

SOBHIT TRIPATHI

tripahtishobhit824@gmail.com || [LinkedIn](#) || [Github](#)

PROFESSIONAL SUMMARY

Results-driven Data Analyst with expertise in Python, SQL, and Power BI, specializing in ETL, data visualization, and analytics-driven decision-making. Proficient in Microsoft Excel, Google Sheets, and stakeholder collaboration. Experience in the payments industry, working with banking transactions and fraud detection. Strong communication skills for engaging with internal and external stakeholders.

EDUCATION

BS in Data Science, IIT Madras

OCT 2021 - Present

SKILLS

Programming	Python, SQL, ETL Pipelines, TensorFlow
Data Processing	Data Cleaning (Pandas, NumPy), Data Mapping & Transformation
Tools & Version Control	Git, Github, Power BI, MS-Excel, Google Sheets
Professional Skills	Requirement Analysis, Communication, Stakeholder Collaboration

PROFESSIONAL EXPERIENCE

Infotact PVT. LTD — Data Analyst Intern

Jan 2025 - Present

- Perform ETL processes and data validation on **150K+ transactions** to enhance fraud detection using **Python and SQL**, improving model efficiency by **15%**
- Designed and implemented **interactive Power BI dashboards to track fraud patterns**, improving early detection and business decision-making.
- Applied machine learning techniques (Random Forest) to **improve fraud detection precision by 18%**, leveraging data preprocessing and advanced analytics.

InternShip Studio — Machine Learning Intern

July 2024 - Aug 2024

- Designed **ETL pipelines** for transforming **15,000+ records** to improve YouTube ad view predictions.
- Conducted regression analysis, to improve data quality reducing errors by **20% and boosting predictive accuracy to enhance YouTube ad performance by 15%**.

ACADEMIC PROJECTS

Telecom's Customer Value Prediction

[Github](#)

- Designed **ETL workflows in python for 250k+ records**, ensuring data quality enabling advanced customer segmentation.
- Developed predictive models in Python (Random Forest) for Customer Lifetime Value (CLV) **forecasting with 90%+ accuracy, aiding in strategic decision-making.** and achieve test **RMSE of 134.83**

Stock Movement Analysis

[Github](#)

- Build a model by integrating **market trends with sentiment analysis**, using BeautifulSoup for **web scraping and processing 300K+ data points.**
- Applied **TextBlob for sentiment scoring**, enhancing stock trend predictions for data-driven investment insights. and **deployed the model on GitHub**, for further enhancements and collaboration

IPL Data Analysis

- Examined 200K+ records from (2008-2024) of IPL data to identify team trends, player performance, and match patterns.
- Designed **Power BI dashboards with 8+ interactive visualizations** for IPL performance analysis, implementing real-time filtering and insights extraction.

CERTIFICATIONS

- Python (HackerRank), Power BI (PwC), Data Visualization (Tata), BCG NLP Job Simulation