

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
<i>Bachelor of Technology, Computer Science and Engineering</i>	<i>CGPA: 9.43 (up to 6th Semester)</i>
MEDHA JUNIOR COLLEGE , Kadapa, Andhra Pradesh	2020-2022
<i>Intermediate - XIIth Standard, PCM</i>	<i>Percentage: 94.8%</i>
SRI SAI HIGH SCHOOL , Kadapa, Andhra Pradesh	2019-2020
<i>SSC - Xth Standard</i>	<i>Percentage: 100%</i>

WORK EXPERIENCE

GEOSPATIAL DATA SCIENCE LAB	Murthy Facility, GITAM University
<i>Research Analyst</i>	Jan 2025 - Mar 2025
<ul style="list-style-type: none">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 Linear 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	<u>GitHub</u>
<ul style="list-style-type: none">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	<u>GitHub</u>
<ul style="list-style-type: none">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	<u>GitHub</u>
<ul style="list-style-type: none">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.Generated 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**)