

Sahiti Chhanam

Data Science Student

✉ sahitichhanam@gmail.com

☎ +91 762 096 0826

📍 Nagpur Maharashtra India

🌐 github.com/Sahiti-web



Education

Bachelor of Technology (B.Tech), *Symbiosis Institute of Technology, Nagpur*
Specialization in Data Science.

Present

Technical Skills

Languages

Python, C

Libraries & Tools

Pandas, NumPy, Matplotlib, Seaborn, GeoPandas,
SHAP, Jupyter Notebook, Google Colab

Developer Tools

Git, GitHub, VS Code

Leadership Experience

Co-Coordinator & Core Member, Literary Club & ACM | SIT Nagpur

Present

- Organized and executed literary and technical events, including workshops and hackathons
- Coordinated teams and managed event logistics
- Promoted student engagement and technical learning initiatives

Courses

Python Programming with Red Hat (AD141) [🔗](#), *Red Hat Training*

Data Structures and Algorithms using Python - Part 1 [🔗](#), *Infosys Springboard*

Projects

India District Education Performance Analysis | Python, Scikit-learn, SHAP, GeoPandas [🔗](#)

- Built an end-to-end ML pipeline (regression, clustering, classification) on India's NAS dataset (1,399 districts) to identify at-risk regions
- Used SHAP with Random Forest ($R^2=0.99$) to identify key drivers; top feature explained ~39% variance
- Applied PCA + K-Means to segment districts into performance tiers (Struggling, Average, High)
- Achieved 95% accuracy in predicting at-risk districts using minimal features
- Visualized geographic disparities using GeoPandas choropleth maps

NYC Bus Delay Prediction | Python, Scikit-learn, Meteostat API, Pandas [🔗](#)

- Built ML model on 47K real-world NYC bus delay records with integrated weather data
- Engineered domain-specific features (rush hour, weather severity, interactions)
- Random Forest achieved MAE 12.3 min (37% improvement over baseline)
- Discovered higher delays during off-peak hours vs rush hour using data analysis

Professional Experience

AI & Machine Learning Intern, Elevatelabs [🔗](#)

- Cleaned and preprocessed datasets by handling missing values, outliers, and data inconsistencies
- Performed Exploratory Data Analysis (EDA) using statistical and visualization techniques to extract insights
- Built and implemented Linear and Logistic Regression models for predictive analysis
- Analyzed model performance and optimized results through feature selection and data transformation