**TEBONG ROLAND**

**DSC680-T301 Applied Data Science (2243-1)**

[**Questions**](https://cyberactive.bellevue.edu/webapps/assignment/uploadAssignment?content_id=_15084585_1&course_id=_522998_1&group_id=&mode=view) **and Answers**

**Questions:**

1. What is the primary business problem being addressed using machine learning in this scenario?

**There has been a revenue decline in the Portuguese Bank, and they would like to know what actions to take. customers are not investing enough in long-term deposits.**

1. What is the target variable in the classification problem addressed in this report?

**The target those customers likely to subscribe to a term deposit.**

1. How many records and columns are there in the 'banking\_data.csv' dataset used for analysis?

**There are 32951 Rows and 15 columns in the Banking\_data\_csv dataset.**

1. What are some key attributes included in the data dictionary for the features used in the classification problem?

**The marketing campaigns were based on phone calls.**

1. What are the assumptions made regarding the data collection process and its representativeness?

**In the context of data collection, it is assumed that the information gathered from the Portuguese Bank's historical records is a representative sample of the broader customer population.**

1. What limitations are highlighted in the analysis, specifically concerning data quality and model predictions?

**Firstly, the findings heavily depend on the quality and completeness of the data contained in the 'banking\_data.csv' file.**

**Incomplete or inaccurate information may compromise the accuracy and reliability of the models developed.**

**Furthermore, the models' predictions are based on historical patterns if future customer behavior will follow similar trends.**

1. What challenges are discussed, particularly related to imbalanced classes and the interpretability of complex models?

**The primary challenges encountered in this analysis revolve around the imbalanced distribution of the target variable 'y,' indicating whether a client subscribed to a term deposit.**

**The imbalanced classes, with a considerably lower percentage of positive cases, pose challenges in model training and may lead to a bias towards predicting the majority class.**

**Additionally, the use of complex models, particularly the Artificial Neural Network, introduces the challenge of interpretability, making it difficult to provide transparent explanations for specific predictions.**

1. What future uses or additional applications are suggested for the developed predictive models?

**Looking ahead, the developed models can be leveraged for various purposes beyond the binary classification task. One potential application is customer segmentation, where the bank can categorize clients based on specific characteristics and behaviors identified by the models.**

1. What key recommendations are provided for enhancing the effectiveness of the predictive models?

**To enhance the effectiveness of the predictive models and the overall decision-making process, it is recommended to conduct a thorough analysis of feature importance.**

1. How is the performance of different models, including Artificial Neural Network, Decision Tree, and Random Forest Classifier, compared in terms of accuracy?

**Looking at the performance of the three models , it can be observed that the Random Forest classifier is outperforming other models with an accuracy of 90%.**