DIPANSU PRUSTY

📞 +91 8328979860   ✉ prustydipansu2k16@gmail.com   in LinkedIn   ⚲ GitHub   ⦾ Portfolio

SUMMARY

Data Analyst with a strong foundation in statistics and hands-on experience in SQL, Python, Excel, Power BI, and Tableau. Skilled in cleaning and analyzing data, building dashboards, and applying machine learning models to solve business problems. Experienced in healthcare operations, sales analytics, and customer behavior analysis through real-world projects. Looking to contribute to data-driven decision-making in an analyst role.

PROJECT

Hospital Operations & Patient Flow Analytics

- Analyzed **9,000+ patient visit records** using SQL to assess wait times, patient flow, and department workload
- Defined **operational KPIs** including average wait time, patient volume, and referral vs walk-in distribution.
- Identified **peak-hour congestion and high-load departments**, highlighting opportunities to reduce wait times by **15–20%**
- Built **Power BI dashboards** to visualize bottlenecks and support data-driven hospital operations decisions

Lead Scoring Prediction Model

- Examined **5,000+ leads** to build classification workflows, achieving **85% accuracy**.
- Decreased **false positives by 22%** through **validation metrics and model tuning**.
- Designed Tableau dashboards that shortened **reporting cycles by 40%**.
- Strengthened **lead prioritization, resulting in a 10-12% increase** in higher-quality conversions.

Cellphone Market Analysis & Sentiment Study

- Processed over **10,000 customer** reviews using **NLP** to **achieve a 92% sentiment classification accuracy**.
- Identified pricing and feature gaps offering a **15-20% expansion opportunity**.
- Delivered interactive dashboards improving **insight communication efficiency by 35%.**

E-Commerce Dark Pattern Classifier using Machine Learning

- Built an **end-to-end NLP pipeline** leveraging **NLTK** for advanced **text cleaning, tokenization, and lemmatization** to detect deceptive user interface patterns.
- Developed a **Linear SVM model** that **achieved 94% test accuracy and a 0.95 F1-score**, outperforming baseline **Naive Bayes and Logistic Regression models.**
- Deployed model explains ability tools (**ELI5**) to visualize feature weights, providing actionable insights into the linguistic characteristics of 5 distinct dark pattern categories

SKILLS

Programming: Python, SQL, Excel
Databases: MySQL, PostgreSQL
Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Statsmodels, Plotly, SciPy, NLTK
Machine Learning: Linear & Logistic Regression, Classification, Decision Tree, Random Forest
Visualization: Tableau, Power BI
Core Skills: Data Cleaning, EDA, Business Intelligence, Dashboarding, Statistical Analysis, Hypothesis Testing, Feature Engineering, Predictive Analytics, KPI Tracking, Reporting

EDUCATION

Post Graduate Diploma in Data Science

International Institute of Information Technology, Bangalore • Apr 2023 - May 2024
- A Grade (3.6/4)
- Coursework: Data Analytics, Machine Learning, Statistics, Data Warehouse, Project Management
- Technologies: Python, SQL, AWS, Tableau, Power BI, Advanced Excel.

Master of Science (Statistics)

Sambalpur University • Aug 2020 - Sep 2022
- CGPA: 8.02

Bachelor of Science (Statistics)

Ispat Autonomous College • Apr 2017 - Jul 2020
- CGPA: 8.2