# 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: 8.73 (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**

- Tech Stack - Python, PyTorch, OpenCV, NumPy, Matplotlib, Jupyter Notebook
- Built and trained a custom CNN model in PyTorch for brain tumor detection from MRI scans, applying deep learning techniques to achieve 100% training accuracy.
- Applied data preprocessing, designed a custom Dataset class with train/validation split, and utilized model interpretability through feature map visualization and confusion matrix analysis.

**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.

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

- Segmented 5,000+ e-commerce users using K-Means, Hierarchical and DBSCAN to identify high spenders, price-sensitive buyers and other groups.
- Preprocessed data with scaling and encoding, and tuned clustering parameters to extract 3–5 actionable segments for targeted marketing.
- Visualized results using scatter plots and dendrograms to compare algorithm effectiveness.
- Tech Stack: Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Jupyter Notebook

## WORK EXPERIENCE

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

- Analyzed Google Play Store app reviews and created an interactive HTML dashboard using Plotly.
- Performed data preprocessing and extracted sentiment trends to guide app improvements.
- Tools: Python, Pandas, Plotly, Matplotlib, Seaborn, Jupyter Notebook

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

- Developed a linear regression model to predict restaurant ratings using customer data
- Conducted data cleaning, feature selection, and visualized insights using Seaborn and Matplotlib.
- Technologies: Python, Pandas, Matplotlib, Seaborn, Scikit-learn

## ADDITIONAL INFORMATION

**Certifications:**

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

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