AUTOMATED MACHINE LEARNING ANALYSIS REPORT

Generated: February 16, 2026 at 03:07

Report Type: Full ML Pipeline Analysis

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

This report presents a comprehensive analysis of an automated machine learning pipeline executed on February 16, 2026 at 03:07. The analysis encompassed data exploration, preprocessing, feature engineering, model selection, and performance evaluation.

Key Findings

Metric	Value
Best Model	GradientBoostingClassifier
Model Score	0.8510
Models Evaluated	3

2. DATA EXPLORATORY ANALYSIS

2.1 Dataset Overview

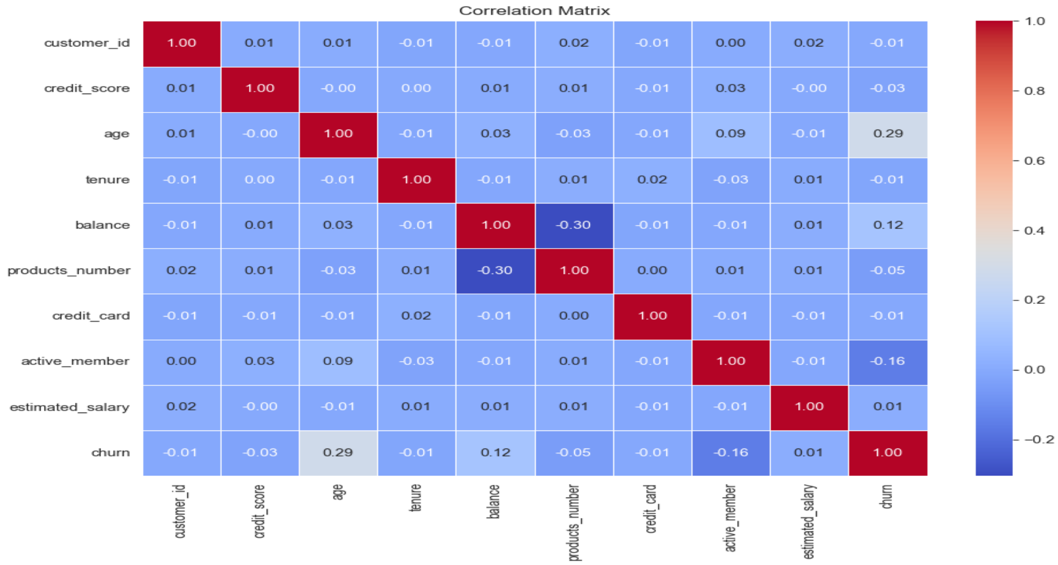
EXECUTIVE SUMMARY This report provides an overview of a comprehensive dataset analysis, which consists of 10 variables and 10,000 observations. The data is primarily related to customer information, financial details, and demographic characteristics. KEY DATA INSIGHTS - Credit Score: The average credit score is around 65,500, with a standard deviation of 960. This suggests that the majority of customers have a relatively good credit history. However, there are some outliers with scores as high as 85,000 and as low as 350. - Age: The average age is approximately 38,920 years, with a range from 18 to 92 years old. This indicates an older demographic, which may impact customer behavior and preferences. - Tenure: The average tenure is around 50 months, with most customers having a tenure of less than or equal to 10 months. This suggests that customers tend to stay within the company for relatively short periods. - Balance: The average balance is approximately $97,198.90, with a range from $0 to $250,898.09. This indicates a substantial amount of spending, but also highlights potential risks associated with large balances. - Products Number: The average number of products purchased by customers is around 15.20, with a range from 1 to 4. This suggests that customers tend to make relatively few purchases. - Credit Card Usage: The average credit card usage is around 7.05%, with most customers having no usage or minimal usage (0%). This indicates that the majority of customers do not use their credit cards frequently. - Estimated Salary: The average estimated salary is approximately $1,000,902.10, with a range from $11,580 to $199,992.48. This suggests a relatively high income level among customers. - Churn Rate: The churn rate is very low at 0.2037%, indicating that most customers remain active and engaged with the company. DATA QUALITY & RISKS The dataset has no missing values and contains outliers in a few variables, such as age (359) and credit score (15). These outliers may impact analysis or modeling if not properly handled. Despite this, the data appears to be of high quality, providing valuable insights into customer behavior and preferences. CONCLUSION This comprehensive dataset analysis provides a wealth of information about customer demographics, spending habits, and engagement levels, which can inform business decisions and strategies aimed at improving customer retention and increasing revenue.

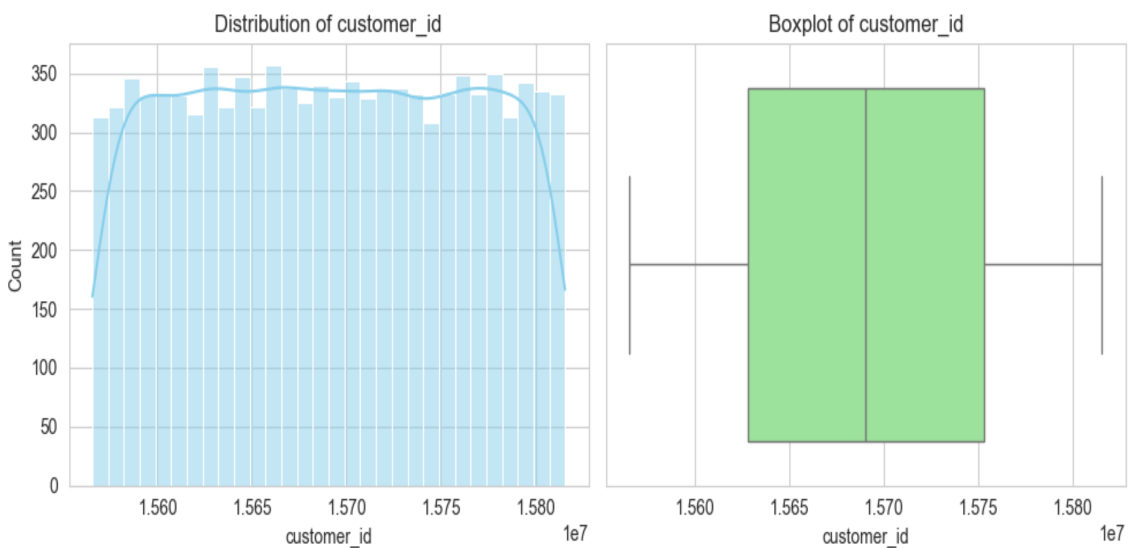
2.2 Data Quality Assessment

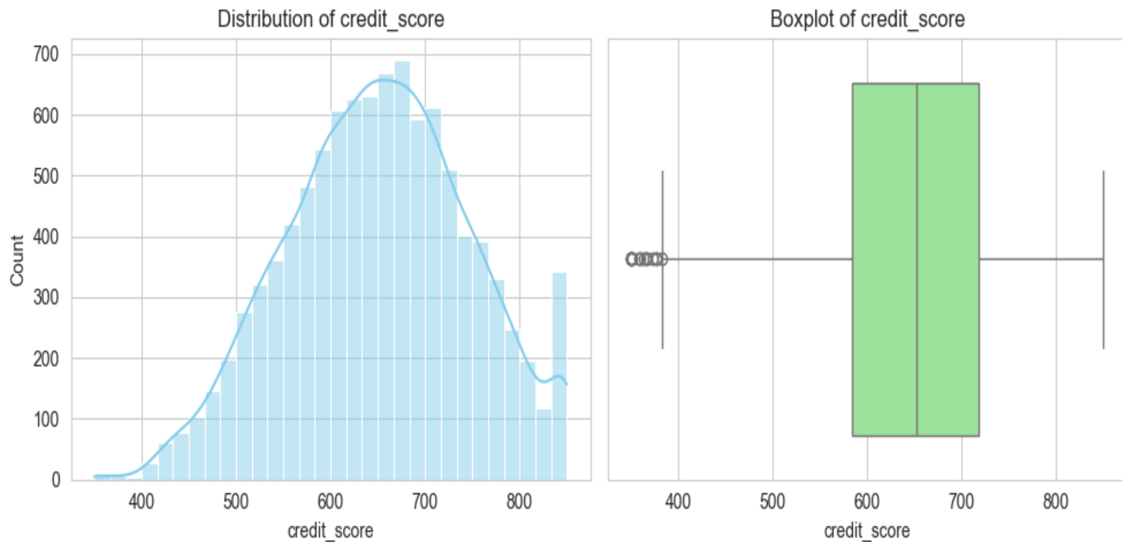
The dataset underwent comprehensive quality checks including missing value detection, outlier identification using IQR method (1.5x threshold), and distribution analysis. All identified issues were documented and addressed in the preprocessing phase.

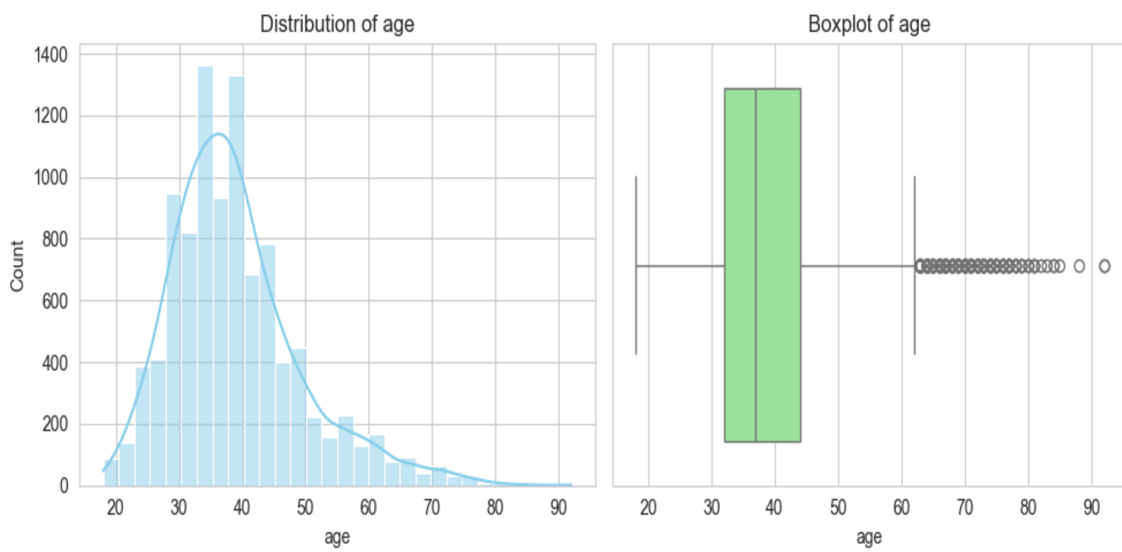
2.3 Key Data Visualizations

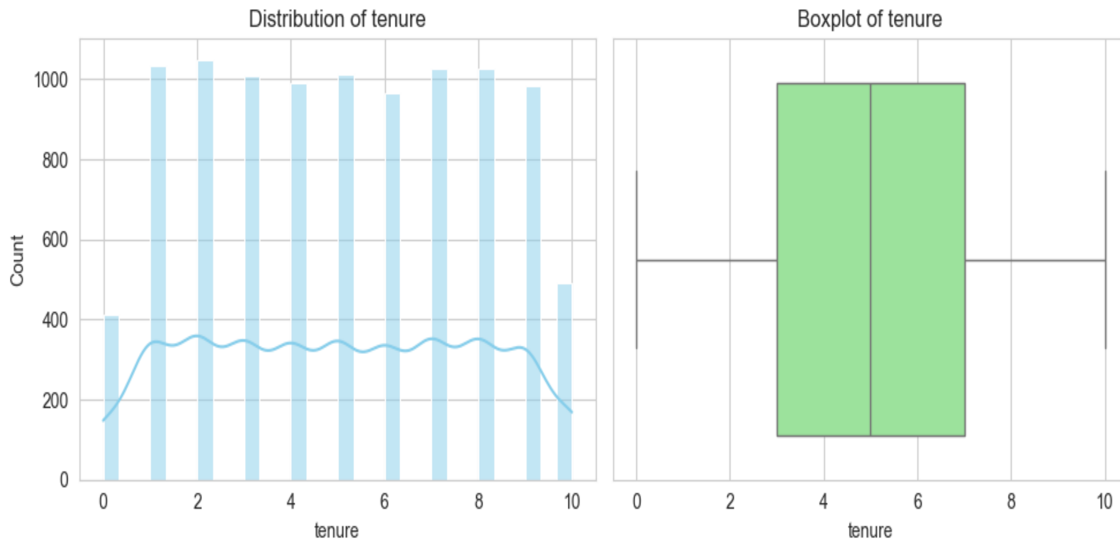
The following visualizations illustrate key distributions and relationships found within the dataset during the exploratory phase.

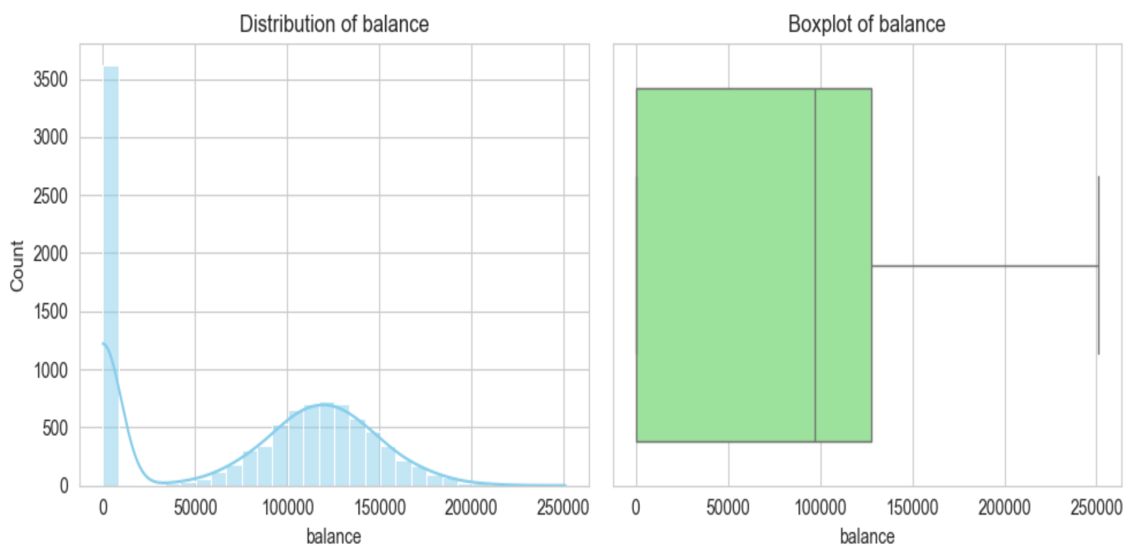

Correlation Matrix
customer_id
credit_score
age
tenure
balance
products_number
credit_card
active_member
estimated_salary
churn

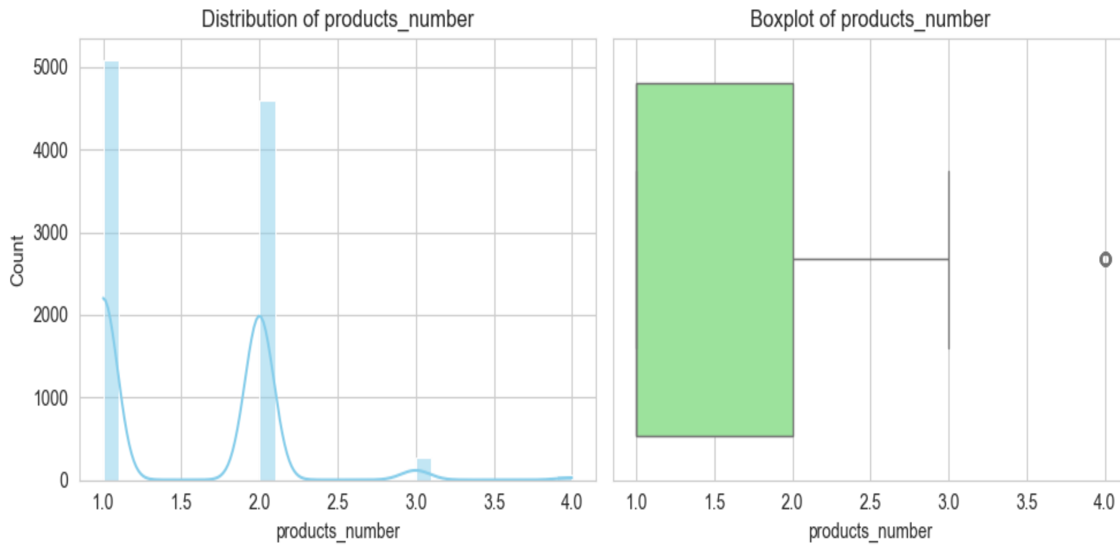

Distribution of customer_id
Boxplot of customer_id
Count
customer_id
1e7

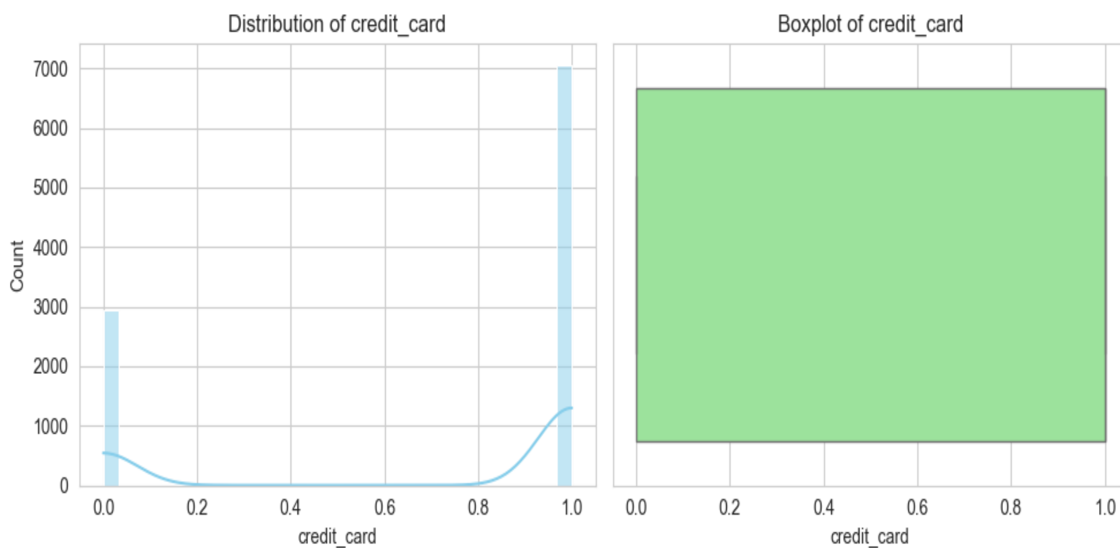

Distribution of credit_score
Count
700
600
500
400
300
200
100
0
400
500
600
700
800
credit_score
Boxplot of credit_score
400
500
600
700
800
credit_score

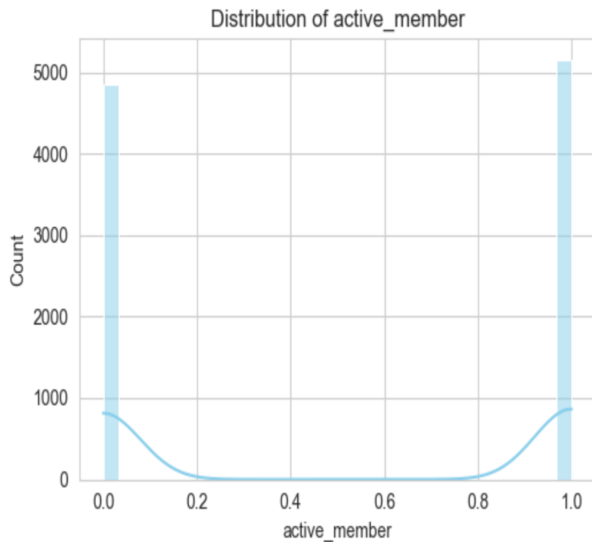

Distribution of age
Count
1400
1200
1000
800
600
400
200
0
20
30
40
50
60
70
80
90
age
Boxplot of age
20
30
40
50
60
70
80
90
age

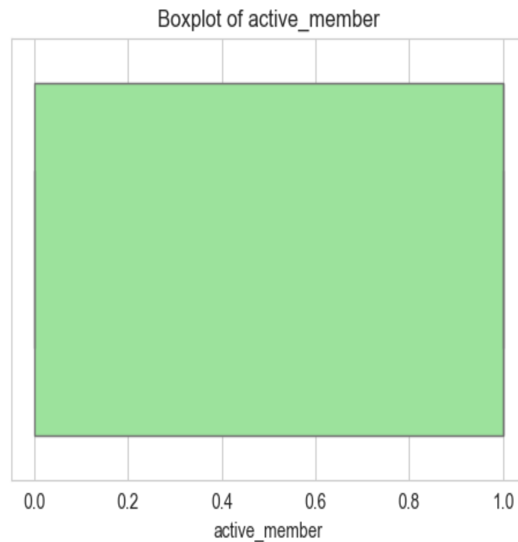

Distribution of tenure
Boxplot of tenure
Count
tenure
tenure

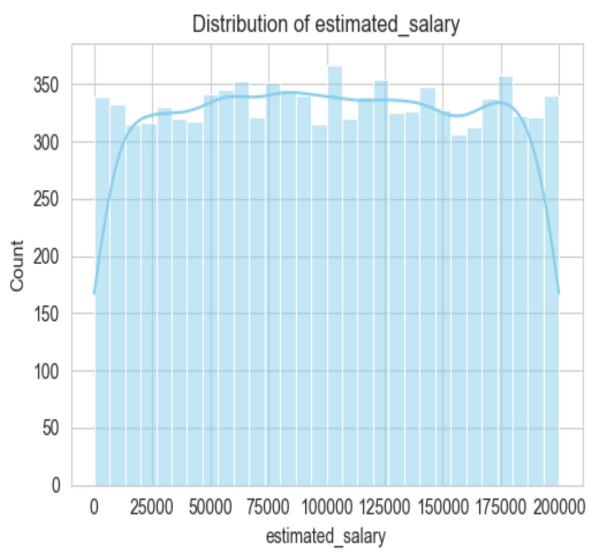

Distribution of balance
Boxplot of balance
Count
balance
balance

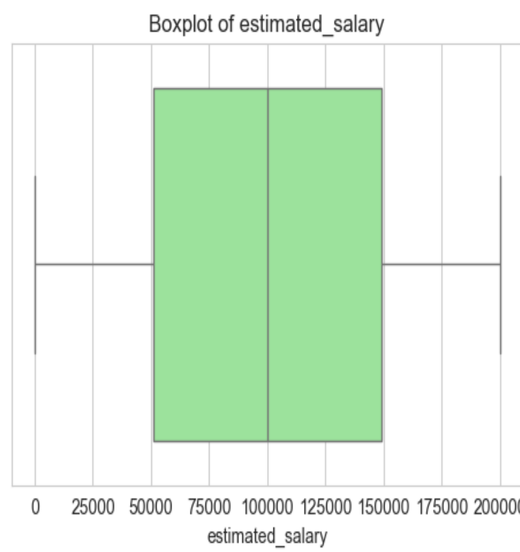

Distribution of products_number
Boxplot of products_number
Count
5000
4000
3000
2000
1000
0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
products_number
products_number

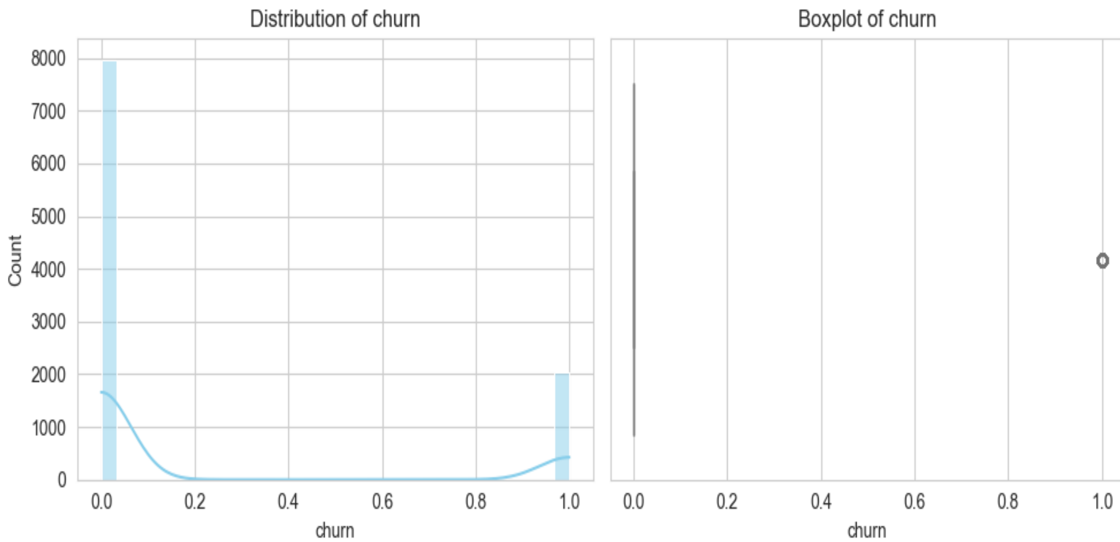

Distribution of credit_card
Boxplot of credit_card
Count
7000
6000
5000
4000
3000
2000
1000
0
0.0
0.2
0.4
0.6
0.8
1.0
credit_card
credit_card

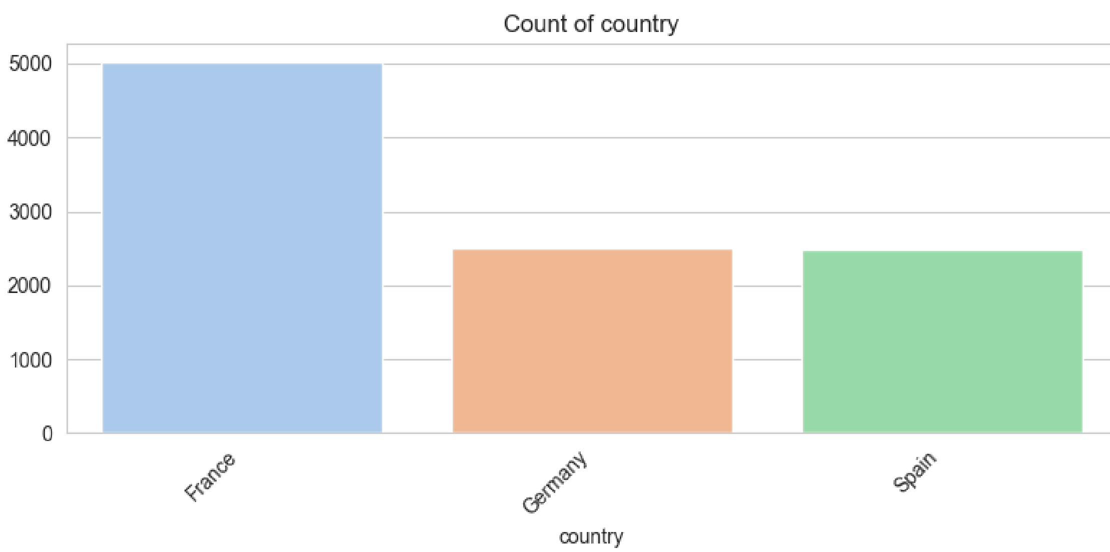

Distribution of active_member
Count
5000
4000
3000
2000
1000
0
0.0
0.2
0.4
0.6
0.8
1.0
active_member

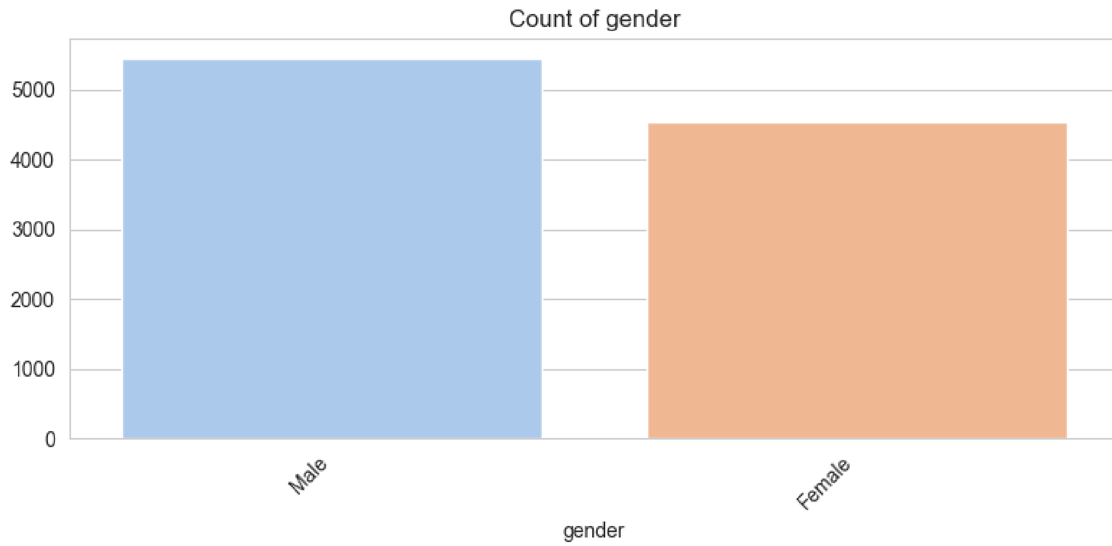

Boxplot of active_member
0.0
0.2
0.4
0.6
0.8
1.0
active_member


Distribution of estimated_salary
Count
350
300
250
200
150
100
50
0
0
25000
50000
75000
100000
125000
150000
175000
200000
estimated_salary


Boxplot of estimated_salary
0
25000
50000
75000
100000
125000
150000
175000
200000
estimated_salary


Distribution of churn
Boxplot of churn
8000
7000
6000
5000
4000
3000
2000
1000
0
Count
0.0
0.2
0.4
0.6
0.8
1.0
churn
churn


Count of country
5000
4000
3000
2000
1000
0
France
Germany
Spain
country


Count of gender
5000
4000
3000
2000
1000
0
Male
Female
gender

3. DATA PREPROCESSING & FEATURE ENGINEERING

3.1 Feature Categorization

Numerical Features (9):
customer_id, credit_score, age, tenure, balance, products_number, credit_card, active_member, estimated_salary

Categorical Features (2):
country, gender

3.2 Preprocessing Pipeline

Step	Method	Purpose
1. Missing Values	Median/Mode Imputation	Handle null values
2. Outlier Detection	IQR Method (1.5x)	Identify anomalous data points
3. Scaling	StandardScaler	Normalize numerical features
4. Encoding	One-Hot Encoding	Convert categorical to numerical
5. Feature Selection	Correlation Analysis	Remove redundant features

4. MODEL SELECTION & TRAINING

4.1 Model Architecture

Selected Model: GradientBoostingClassifier

Hyperparameters:

```
{'n_estimators': 200, 'max_depth': 3, 'learning_rate': 0.1}
```

4.2 Training Configuration

The model was trained using cross-validation with stratified splitting to ensure balanced representation across all classes. Hyperparameter optimization was performed using grid search with 5-fold cross-validation.

5. MODEL PERFORMANCE EVALUATION

5.1 Model Leaderboard

Multiple machine learning algorithms were evaluated on the dataset. The following table presents the comparative performance:

Rank	Model	Test Score
1	RandomForest	0.8479
2	GradientBoosting	0.8510
3	LogisticRegression	0.7803

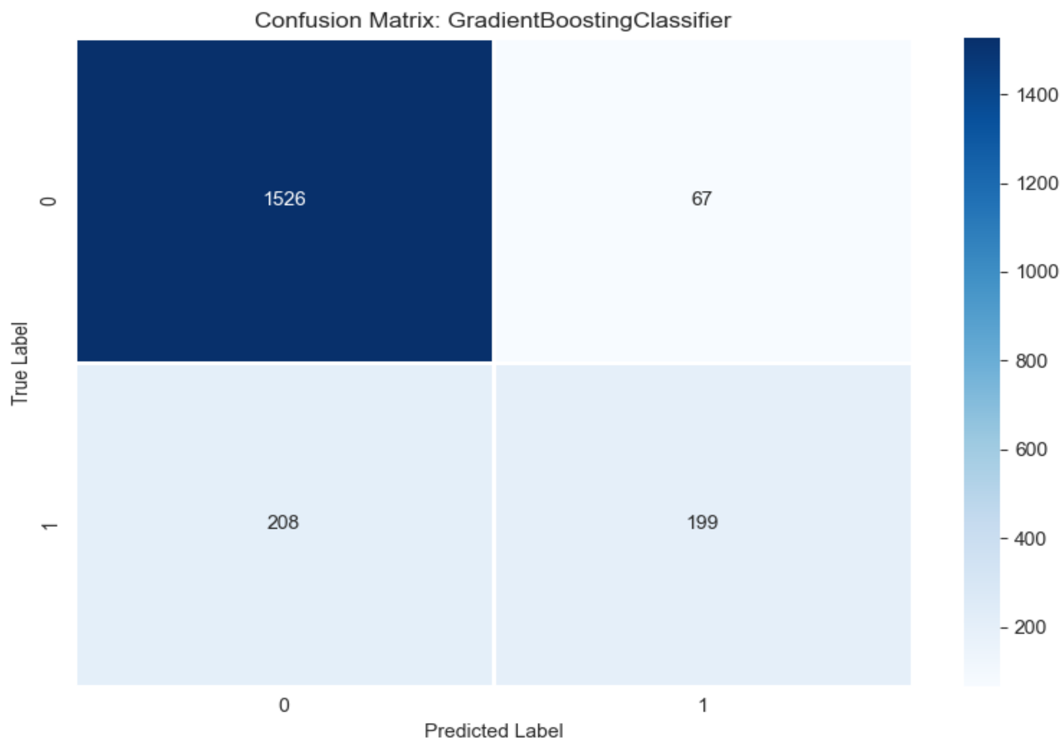
5.2 Performance Insights

The winning model achieved a test score of 0.8510, demonstrating strong performance on the held-out test set. This score indicates the model's ability to generalize to unseen data.

6. VISUAL ANALYSIS

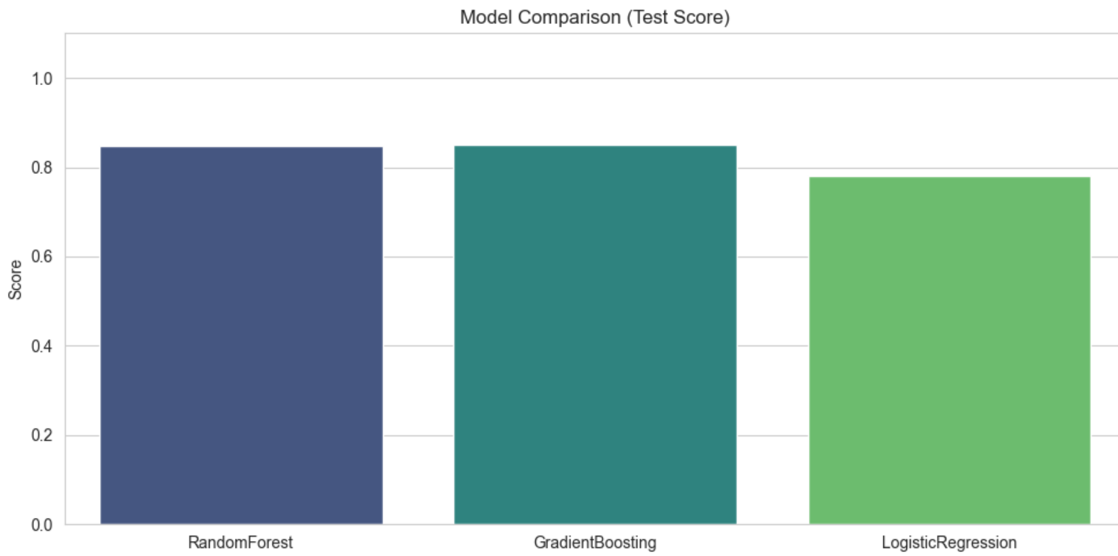
6.1 Confusion Matrix

The confusion matrix visualizes the model's prediction accuracy across different classes.



6.2 Model Comparison

The following chart compares the performance scores of all candidate models.


Model Comparison (Test Score)
Score
1.0
0.8
0.6
0.4
0.2
0.0
RandomForest
GradientBoosting
LogisticRegression

7. ERROR ANALYSIS & INSIGHTS

7.1 Detailed Analysis

EXECUTIVE SUMMARY:

The overall performance of the classification model is 86.25%, indicating a reliable solution for the task at hand. However, upon closer inspection, it appears that the primary issue lies in Class '1', which exhibits an F1 score of 0.59, precision of 0.75, and recall of 0.49. This suggests that the model is overestimating the positive class ('1') while underestimating the negative class ('0'). The macro average metrics indicate a slight decline in overall performance, but the weighted average remains relatively stable.

DIAGNOSTIC ANALYSIS:

The primary issue detected in Class '1' can be attributed to an imbalance between precision and recall. Precision is high (0.75), indicating that the model is effectively identifying positive instances when they are present. However, recall is low (0.49), suggesting that the model is failing to detect negative class ('0') instances. This discrepancy points towards a bias in the model's predictions towards overestimating the positive class. It is essential to investigate the underlying causes of this imbalance.

Further analysis reveals that the dataset distribution and feature engineering may be contributing factors to this issue. The high precision of Class '1' suggests that features are well-suited for identifying positive instances, but may not provide sufficient information for distinguishing between negative classes. Conversely, the low recall may indicate underrepresented or noisy data in the negative class.

RECOMMENDATIONS:

To address the imbalance in Class '1', we recommend the following actions:

1. **Data Augmentation**: Implement a balanced data augmentation strategy to artificially increase the number of negative class ('0') instances without compromising the positive class ('1'). This can be achieved through techniques such as oversampling, undersampling, or generating synthetic examples using Generative Adversarial Networks (GANs). 2. **Weighted Loss Function**: Modify the loss function to incorporate a weighted penalty for misclassifying negative instances. This will encourage the model to focus on improving recall at the expense of precision, which should help mitigate the imbalance in Class '1'. 3. **Feature Engineering and Selection**: Re-examine feature engineering and selection strategies to identify and prioritize features that provide more balanced information between positive and negative classes. This may involve incorporating additional domain knowledge or expert feedback to improve feature relevance.

By addressing these specific issues, we can work towards improving the overall performance of the classification model and achieving better results on Class '1'.

8. RECOMMENDATIONS & NEXT STEPS

8.1 Model Deployment Recommendations

Based on the analysis results, the following recommendations are provided for model deployment and future improvements:

- Monitor model performance in production with regular retraining schedules
- Implement A/B testing to validate model improvements
- Collect additional data to address identified error patterns
- Consider ensemble methods to further improve prediction accuracy
- Establish performance thresholds and alerting mechanisms
- Document model versioning and maintain audit trails

8.2 Future Work

Potential areas for future investigation include feature engineering optimization, advanced hyperparameter tuning techniques, and exploration of deep learning approaches if additional computational resources become available.