RISHITA PRIYADARSHINI SARAF

rishitasarafp@gmail.com | P: +91-9490463841 | www.linkedin.com/in/rishita-priyadarshini-saraf | github.com/Rishita-P-Saraf

EDUCATION

Vellore Institute of Technology, Bhopal 2022- 2026
BTech In Computer Science and engineering CGPA: 8.71 (expected)

Narayana Junior College 2020-2022
Narayana Junior College, Narayanguda SSC (CLASS XII) Aggregate: 92.8%

High School 2010-2020
St. Joseph's School, Habsiguda ICSE (Class X) Aggregate: 96.4%

TECHNICAL SKILLS

- **Programming Languages:** Python, C++
- **Databases:** MySQL
- **Technical Skills:** Data Analysis, Data Visualization, Microsoft Excel, Power BI, SQL, Machine Learning Algorithms, Deep Learning, NLP, Image Processing, Generative AI
- **Tools & Frameworks:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, OpenCV

WORK EXPERIENCE

Data Analyst Intern - NullClass (Remote) Jan 2025 – Feb 2025
- Analyzed Google Play Store app reviews and performed sentiment analysis to extract user behavior insights.
- Cleaned and preprocessed large datasets using Python and Pandas for accurate trend identification.
- Built and deployed an interactive HTML dashboard using Plotly to visualize sentiment and feature feedback.
- Tech Stack: Python, Pandas, NumPy, Plotly, Jupyter Notebook, HTML

Data Science Intern - Cognifyz Technologies (Remote) Dec 2024 – Jan 2025
- Built a linear regression model to predict restaurant ratings based on customer data and key features.
- Performed data cleaning, feature selection, and exploratory analysis to enhance model accuracy.
- Visualized patterns and correlations using Seaborn and Matplotlib to support business insights.
- Tech Stack: Python, Pandas, Scikit-learn, Seaborn, Matplotlib, Jupyter Notebook

PROJECTS

SoEfficient: ML-Based Solar Panel Performance Forecasting June 2025
- Engineered and preprocessed large-scale sensor datasets by imputing missing values, encoding categorical variables, and generating advanced features to improve model interpretability and predictive accuracy.
- Developed and evaluated regression models (XGBoost, Ridge, RidgeCV) for solar panel efficiency prediction, achieving a validation RMSE of ~0.10 and conducting hyperparameter tuning using GridSearchCV to optimize performance.
- Tech Stack: Python, Pandas, NumPy, Scikit-learn, XGBoost, GridSearchCV, Jupyter Notebook

E-Commerce Customer Segmentation Using Clustering
- Engineered an e-commerce customer segmentation solution using Python, scikit-learn, and pandas to analyze user behavior.
- Implemented and compared K-Means, Hierarchical, and DBSCAN clustering, visualizing segments with matplotlib and seaborn.
- Translated model insights into actionable strategies, driving targeted marketing and enhancing customer engagement.
- Tech Stack: Python, Pandas, NumPy, Scikit-learn, K-Means, Hierarchical, and DBSCAN clustering, Jupyter Notebook

ADDITIONAL INFORMATION

Certifications:
- Oracle Data Science Professional Certificate
- IBM GEN AI Using IBM Watsonx certificate

Achievements: Ranked 78th in the Zelestra Hackathon on HackerEarth for building a solar panel efficiency prediction model using XGBoost, achieving 89.88% accuracy.