Arpit Dubey

Mob: +91 8303994991 | Email:dubeyarpit0608@gmail.com|
LinkedIn: **linkedin.com/in/arpitd06**
GitHub: **github.com/ArpitD06**

Profile Summary

A Data Science professional with experience in SMT operations. Skilled in Python, SQL, Power BI, Tableau and Machine Learning, with a focus on data-driven insights for Ecommerce Problems, Predict Profit Startup and Analysis of Ted Talks. Proven ability to optimize processes, collaborate with teams, and solve real-world business problems using tools likePandas, Matplotlib, Seaborn, and Scikit-learn. Seeking to leverage technical expertise and analytical skills to contribute to business growth and efficiency in a data-centric role.

Work Experience

Current Organization: Harman-Pune **September 2022 - Present**

Role: SMT Operation

- Conducting regular analysis of the SMT machine performance and cycletimes to identify bottlenecks and inefficiencies.
- Monitoring production output and analyzing trends in product yields and defects.
- Leveraging tools like Excel and data visualization techniques to track key performance metrics such as machine uptime, throughput, and defect rates. I regularly report these metrics tomanagement, providing insights that inform decision-making.
- Analyzing material handling and component availability, using data topredict shortages and streamline material arrangement. This ensures smooth production with minimal delays.

Internship

Data Science Intern : Access million (2024)

- Worked under the marketing department, analyzing marketing campaigns and customer behaviors using Python, SQL And Tableau.
- Developed data-driven insights to improve marketing strategies and optimize lead generation.
-

Top Mentor Training and Certification

- **Master** in **Data Science** and **Analytics**
- **Tableau** and **Power BI**. (March 2024 – April 2024) Guided by Mr. Bose
- **SQL** and **Database Management System**. (February 2024 – March 2024)
- **Python** with **Data Analysis**. (January 2024- July 2024)
- **Machine Learning** (January 2024- July 2024)

Academics

Examination	Percentage	Name of the Institute	University	Year
Diploma (Electronics)	77%	Gov. Polytechnic Rampur	Board of Technical Education Uttar Pradesh	2022
HSC (Science)	63%	Sukhpura Inter College, Ballia	Uttar Pradesh State Board	2019
SSC	85%	S J N Inter College, Ballia	Uttar Pradesh State Board	2017

Projects

Analyzing eCommerce Problems:

- **Problems -** Customers satisfaction due to missing reviews, predicting sales accurately, optimizing delivery performance, analyzing product categories for dissatisfaction, handling incomplete data, and delivery times without a clear optimization strategy.
- **Tools -** Using Pandas for data manipulation, NumPy for numerical operations, Matplotlib and Seaborn for data visualization, Scikit-learn for machine learning modeling, Stats models for statistical analysis, NLTK and Text Blob for natural language processing.
- **Outcomes -** Achieved **75%** of customer reviews highlighted satisfaction factors like fast delivery and product quality. Forecasting predicted a **30%** revenue increase for 2019. Delivery time analysis indicated a **20%** opportunity for improvement. Optimizing product categories identified **25%** of categories as key focus areas for enhanced performance

Predict Profit Startup:

- **Problems -** The main challenges included handling missing values, transforming categorical data for regression analysis, and ensuring the model effectively predicts profits based on varying expenditures.
- **Tools -** Using Python libraries like Pandas for data manipulation, Scikit-learn for preprocessing and model building, and Matplotlib for data visualization. Techniques such as Label Encoding, One-Hot Encoding, and Standard Scaling were employed for feature preparation.
- **Outcomes -** The model achieved an R^2 score of approximately 0.99, indicating a strong predictive ability. The Mean Absolute Error for the full model was around 2969.05, demonstrating effective profit prediction based on expenditures.

Analysis of Ted Talks:

- **Problems -** The analysis aimed to explore the trends and insights from the TED talks dataset, including popularity, engagement through views and comments, speaker characteristics, and the distribution of talks over time.
- **Tools -** The project utilized Python libraries such as Pandas for data manipulation, NumPy for numerical operations, Matplotlib and Seaborn for data visualization, and Word Cloud for text analysis.
- **Outcomes -** The analysis revealed the most viewed TED talks and their correlation with comments. It highlighted trends in talk frequency by month and day, popular speaker occupations, and themes. Visualizations showcased audience engagement through view counts and talk durations, offering insights into the factors driving TED talk popularity

Skills

- Power BI and Tableau
- Python
- Business Analysis
- Data Analysis and Visualization
- Machine Learning
- SQL and Database Management
- Microsoft Office
- Communication

Declaration

I hereby declare that the details mentioned above in my resume are correct to the best of my knowledge and belief.

Date:

Place: Arpit Dubey