TANISHA SINHA

tanisha02sinha@gmail.com| +91-9691980188 | https://github.com/tanisshaaa |WWW.LINKEDIN.COM/IN/TANISHASINHAA

Education

Chhattisgarh Swami Vivekanand Technical University, University Teaching Department, India   **2021 - 2025**
- B.Tech-Hon's CSE (Data Science) | CGPA: **8.08(Present)**

Narayana High School, India   **2020 - 2021**
- *AISSCE (Class XII), Percentage: 69.2%:*

St. Francis Senior Secondary School, India   **2018 - 2019**
- *AISSE (Class X), Percentage: 64.8%:*

Skills

C++| Python |Machine Learning |SQL |Data Visualisation | Power BI | Excel | Git

Work Experience

IIIT Naya Raipur | Research Intern   **Jan'24 - Jun'24**
- Conducted in-depth **Performance analysis of Apache Spark job schedulers** for big data processing, focusing on optimizing resource utilization and execution efficiency.
- Published findings in a research paper, contributing to advancements in the field of **big data analytics**.

Projects

Dairy Product Analysis   **Jan'25**
Analyzed a dairy dataset using Power BI to extract insights on farm efficiency, sales performance, and product distribution. Focused on revenue trends, customer preferences, and the impact of shelf life and storage conditions.
Tech Stack: Power BI Desktop, DAX, Power Query

Predictive Maintenance for Industrial Machinery **Live Demo**   **Jan'25**
This project is a Streamlit-based web application that leverages machine learning models to predict potential failures in industrial machinery using historical sensor data. It helps manufacturers and engineers take proactive maintenance actions, reducing downtime and operational costs.
Tech Stack: Python, MySQL, Machine Learning, Streamlit, Matplotlib, pandas, NumPy, Scikit-Learn, Seaborn

Performance Analysis of Apache Spark Job Schedulers for Big Data   **Jun'24**
Conducted a comparative analysis of Apache Spark's job schedulers, including FIFO, Fair, and Capacity. Evaluated performance based on execution time, resource utilization, and workload distribution under varying conditions. The study provides insights to optimize scheduler selection for efficient big data processing.
Tech Stack: Hadoop, Apache Spark, Big Data Analytics, Scheduling, Big Data Processing, PySpark

Academic and Extracurricular Achievements

- Paper on performance analysis of Apache Spark job scheduler accepted for publication at TENSYMP 2024.
- Event Coordinator at TPO, overseeing placement drives and career development events, ensuring seamless execution.

Certifications

- SQL Basic from Hackerrank
- Data Visualisation from Kaggle