Week 10 Deliverable
Group Name: Neural Nomad

Name: Nitish Jena
Email: nitishkumar.jena.2023@student.ism.de
Country: Germany
College: ISM Hochschule
Specialization: Data Science

Github Repo link: https://github.com/nitishjena/Week-10

Problem Description:
ABC Bank seeks to develop a predictive model to determine the likelihood of customers subscribing to their term deposit product based on historical interactions. The objective is to leverage machine learning techniques to identify potential customers who exhibit a higher propensity to purchase the product. This will enable the bank to efficiently target and prioritize customers with a greater likelihood of engagement, thereby optimizing marketing efforts and improving conversion rates.

Exploratory Data Analysis (EDA):
I conducted an extensive exploratory data analysis (EDA) to gain deeper insights into the dataset, focusing on both categorical and continuous variables. The primary objective was to understand the relationships and patterns within the data, enabling us to extract meaningful insights.

My analysis aimed to examine how various demographic factors such as age, education level, marital status, and occupation, correlate with customers' likelihood of subscribing to the term deposit. To achieve this, I employed bar chart visualizations, which effectively illustrate the distribution of subscribers and non-subscribers across different attributes.

Specifically, I generated three bar charts displaying the number of customers who subscribed versus those who did not, segmented by age, job, education, and marital status. The analysis revealed a consistent pattern: across all these attributes, the number of non-subscribers exceeded that of subscribers.

Additionally, I performed a conditional probability analysis to evaluate the likelihood of customer subscription based on various categorical attributes, such as age, education, housing status, and loan status. This approach provided a more nuanced understanding by focusing on success rates within specific groups rather than absolute counts. This method is particularly effective in mitigating issues related to skewed distributions, over-representation, and under-representation, allowing for a more balanced and accurate interpretation of the dataset.

Final Recommendation:
To refine my analysis, I will conduct a more comprehensive examination of the relationships between key features. This deeper insight will enable us to draw more precise conclusions from the data, ultimately enhancing the effectiveness of a machine learning model in predicting whether a customer is likely to subscribe.