

SHUBHAM GOYAL

+91 9250733923

goyalshubham575@gmail.com

[LinkedIn](#) | [GitHub](#)

[Portfolio](#)

EDUCATION

Bhagwan Parshuram Institute of Technology

B-Tech, Electrical and Electronics Engineering

SKILLS

Analytical Tools:

Power Bi
Excel
Tableau

Programming Languages:

SQL (Structured Query Language)
Python

Database:

MySQL

Machine Learning

Supervised and Unsupervised Learning
Hypothesis Testing
Time Series Analysis
Natural Language Processing (NLP)

PROJECTS

Spam Detection [Link](#)

- Built a spam detection model using NLP, classifying messages as ham or spam.
- Tested multiple classifiers, with SVC achieving **98%** accuracy and a **1.0** precision score.
- Implemented preprocessing techniques to improve text classification performance.

Ecommerce Product Categorization (Data Science Hackathon) [Link](#)

- Developed a **text classification model** to accurately categorize eCommerce products based on product descriptions.
- Applied **NLP techniques** such as TF-IDF and word embeddings for feature engineering and improved categorization accuracy.
- Built and fine-tuned **machine learning models**, optimizing hyperparameters to enhance precision, recall, and F1 score.
- Presented findings through **data visualizations** (word clouds, category distributions) and a structured Jupyter Notebook.

Virgin Airlines (Sentiment Analysis) [Link](#)

- Developed a tweet sentiment classification model to analyze customer perceptions of Virgin Airlines on social media.
- Tested Logistic Regression, MultinomialNB, Random Forest, and SVC, with **Logistic Regression** achieving best accuracy of **84%**.
- Created a word frequency chart to visualize the most commonly used words in positive and negative tweets.

Market Basket Analysis for Wallkart [Link](#)

- Analyzed transaction data to identify frequently bought-together items and optimize product placement.
- Discovered key product combinations, such as **rolls/buns & whole milk, yogurt & whole milk, and soda & vegetables**, to enhance visibility and sales.
- Identified most frequently purchased items & customer purchase patterns by day, aiding in better inventory & marketing strategies.

Hypothesis Testing (Ad Performance Analysis: Facebook vs. AdWords) [Link](#)

- Compared Facebook and AdWords ad campaigns based on clicks, conversions, and cost-effectiveness to optimize ROI.
- Identified a stable and proportional relationship between ad spend and conversions, demonstrating long-term effectiveness.
- Highlighted influence of external factors like market conditions & audience behavior, emphasizing need for deeper validation.

Time series forecasting (Netflix Stock Price Time Series Analysis) [Link](#)

- Analyzed Netflix stock prices (2018–2022) using various time series models to predict future stock trends.
- Evaluated MA, ARMA, ARIMA, SARIMA, and Smoothing models, with Double Exponential Smoothing achieving the lowest RMSE of **9.81**.
- Provided insights into stock price patterns and volatility, aiding in better financial forecasting.

IPL insight [Power BI | SQL] [Link](#)

- Developed an interactive dashboard for a sports magazine analyzing three years of IPL data to evaluate batter bowler performance and overall team statistics.
- Provided comprehensive player insights by tracking metrics such purple cap and orange cap winners, match details, and win counts, enhancing editorial content and strategic sports analysis.
- Help to identify trends and patterns in players and team performances, offering valuable insights for editorial content.

Hospitality Domain [Power BI | SQL] [Link](#)

- Mumbai city has generated highest revenue **39%**, while Delhi has highest occupancy rate **26%** and cancellation rate **25%**

- Weekdays generated nearly **2 times** revenue than weekends.
- Atliq Exotica accounted for **18.7%** of total revenue, making it the highest revenue generating hotel.
- Highest numbers of booking done from Other platform (**41%**) followed by MakeMyTrip (**20%**).

FMCG Domain [SQL | Power Point]

[Link](#) Feb 2024

- Presented Insights on ad-hoc business requests, by querying the million rows database using SQL. Designed a Power Bi dashboard to understand sales.
- Revenue increased by more than **2x** in both campaigns while the units sold was **4x** in Diwali and **2x** in Sankranti sales.
- **500 Cashbacks** outperform in Diwali campaign and **BOGOF** outperform in Sankranti campaign. **25% Off** and **50% Off** discounts underperform in both campaigns.

INTERNSHIP

Med Tour Easy

- Analysed a cosmetics dataset using Python (Pandas, NumPy, scikit-learn) to uncover insights.
- Applied t-SNE for dimensionality reduction and used Bokeh for interactive data visualization.

4-Week Virtual Internship | AtliQ Technologies

May 2024

- Conducted data cleaning, transformation, and visualization tasks to support business decision making process.
- Assisted in developing interactive dashboards and reports, contributing to strategic business insights.
- Created presentations to communicate analytical findings & insights to stakeholders, enhancing data-driven decision making.

EXPERIENCE

Collegetravellers (Co-Founder)

[Link](#) 2020-2022

- Orchestrating over 20 trips, fostering memorable experiences, and cultivating a vibrant community of travelers.
- Managed logistics, Marketing, and Customer relations, ensuring high satisfaction and repeat business.

Delhi NCR Memes

[Link](#) 2018-2022

- Managed Instagram page with over 16000 followers, curating content and engaging with followers to drive growth.
- Developed and executed content strategies, increasing followers' engagement and page visibility.