SHAIK FAIZUL GAFFAR

+91 9110391236 | **fyzul1830@gmail.com** | Kadapa, Andhra Pradesh | **LinkedIn** | **Portfolio** | **GitHub**

OBJECTIVE

Data Science undergraduate with hands-on experience in **Python, C, SQL, and Machine Learning**. Proficient in **data analysis, visualization** (Power BI, Excel), and building real-world ML solutions. Eager to contribute to impactful data-driven projects in a collaborative environment.

EDUCATION

GITAM UNIVERSITY, Bengaluru, Karnataka — 2022-2026
Bachelor of Technology, Computer Science and Engineering — *CGPA: 9.43 (up to 6th Semester)*

MEDHA JUNIOR COLLEGE, Kadapa, Andhra Pradesh — 2020-2022
Intermediate - XIIth Standard, PCM — *Percentage: 94.8%*

SRI SAI HIGH SCHOOL, Kadapa, Andhra Pradesh — 2019-2020
SSC - Xth Standard — *Percentage: 100%*

WORK EXPERIENCE

GEOSPATIAL DATA SCIENCE LAB — Murthy Facility, GITAM University
Research Analyst — Jan 2025 - Mar 2025

- Analyzed key **remote sensing indices** including Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), and Normalized Difference Built-up Index (NDBI) using **QGIS software.**
- Applied a L**inear Regression model** using **NDVI as the independent variable** and **LST as the dependent variable.** The model achieved an R^2 value ranging from 0.55 to 0.60, **showing that up to 60% of LST variation is explained by vegetation cover**, reinforcing the cooling influence of green spaces in urban areas.

PROJECTS

Exploratory Data Analysis of Google Play Store Apps — **GitHub**

- **Analyzed 10K+ Android apps** to extract insights on ratings, installs, and categories
- Cleaned data by **handling missing values, and outliers** using Python (Pandas & NumPy)
- Performed **univariate and bivariate analysis** to explore trends in various aspects.

Sentiment Analysis Using Machine Learning — **GitHub**

- Preprocessed 10,000+ tweets **to classify sentiments as Positive, Negative, or Neutral**
- Applied **NLP techniques (tokenization, TF-IDF)** using NLTK and Scikit-learn
- Trained **Logistic Regression and SVM models**, achieving up to 70% accuracy

Big Data Analysis of NYC Taxi Trips — **GitHub**

- **Analyzed 1.2M+ taxi trip records** using PySpark for time-based and spatial analysis.
- Performed data cleaning, **feature engineering**, and **calculated metrics like average trip time and tips.**
- Ge**nerated visualizations using Matplotlib** to extract key insights such as fare patterns, trip density by hour.

SKILLS

- Programming languages: Python, R, C, SQL
- Tools: Excel, Git, Jupyter Notebook, VSCode
- Libraries: Numpy, Pandas, Matplotlib, Seaborn, Scikit Learn
- Visualization: PowerBI
- Machine Learning & Deep Learning
- Probability & Statistics

ACHIEVEMENTS & CERTIFICATIONS

- Secured Top 5% Rank in **AP EAMCET 2022** (Engineering stream)
- Certificate of Merit – Aptitude & Programming Test Series, **Sanfoundry** Company Readiness Program
- Certified in Python, SQL, and Excel **(Udemy)**; Machine Learning and Google Data Analytics **(Coursera)**