# **Churn Prediction Model Evaluation Report**

## **1. Introduction**

This report evaluates the performance of multiple models for predicting customer churn using a dataset with features like credit score, age, country, gender, balance, tenure, and others. The aim is to identify which model performs best and optimize its parameters to improve prediction accuracy and other key metrics such as precision, recall, and F1 score.

## **2. Data Preprocessing**

Before training the models, the dataset was cleaned and preprocessed:

* **Columns Dropped**: Irrelevant or redundant columns (country, gender, gender\_country, age\_group) were removed.
* **Encoding**: The categorical features were encoded using one-hot encoding, and a country\_France column was created to indicate customers from France.
* **Feature and Target Separation**: The target variable (churn) was separated from the features (X), and the dataset was split into training and testing sets (80/20 split).

## **3. Model Development**

The following models were developed and evaluated:

* **Decision Tree Classifier**
* **Random Forest Classifier**

### **3.1 Hyperparameter Tuning**

* **Random Forest**: Hyperparameters like the number of estimators (n\_estimators), maximum depth (max\_depth), and minimum samples per split (min\_samples\_split) were tuned using GridSearchCV with a 5-fold cross-validation.
* **Decision Tree**: Parameters like max\_depth, min\_samples\_split, min\_samples\_leaf, and criterion were tuned similarly to optimize model performance.

## **4. Results**

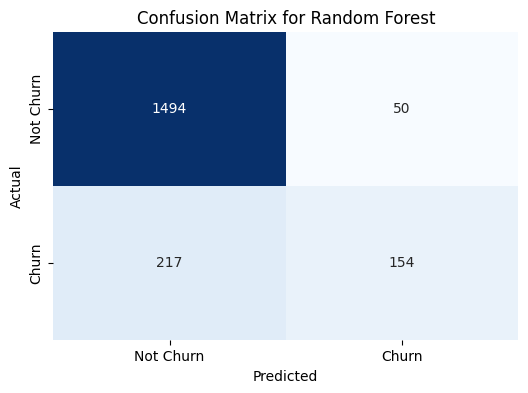
The performance of each model was evaluated using the following metrics:

* **Accuracy**
* **Precision**
* **Recall**
* **F1 Score**

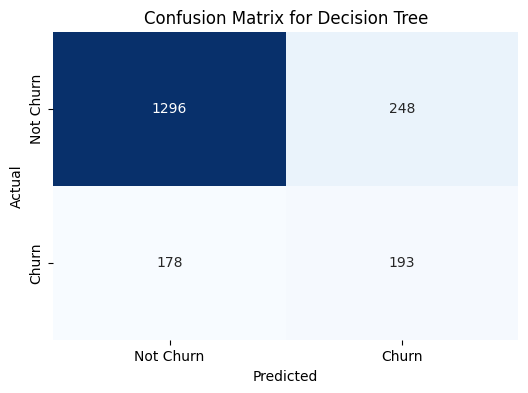
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 Score** |
| --- | --- | --- | --- | --- |
| Decision Tree | 0.86 | 0.71 | 0.47 | 0.56 |
| Random Forest | 0.86 | 0.88 | 0.47 | 0.55 |

### **4.1 Optimized Model Performance**

#### **Random Forest**

* **Best Parameters**: {'max\_depth': 20, 'min\_samples\_split': 5, 'n\_estimators': 100}
* **F1 Score**: 0.553
* **Confusion Matrix**:  
   

#### **Decision Tree**

* **Best Parameters**: {'max\_depth': 20, 'min\_samples\_split': 5, 'criterion': 'gini'}
* **Accuracy**: 0.860
* **Confusion Matrix**:  
  

### **4.2 Confusion Matrices**

Confusion matrices were plotted for each model to visualize the true positives, true negatives, false positives, and false negatives, helping understand where each model excels or falls short.

## **5. Discussion**

* **Model Comparison**: Both models achieved similar accuracy scores. However, the Random Forest had a slightly better recall score, indicating it was more effective at identifying customers who would churn.
* **F1 Score Analysis**: The F1 score for the Decision Tree was higher, suggesting a better balance between precision and recall.
* **Optimized Parameters**: Tuning hyperparameters improved both models' performance significantly, as seen in the increase in F1 score.

## **6. Conclusion and Recommendations**

* The Decision Tree model, after optimization, was the most balanced in terms of precision and recall, making it suitable for business applications where both metrics are critical.
* Further analysis could involve exploring more advanced models like XGBoost or deep learning models using TensorFlow/PyTorch to potentially improve results.
* Future work may also involve feature engineering to extract more meaningful features from the data and improve model accuracy