Sarvesh Mishra [Kaggle Expert]

Kaggle • Github • LinkeDIn

sarveshmishra.bhu20@gmail.com
+91 8957883380

Educational qualification

Bachelor of Science, Banaras Hindu University — *2020-2023*
CGPA - 7.6/10

Data Science Course & Certification

iNeuron Full Stack Data Science Masters — *06/2023 - 02/2024*

Technical skills

- **Machine Learning** : Supervised and Unsupervised
- **Deep Learning**: ANN, CNN
- **Language and Framworks** : Python, R, Scikit-learn, Flask
- **Database:** SQL, Mysql, Snowflake
- **Libraries** : Sklearn , Numpy , Pandas , Beautifulsoup
- **Natural Language Prcessing (NLP)** : NLTK, Word2Vec
- Statistics, Linear algebra, Probability
- **Data analysis & visualization tools**: Excel, PowerBI
- **Others:** AWS, DSA

Projects

1. Swiggy Customer order details analysis of the 9 different cities

View Project

- **Objective:** The primary goal was to derive meaningful insights, both in terms of numbers and graphical representations, to facilitate informed business decisions and improvements.
- In total, there were **8,680** orders across nine cities, with the highest percentage **15.50% originating from Kolkata**.
- Revenue analysis reveals that Mumbai contributes to **17%** of the total revenue, which amounts to **0.56 million IN**R.
- Notably, the maximum revenue per order is from Mumbai at **393 INR**, while the minimum is from Surat at **270 INR.**
- The average delivery time across all cities is **54 minutes**. The slowest delivery time is in Kolkata, taking **68 minutes**, whereas the fastest is in Mumbai, with an impressive **48 minutes**.
- **Tools used :** Microsoft Excel, Power BI, SQL, Snowflake

2. Expense Data Analysis and Visualization Using Power BI

View Project

- **Objective:** This project focuses on leveraging Power BI and Excel to analyze and visually represent expense data, providing stakeholders with actionable insights for informed decision-making. The primary objective is to present expense-related information in a meaningful and easily understandable format. The project employs various visualization techniques, including **Bar Charts, Pie** Charts, and Line Charts, to effectively communicate key findings.
- Males who smoke incur expenses that are **25.6%** higher than those who don't smoke.
- Females who don't smoke experience **35% higher expenses than their smoking counterparts.**
- **A substantial 63%** of the revenue comes from customers who either have **one child or no children**.

3. Customer Churn Prediction System Using Machine Learning and Flask

View Project

- **Objective:** In this machine learning project, the primary objective was to develop a robust customer churn prediction system capable of forecasting whether a customer is likely to cancel their subscription or remain a loyal subscriber. The project employed classification machine learning algorithms to address this critical business challenge, providing valuable insights for customer retention strategies.
- Developed meaningful insights in data exploration part using pandas
- Developed a web application using Flask to deploy the machine learning model locally.
- Created a user-friendly interface for users to input relevant data and receive predictions regarding customer churn.
- Integrated the trained model into the Flask application for real-time predictions.

4. Diabetes Data Analysis and Prediction *View Project*

- **Objective:** The primary objective was to analyze and visualize a dataset related to diabetes, aiming to identify the factors that contribute to the occurrence of diabetes.
- Conducted thorough exploration of the dataset to understand its structure and characteristics. Utilized statistical and graphical methods to summarize and interpret the main features of the data. Visualized relationships between different independent variables and the target variable (Outcome) to identify potential patterns and correlations.
- Implemented a simple machine learning model using regression techniques. Chose an appropriate regression algorithm based on the nature of the problem and dataset.

Achievements

Achieved Kaggle 2x Expert status through consistent participation and high-performance rankings in data science competitions. Demonstrated expertise in applying advanced machine learning techniques and collaborating with a global community of data scientists

4 Star SQL on HackerRank

About me:

With a solid foundation in mathematics, physics, and geology acquired during my BSc studies at Banaras Hindu University, I have developed a strong analytical mindset and a keen interest in problem-solving. Now, I am enthusiastically transitioning into the field of data science, driven by my passion to tackle complex data-related challenges and harness the power of data-driven insights to make a meaningful impact.