Nanda Varshitha

Hyderabad nandavarshitha722@gmail.com 9701328295

github linkedin

Summary

Analytical and results-driven Data Analyst with expertise in Python, SQL, and advanced analytics, specializing in machine learning, data visualization, and predictive modeling. Proven track record in building end-to-end solutions from automated ETL pipelines and dashboards to interactive ML applications using Power BI and Gradio. Skilled at translating complex data into actionable insights and driving measurable business impact in collaborative environments.

Technical Skills

Programming & Data: Python (Pandas, NumPy, Seaborn), SQL, ETL pipelines, Data Cleaning, Data Wrangling

Machine Learning: Regression, Classification, Random Forest, SVM, KNN, CNN (TensorFlow/Keras)

Visualization & Dashboarding: Power BI, Matplotlib

Web & Automation: Web Scraping (BeautifulSoup, Requests, Selenium), Gradio deployment,Git

Reporting: Executive reporting, business analytics

Education

Gokaraju Lailavathi Women's Engineering College *2021–2025*
Bachelor's Degree in Computer Science
CGPA: **7.85**

Sri Chaitanya Junior College (Intermediate) *2019–2021*
Percentage: **95%**

Pragathi Model High School *2018–2019*
CGPA: **9.3**

Projects

Rainfall Prediction Using Machine Learning GitHub *2025*

- Developed and deployed a Random Forest regression model to predict rainfall using multiple weather variables and real-world data.
- Visualized feature relationships and rainfall trends with Matplotlib for transparent model interpretation.
- Deployed an interactive Gradio web application for real-time rainfall prediction based on user input.

ATM Transaction Dashboard & Cost Analysis GitHub *2024*

- Built an interactive dashboard analyzing ATMs across states for performance, cost, and maintenance insights using Power BI/Tableau.
- Engineered ETL pipelines and performed clustering and margin analysis to identify underperforming ATMs.
- Automated revenue forecasting and delivered actionable recommendations for operational improvement.

Lung Cancer Detection via Image Classification GitHub *2024*

- Developed a cancer detection system using CNN and classical machine learning (Random Forest, SVM, KNN) on medical imaging data.
- Engineered image preprocessing, augmentation, and benchmarking through ROC, AUC, and accuracy metrics.
- Automated single-image diagnostic predictions and compared models for optimal accuracy.

HealthCare Product Analysis (Nykaa) GitHub *2023*

- Scraped and analyzed 150+ healthcare product listings from Nykaa using BeautifulSoup and Requests.
- Conducted detailed statistical analysis on price, discount, and brand trends, identifying outliers and customer patterns.
- Visualized insights using Pandas, Matplotlib, and Seaborn to inform pricing and marketing strategies.

Grocery Store Management System (SQL) GitHub *2023*

- Designed and normalized relational database schemas for inventory, billing, and employee modules in a retail setting.
- Wrote and optimized advanced SQL queries to ensure accurate and efficient store data management.
- Enhanced data integrity and query speed with indexing and relational modeling.

Certifications & Achievements

Advanced Data Science Program — Innomatics Research Labs (Ongoing, 2025)

The Joy of Computing using Python — NPTEL

Machine Learning and Data Science — GeeksForGeeks