

Hitha S

Portfolio: hithasethu.github.io/portfolio/
Github: github.com/Hithasethu
LinkedIn: linkedin.com/in/hitha-s-

Email: hithasethu@gmail.com
Mobile: +91-7306274462
Kochi, Kerala

SUMMARY

Data Analyst with a BSc in Chemistry (Honors) and a strong analytical mindset. Transitioning from a scientific background to data analytics, leveraging expertise in data interpretation, problem solving, and statistical analysis. Skilled in SQL, Python (Pandas, NumPy) and Power BI, with hands on experience in data cleaning, visualization, and reporting. Passionate about uncovering insights from complex datasets and applying data driven approaches to real world challenges. Quick learner, adaptable, and eager to contribute analytical skills in a dynamic environment.

TECHNICAL SKILLS

- **Languages:** Python, SQL
- **Frameworks:** Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn, TensorFlow, Keras
- **Tools:** MySQL, Power BI, Excel
- **Platforms:** Linux, Windows, Kaggle, Google Colab, Jupyter Notebook, AWS (S3, EC2, IAM)
- **ML Algorithms:** Linear Regression, KNN, Naive Bayes, SVM, Decision Trees, Random Forest
- **DL Algorithms:** ANN, CNN, LSTM, OCR, YOLO
- **Soft Skills:** Leadership, Time Management, Communication, Decision Making, Problem Solving

EXPERIENCE

- **Luminar Technolab** Kochi, Kerala
Data Science Intern (Present) Dec 2024-
 - **Exploratory Data Analysis (EDA):** Identified trends and patterns using Python (Matplotlib, Seaborn, Plotly), SQL and Power BI
 - **Data Cleaning & Preprocessing:** Processed and cleaned large datasets using Python (Pandas, NumPy), SQL, improving data accuracy.
 - **Dashboard Creation:** Designed interactive dashboards (Power BI) to visualize key metrics for stakeholders
 - **Business Insights:** Analyzed sales/customer data, providing actionable recommendations to stakeholders.
 - **Machine Learning:** Developed predictive models using scikit-learn for [specific task, e.g., price prediction, sales forecasting].

PROJECTS

- **Flight Price Prediction Model - (Linear Regression, Encoding, Correlation):** Developed a Linear Regression model to predict flight price based on structured data. Preprocessing included statistical correlation (feature selection), one-hot encoding (get-dummies). Evaluated model performance using RMSE and interpreted coefficients to derive insights into price driving factors. (Github)
- **Analyzing International Education Costs using SQL:** Conducted a comprehensive SQL-based analysis of over 900 academic programs across 69 countries to identify affordable study destinations and highlight cost-intensive regions. Extracted insights on tuition, rent, visa, and insurance costs using advanced SQL queries and aggregated results to support student decision-making. Visualized findings using Power BI for clear, data-driven recommendations. (Github)
- **E-commerce Sales Classification Model – (KNN, Naive Bayes, SVM):** Built classification models to predict product categories based on customer and sales features. Performed data preprocessing including outlier detection and feature engineering. Evaluated model performance using accuracy and classification metrics, and compared algorithms to identify the best-performing model. (Github)
- **SMS Spam Classification using NLP:** Built an SMS classification model using Natural Language Processing (NLP) techniques to distinguish spam from ham messages. Preprocessed text data (tokenization, stopword removal, TF-IDF vectorization), and trained classification models, including knn, naive Bayes, SVM, Decision Tree and Random Forest. Achieved high accuracy and precision, demonstrating the ability to handle unstructured textual data and apply ML techniques effectively. (Github)

EDUCATION

- **University Of Delhi** Delhi, India
Bachelor of Science- Chemistry; CGPA: 7.757 Nov 2021 - Jun 2024

CERTIFICATIONS

- Python Basic (HackerRank)- Feb, 2025
- Python for Data Science (Reliance Foundation)- Mar 2025
- Data Visualisation: Empowering Business with Effective Insights – TATA (Forage, 2025)