supervised Machine Learning: Regression and Classification, DeepLearning.ai, Coursera.
- Python for Data Science, AI and Development, IBM, Coursera.

DEVELOPERTOOLS

DATAIKU JUPYTER VS CODE

ACHIEVEMENTS

- Collaborated with a team to create dynamic Power BI dashboards for reporting purposes, resulting in more efficient and effective reporting.
- Managed finances for the team and oversaw budgets for a year.
- Analyzed financial data to ensure alignment with team objectives.
- Experience in working with ARIBA and INSPIRE, which are finance tools utilized for raising Goods Receipt Notes (GRNs) and processing Purchase Requisitions (PRs) and Purchase Orders (POs).

LANGUAGE

ENGLISH

HINDI

Tender Benchmark(TBT)

Goal: The TBT project was designed to predict manufacturing costs using client specifications, estimating production time and applying logical cost calculations.

- Developed a prediction pipeline using multiple regression models (linear, SVM, polynomial, XGBoost, random forest) to predict hours and calculate costs. Selected the optimal model based on R^2 score, MAPE, and MAE.
- Responsible for model building and developing a prediction pipeline for accurate time estimation.
- Tuned hyperparameters to optimize the prediction pipeline, selecting the most accurate model for each target scenario.
- Implemented a cost determination pipeline using a predefined hard logic methodology following the time prediction.

PERSONALPROJECTS

Automated Data Processing and Query Generation System.

Goal: Automate data processing and query generation using Dataiku and advanced AI to achieve accurate and efficient results

- Uploaded and synchronized datasets into a blob file in Dataiku for enhanced data processing capabilities.
- Embedded metadata into the certain column on the blob file to ensure detailed data representation.
- Integrated with a knowledge base to perform LLM operations, enabling advanced data analysis and insights..
- Implemented AI-driven query generation using Dataiku AI, leveraging LLM for efficient and accurate results based on provided data.

RESEARCHEXPERIENCE

Goal: Non-Redundant Cover Set Formation to Improve Full Coverage in WSNs.(Sept 2021 - Apr 2022)

- Created a novel algorithm to optimize Wireless Sensor Networks, achieving longer network lifespan and efficient coverage through non- redundant sensor distribution.
- Implemented the algorithm using Python, utilizing libraries like Numpy and Pandas for efficient computation and practical deployment.
- **Significant Impact:** Worked with experts from different fields, validated the algorithm thoroughly, continuously learned new methods, and helped advance WSN research and applications by clearly sharing the results.