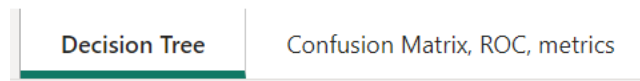


## Summary Report

### 1. Guide to Navigating the Live Reports in Power BI for the Marketing Team:

- The Power BI dashboard is designed to provide a visual representation of data analysis, in this case, a decision tree classifier.
- To interact with the report, you can select different nodes of the decision tree to view the criteria used to split the data at each decision point.

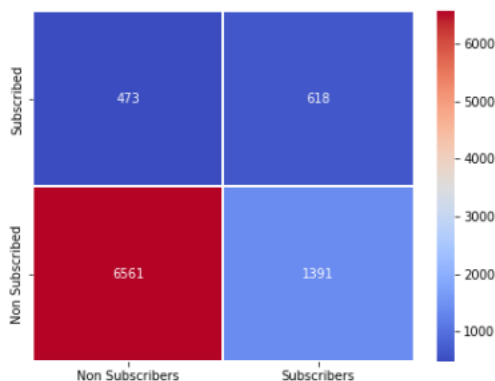


- The table shows the 'Level of Feature Importance' to see the exact importance score of each feature.

Most Informative Feature	Score
housing	0.112064805
marital	0.070062775
poutcome_failure	0.027761639
poutcome_unknown	0.022360529
poutcome_other	0.017962824
job_technician	0.014654581
contact_unknown	0.011915813
month_may	0.011218418
job_blue-collar	0.009226268
pdays	0.003128137
education	0.001858729
job_management	0.000533256

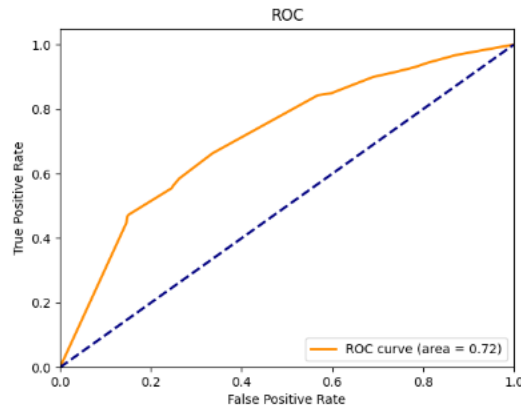
- Use the confusion matrix to understand the model's performance in terms of actual vs. predicted results.

Confusion Matrix Non Subscribers and Subscribers



- The ROC curve provides insight into the trade-off between true positive rate and false positive rate for different threshold settings.

ROC Curve



## 2. Analysis of the Features and Outcome:

- The feature importance chart shows that 'housing\_no' is the most significant feature, followed by 'contact\_unknown' and 'poutcome\_unknown'.
- Considering the operational objective, focus on these top features as they have the most influence on the outcome.
- The decision tree visualization helps identify the thresholds and combinations of feature values that lead to different predictions.

## 3. Advantages and Drawbacks of Power BI and Python:

### Advantages:

- Power BI provides an interactive user interface that's easy for non-technical stakeholders to navigate and understand.
- Python, on the other hand, offers robust data processing capabilities, flexibility in statistical analysis, and machine learning tasks.

### Drawbacks:

- Power BI has limitations in advanced data processing and customization, which Python can handle with more control.
- Python requires a higher level of technical expertise, which might not be accessible to all team members.

#### 4. Challenges Encountered:

- Data cleaning and preprocessing was slightly challenging, ensuring data quality is a prerequisite for accurate modeling.
- Integrating Python scripts with Power BI to run the decision tree model and update the live reports seamlessly was technically demanding as this was not previously completed.

#### 5. Development Process and Technical Documentation:

- The development process included data collection, cleaning, preprocessing, model training (decision tree), and visualization – see attached 'Technical Report'
- The technical documentation details each step, with the Python scripts used, parameters selected for the decision tree model, and guidelines on interpreting each component of the Power BI dashboard.

#### 6. Future Recommendations:

- Continuously update the model with new data to maintain its accuracy.
- Explore modeling techniques like deep learning as these yielded the best results.
- Consider implementing A/B testing to compare the performance of different models in a live environment.