DISHANT CHOUHAN

chouhandishant921@gmail.com | +91-9713996565

LinkedIn |Portfolio| GitHub

BRIEF SUMMARY

"Passionate data enthusiast with a knack for turning complex datasets into actionable insights. Experienced in data analysis, machine learning, and statistical modeling. Thrives in dynamic environments, leveraging data-driven strategies to drive impactful decisions. Ready to tackle new challenges and unlock the potential of data."

EDUCATION

Medi-Caps University   Indore, M.P.
Bachelor of Technology Computer Science and Engineering   2021 - 2025
CGPA: 7.93/10
Shri Sai Academy   Mhow, M.P.
Class 12th | CBSE | Percentage: 74.2%   2021
Shri Sai Academy   Mhow, M.P.
Class 10th | CBSE | Percentage: 78.6%   2019

SKILLS

Language:- Python, C++, SQL, OOP
Machine Learning & Deep Learning:- Scikit-learn, Pandas, NumPy , TensorFlow ,Keras, NLP,Computer Vision, Algorithms
Databases :- MySQL, NoSQL, PostgreSQL
Data Analysis: Tableau, Exploratory Data Analysis, Dashboard Development
Other Skills:- Big Data, Docker, Flask, Data Engineering(Basic), Teamwork

TRAINING/ WORK EXPERIENCE

INDUSTRIAL TRAINING | ALTERYX   June 24- July 24

- Participated in industrial training focused on data transformation techniques, enhancing skills in efficiently manipulating and analyzing complex datasets.
- Gained experience in creating basic data workflows, which improved the efficiency of data processing tasks using Alteryx.

PROJECTS

Crop Yield Prediction & Recommendation System GitHub

- Developed a machine learning-based Crop Recommendation and Yield Prediction System capable of predicting yields and recommending 105 crops with 88% accuracy, leveraging SMOTE to address dataset imbalance.
- Utilized Gradient Boosting for crop recommendation and Random Forest for yield prediction, analyzing soil properties, climate conditions, and agronomic practices to optimize agricultural productivity.

ASKmYDocs – LLM-Powered PDF Question Answering System GitHub

- Built a document Q&A chatbot using LangChain, FAISS, and Mistral-7B-Instruct-v0.3 (Hugging Face) to generate context-aware answers from uploaded PDFs with source attribution.
- Implemented a semantic search pipeline using MiniLM-L6-v2 embeddings and recursive text chunking for fast and accurate retrieval across document content.
- Enhanced UX with features like dynamic follow-up question suggestions, chat history memory, keyword highlighting, and vector store management (upload/delete/reset).

Musculoskeletal Abnormalities Detection GitHub

- Designed and implemented a deep learning system for detecting musculoskeletal abnormalities in X-ray images, focusing on body parts like the wrist, finger, and forearm, using TensorFlow / Keras for accurate model inference.
- Developed a Flask-based web application to provide medical professionals with a user-friendly interface for real-time diagnostic analysis, enhancing accessibility and efficiency in clinical workflows.

Real-Time Data Streaming | End-to-End Data Engineering Project GitHub

- Designed and implemented a full-scale data engineering pipeline using Apache Airflow, Kafka, Spark, and Cassandra for seamless data ingestion, processing, and storage.
- Integrated data ingestion from randomuser.me API and orchestrated workflows with Airflow, enabling real-time data processing with Kafka and Spark.

ACHIEVEMENTS

- Amazon ML Challenge: Ranked 357th out of 75,000 participants in a national-level, two-stage competition.

CERTIFICATION

Google Data Analytics Specialization(Coursera) , IBM Data Engineering(Coursera) , CISCO Python Essentials