

# Aditi Sahu

📍 Surat, Gujarat 📩 aditisahu3845@gmail.com ☎ 7377995640 💬 aditi-sahu-080773266  
👤 Aditi-Sahu-1104

## Summary

Aspiring Data Scientist and Machine Learning enthusiast with hands-on experience in building predictive models and deploying ML applications.

## Education

<b>B.Tech (Computer Science Engineering with Big Data Analytics)</b> Parul Institute of Engineering and Technology	2021 - 2025 CGPA: 7.95
<b>Higher Secondary Education (CHSEO - Intermediate)</b> Rama Devi Womens' HS School	2018 - 2020 Percentage: 78.83
<b>Secondary Education (BSEO - SSC)</b> Saraswati Vidya Mandir	2018 Percentage: 88.33

## Skills

**Languages:** HTML, CSS, JavaScript, C++, Python, SQL  
**FrameWorks:** pandas, NumPy, scikit learn, matplotlib, seaborn, TensorFlow, Streamlit, Django  
**Tools & Technologies:** Data Cleaning & Preprocessing, Feature Engineering, ETL, EDA, Machine Learning Algorithms, Statistics, Web Scraping, Excel, Power BI, Tableau, Docker, Git, AWS/Azure(Basics)  
**Soft Skill:** Good Communication, Team Management

## Projects

### Movie Recommendation System [GitHub Repository](#)

- Developed a web-based movie recommendation app with interactive UI using Streamlit. Trained a content-based filtering model to suggest movies using Python(Pandas, Numpy, Scikit-learn, nltk) and serialized the trained model using .pkl for efficient loading.
- Used data preprocessing techniques to enhance accuracy. Deployed live on Streamlit Community Cloud, ensuring scalability and accessibility for users. [Live App](#)

### Customer Churn Prediction [GitHub Repository](#)

- Designed a predictive model identifying potential churners using Support Vector Classifier (SVC).- Improved model accuracy through feature selection & data balancing techniques.

### Vendor Performance Analysis [GitHub Repository](#)

- Designed and implemented a Vendor Sales Analysis pipeline by importing CSV data into a relational database.Utilized SQL functions such as JOIN, GROUP BY, WINDOW FUNCTIONS, and aggregate functions to extract and summarize key metrics.
- Applied Python for data cleaning, feature engineering, and data transformation. Created interactive Power BI dashboards visualizing vendor performance, sales trends, and product category analysis.

## Experience

### Machine Learning Internship

Jan 2025 – Apr 2025

Infotact Solutions

- During the internship I Worked on 3 ML projects involving classification, prediction, and recommendation systems, applying Logistic Regression, SVM, and Collaborative Filtering. Performed preprocessing, feature engineering, model evaluation, and deployed models using Streamlit for end-to-end implementation.

## Certifications

Cleared the 1st round of Vultr Cloud Hackathon, Data Science Course (Udemy), Data Science with Python(Coincent), Microsoft Power BI Course (Infosys), Data Analytics Foundation (LinkedIn)