Optimizing Loan Acceptance Prediction

Faruk Yasar

Objective

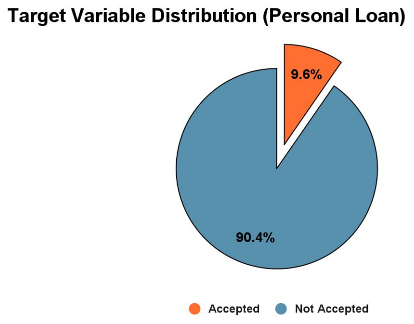
🕵️ **The problem:** Predicting customer loan acceptance to identify factors influencing their decisions.

🎯 **The Objective:** Minimize False Negatives (maximize Recall) to improve targeting of customers likely to accept offers.

📈 **Business Impact:**

- Enhanced targeting efficiency.
- Reduced missed opportunities for loan conversions.

🤔 **The Challenge**:The data is significantly imbalanced.



Data and Methodology

Dataset:

- 5,000 records, 11 features.
- Target: Personal Loan (Accepted = 1, Not Accepted = 0).
- Features:
- Income, Education, Family Size, Credit Card Spending.

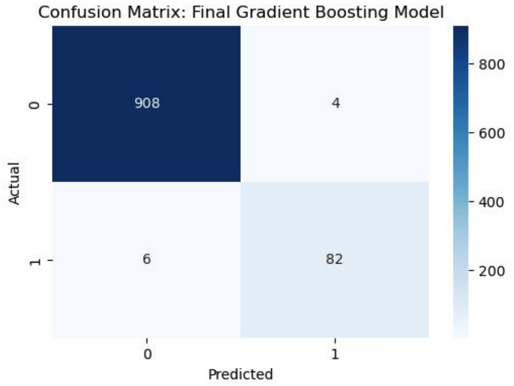
```
Dataset shape: (5000, 14)

Dataset info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999
Data columns (total 14 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   ID                  5000 non-null   int64
 1   Age                 5000 non-null   int64
 2   Experience          5000 non-null   int64
 3   Income              5000 non-null   int64
 4   ZIP Code            5000 non-null   int64
 5   Family              5000 non-null   int64
 6   CCAvg               5000 non-null   float64
 7   Education           5000 non-null   int64
 8   Mortgage            5000 non-null   int64
 9   Personal Loan       5000 non-null   int64
 10  Securities Account  5000 non-null   int64
 11  CD Account          5000 non-null   int64
 12  Online              5000 non-null   int64
 13  CreditCard          5000 non-null   int64
dtypes: float64(1), int64(13)
```

Methodology:

- **Data Preprocessing:** Scaling, and SMOTE for balancing.
- **Model Comparison:** Logistic Regression, Random Forest, Gradient Boosting, XGBoost.
- **Hyperparameter Tuning:** Focused on Gradient Boosting and XGBoost for Recall.

Model Evaluation and Selection

Model	Recall	ROC-AUC	Precision	F1 Score
Logistic Regression	0.8523	0.9529	0.4870	0.6198
Random Forest	0.9091	0.9963	0.9524	0.9302
Gradient Boosting (Pre-tune)	0.9205	0.9979	0.9101	0.9153
Gradient Boosting (Post-tune)	**0.9318**	**0.9983**	0.9123	**0.9219**
XGBosst (Pre-tune)	0.9205	0.9962	0.9000	0.9101
XGBoost (Post-tune)	0.8864	0.9966	0.9176	0.9012

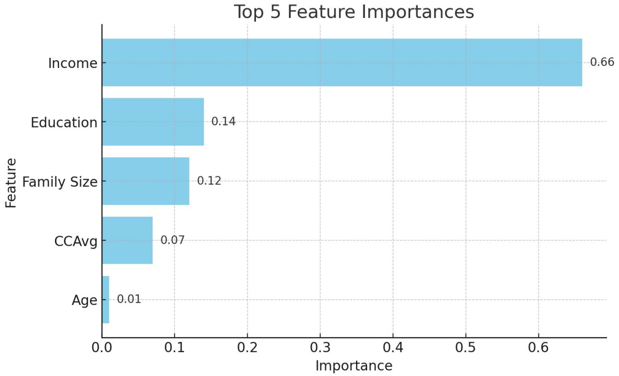


Confusion matrix for Gradient Boosting Post-Tune model showing minimal False Negatives (6) and False Positives (4).

Feature Importance and Business Insights

Features

1. **Income 💵 :** Signals financial stability.
2. **Education 🎓:** Higher education levels linked to better financial literacy.
3. **Family Size 👨‍👩‍👧‍👦:** Larger families often have higher financial needs.
4. **CCAvg (Credit Card Spending) 💳 :** Indicates financial engagement and creditworthiness.



Business Insights

- High-income customers are most likely to accept personal loan offers.
- Education level is a strong predictor of loan acceptance, indicating trust in financial products.
- Larger families show increased loan acceptance, reflecting higher financial needs.
- Credit card usage signals financial engagement and creditworthiness.

Recommendations and Impact

🎯 Target High-Potential Segments

- Focus on high-income, highly educated customers.
- Use family size and spending patterns for personalized marketing.

💡 Customize Marketing Strategies

- Emphasize financial benefits to educated customers.
- Promote flexible repayment plans for families.

📊 Business Impact

- Improve conversion rates.
- Reduce missed opportunities by targeting the right customers.
- Improve marketing ROI through better-aligned strategies.

Summary

Selected model: Post-tuned Gradient Boosting

Achieved Metrics

- Recall: 93.18% (minimizing False Negatives).
- ROC-AUC: 99.83% (excellent discriminative ability).

Impact Summary

- Enhanced targeting will reduce missed opportunities and improve conversion rates.
- Tailored marketing strategies will boost ROI for personal loan campaigns.

Next Steps

1. Deploy the Gradient Boosting model into production.
2. Monitor model performance and retrain periodically to maintain accuracy.
3. Use insights to refine marketing strategies and optimize loan products.