DISHANT CHOUHAN

DONGARGAON,MHOW(M.P.) | P: +91 9713996565

Data Engineer| Data Analyst | Machine Learning | Data Science

chouhandishant921@gmail.com |www.linkedin.com/in/dishant-chouhan |https://github.com/Dishant-Chouhan

BRIEF SUMMARY

"A passionate data engineer with expertise in transforming complex data into actionable insights. Skilled in Python and Apache Spark for big data processing and machine learning. Proficient in managing relational SQL and NoSQL databases, and experienced in building data pipelines with Airflow and Kafka. Adept at creating visualizations with Tableau and Power BI, and optimizing ETL workflows. Committed to solving complex data challenges and delivering innovative solutions."

EDUCATION

MEDICAPS UNIVERSITY   RAO,M.P.
Bachelor of Technology : Computer Science Engineering - CS(DATA SCIENCE)   Expected May 2025

SHRI SAI ACADEMY   MHOW, M.P.
HIGH SCHOOL – PCM+CS   APR 2019 - Jul 2021

SKILLS

Languages: Python, R, C++, SQL, Flask, OOP
Database Management: PostgreSQL, MongoDB, MySQL, IBM Db2, Database Administration (DBA)
Big Data Technologies: Hadoop, HIVE, Cassandra Apache Spark (Spark SQL, Spark ML, Spark Streaming),Docker
ETL & Data Pipelines: Apache Airflow, Apache Kafka
Machine Learning: Data Mining, Scikit-learn, Pandas, NumPy , TensorFlow
Data Analysis: Tableau, Power BI, Matplotlib, Seaborn, Exploratory Data Analysis, Dashboard Development
Business Skills: Decision Making, Problem Solving, Presentation Skills, Effective Communications, Critical Thinking

TRAINING/ADDITIONAL 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 AND RECOMMENDATION SYSTEM
Achievements:
- **Increased Model Accuracy:** Enhanced accuracy from 20% to 88% by applying SMOTE to address dataset imbalance.
- **Comprehensive Crop Yield Prediction and Recommendation:** Capable of predicting yields and recommending 105 crops, surpassing other models.
- **Integration:** Developed and integrated the solution using Flask, unifying both models into a seamless web application for an improved user experience.

Real-Time Data Streaming | End-to-End Data Engineering Project
Achievements:
- Built a comprehensive data engineering pipeline utilizing Apache Airflow, Kafka, Spark, and Cassandra.
- Integrated data ingestion from randomuser.me API, orchestrated workflows with Airflow, and processed real-time data streams with Kafka and Spark.
- Stored processed data in Cassandra and PostgreSQL, with the entire setup containerized using Docker for scalability and ease of deployment.

Case Study: COVID-19 Data Analysis Across Multiple Platforms

Achievements:

- Created interpretable visualizations with Tableau to illustrate pandemic trends.
- Performed SQL queries for data extraction and manipulation, ensuring robust preprocessing and analysis.
- Analyzed COVID-19 records in Excel, focusing on death and vaccination trends.
- Demonstrated proficiency in data visualization, SQL querying, and managing large datasets to derive high-level insights during the global pandemic.

Case Study: Data Analysis for Sales Dashboard

Achievements:

- Developed a comprehensive sales dashboard in MS Excel to track and analyze performance across furniture, office supplies, and technology categories.
- Leveraged advanced Excel functionalities including pivot tables, charts, and slicers to visually represent sales data and trends effectively. Implemented formulas and conditional formatting to automate data analysis and highlight key insights such as top-selling products, revenue trends, and regional sales distribution.
- This project honed my skills in data manipulation, visualization, and dashboard creation in Excel, showcasing my ability to transform raw data into actionable business intelligence.

CERTIFICATION

Data Visualization: Empowering Business with Effective Insights Job Simulation   **Jan 2024 - Jan 2024**
Tata Forage

Participated in TATA's Data Visualization job simulation, focusing on empowering businesses with effective insights through data visualization techniques.
Utilized advanced skills in Tableau and Excel to analyze complex datasets and provide valuable insights for business decisions.
Demonstrated proficiency in presentation skills, effectively communicating insights to stakeholders.

IBM Data Engineering Professional Certificate   **2024**
IBM(Coursera)

Mastered data engineering concepts including big data technologies (Hadoop, Spark), ETL pipeline development (Airflow, Kafka), and database management (PostgreSQL, MongoDB). Gained hands-on experience in data analysis, machine learning, and visualization using tools like Tableau and Power BI, demonstrated through a capstone project that applied skills to real-world data challenges.

Google Data Analytics Specialization   **2023**
Google(Coursera)

Completed Google's Data Analytics Specialization on Coursera, mastering data cleaning, analysis, and visualization using Google Sheets, SQL, and Tableau. Proficient in real-world data analysis through hands-on projects, enhancing knowledge of key analytics concepts applicable across diverse industries.

Python Essential 1 and 2 – Cisco   **2023**
Cisco

Mastered Python Essential 1 & 2 at Cisco, proficient in analytical thinking, procedural programming, and OOP. Acquired practical experience with Python packages.

STRENGHTS

- Goal Oriented
- Problem Solving
- Attention to Details
- Time Management
- Communication Skills

HOBBIES & INTERESTS

- Playing video games
- Watching movies
- Collecting old coins
- Stock Trading
- Technology Enthusiast

LANGUAGES

- HINDI
- ENGLISH