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: 8.77 (expected)

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

UNIVERSITY PROJECTS

Brain Tumor Detection using CNN May, 2025

- Built and trained a 7-layer CNN achieving 100% training accuracy and 96.8% validation accuracy for MRI-based tumor detection.
- Processed MRI images, performed augmentation and normalization, and visualized feature maps for interpretability.
- Improved model performance by reducing overfitting by 12% using dropout and data balancing techniques.
- Tech Stack - Python, PyTorch, OpenCV, NumPy, Matplotlib, Jupyter Notebook

SoEfficient: ML-Based Solar Panel Performance Forecasting June, 2025

- Engineered and cleaned 1,20,000 sensor records, imputing missing values and encoding categorical features to enhance interpretability.
- Trained and evaluated 3 regression models (XGBoost, Ridge, RidgeCV) achieving RMSE ≈ 0.10
- Applied GridSearchCV (20+ parameter combinations) to optimize hyperparameters and boost model accuracy by 15%**.**
- Tech Stack: Python, Pandas, NumPy, Scikit-learn, XGBoost, GridSearchCV, Jupyter Notebook

ClusterCart: Multi-Model Customer Segmentation for E-Commerce Analytics June, 2025

- Segmented 5,000+ customer records using K-Means, Hierarchical, and DBSCAN, identifying 3–5 actionable segments for marketing.
- Tuned clustering parameters (silhouette score improved from 0.41 → 0.63) to enhance separation between clusters.
- Visualized segmentation patterns through scatter plots and dendrograms**,** improving stakeholder interpretability.
- Tech Stack: Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Jupyter Notebook

WORK EXPERIENCE

Data Analyst Intern - NullClass (Remote) Jan 2025 – Feb 2025

- Analyzed 10,000+ Google Play Store app reviews and created an interactive dashboard using Plotly with 4+ dynamic charts.
- Extracted sentiment trends (positive, neutral, negative) achieving 85% accuracy using text preprocessing and polarity scoring.
- Improved client insight reporting speed by **30%** through automated data pipelines.
- Tools: Python, Pandas, Plotly, Matplotlib, Seaborn, Jupyter Notebook

Data Science Intern - Cognifyz Technologies (Remote) Dec 2024 – Jan 2025

- Built a linear regression model predicting restaurant ratings with $R^2 = 0.84$.
- Processed 1,500+ data records, handled missing values, encoded categorical features, and visualized insights using Seaborn.
- Optimized feature set, reducing model error by 18% and improving interpretability through correlation heatmaps.
- Tools: Python, Pandas, Matplotlib, Seaborn, Scikit-learn

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.