# Model Performance Report

# Description of data

This report incorporates a detailed analysis of a key data set used to predict the risk of corporate customer defection. The data set excludes direct corporate identifiers, focusing predictive modeling on significant operational variables. These variables include the monthly and annual activity of the client, which reflects the level of recent and sustained interaction, as well as the size of the client, represented by the number of employees. These metrics are essential to feed the machine learning model, whose objective is to anticipate the termination of contracts. By accurately forecasting churn risk, account managers can strategically prioritize customer retention. The dynamic approach to this analysis ensures that the model stays up-to-date in the face of changing market trends, allowing the company to minimize customer churn and maximize revenue retention.

## Ingested Files Information:

dataset3.csv: - Ingestion date: 05/02/2024 19:59:37
dataset4.csv: - Ingestion date: 05/02/2024 19:59:37
dataset3.csv: - Ingestion date: 05/02/2024 20:03:11
dataset4.csv: - Ingestion date: 05/02/2024 20:03:11
dataset3.csv: - Ingestion date: 05/02/2024 20:03:48
dataset4.csv: - Ingestion date: 05/02/2024 20:03:48

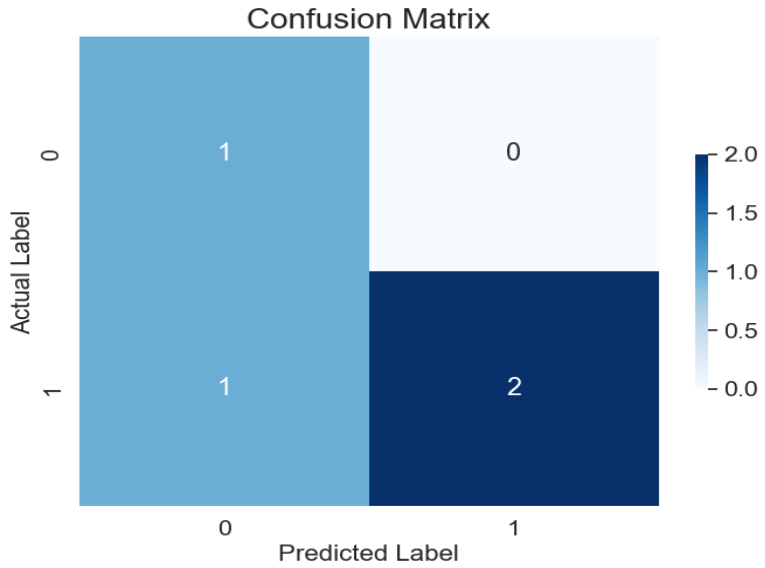
## Dataframe statistics:

| Statistic | Mean | Median | StdDev |
|---|---|---|---|
| 26.0 | 26.0 | 26.0 | 26.0 |
| 5625.923076923077 | 763.5384615384615 | 457.46153846153845 | 0.5 |
| 19444.774621317832 | 2016.6093569309696 | 800.6048328991891 | 0.5099019513592785 |
| 0.0 | 0.0 | 2.0 | 0.0 |
| 38.25 | 14.0 | 12.75 | 0.0 |
| 425.0 | 97.5 | 99.0 | 0.5 |
| 1804.75 | 474.75 | 626.5 | 1.0 |
| 98765.0 | 9982.0 | 3782.0 | 1.0 |

## F1 Score Interpretation and Current Value:

The F1 score is the harmonic mean of precision and recall, providing a balance between them. It is a good measure to use if you need to seek a balance between precision and recall and there is an uneven class distribution (large number of actual negatives).
The actual F1 score is 0.8571428571428571

## Confusion Matrix:



## Previous Run Times:

| Data Ingestion Time (seconds) | Model Training Time (seconds) |
|---|---|
| 0.49539449999999996 | 42.241368 |
| 0.5018412000000001 | 42.6331844 |
| 0.6041022999999999 | 44.7371179 |
| 0.4977117 | 42.2124315 |
| 0.5063838 | 42.058130000000006 |
| 0.531963600000001 | 43.63725839999999 |
| 0.5053149000000001 | 41.8649879 |
| 0.4946219999999997 | 41.77022049999999 |

## Data Ingestion and Model Training Times Graph: