Boosting Loyalty: Key Strategies to Reduce **Churn**

Why Customer Retention Is Important

Reducing Marketing Costs

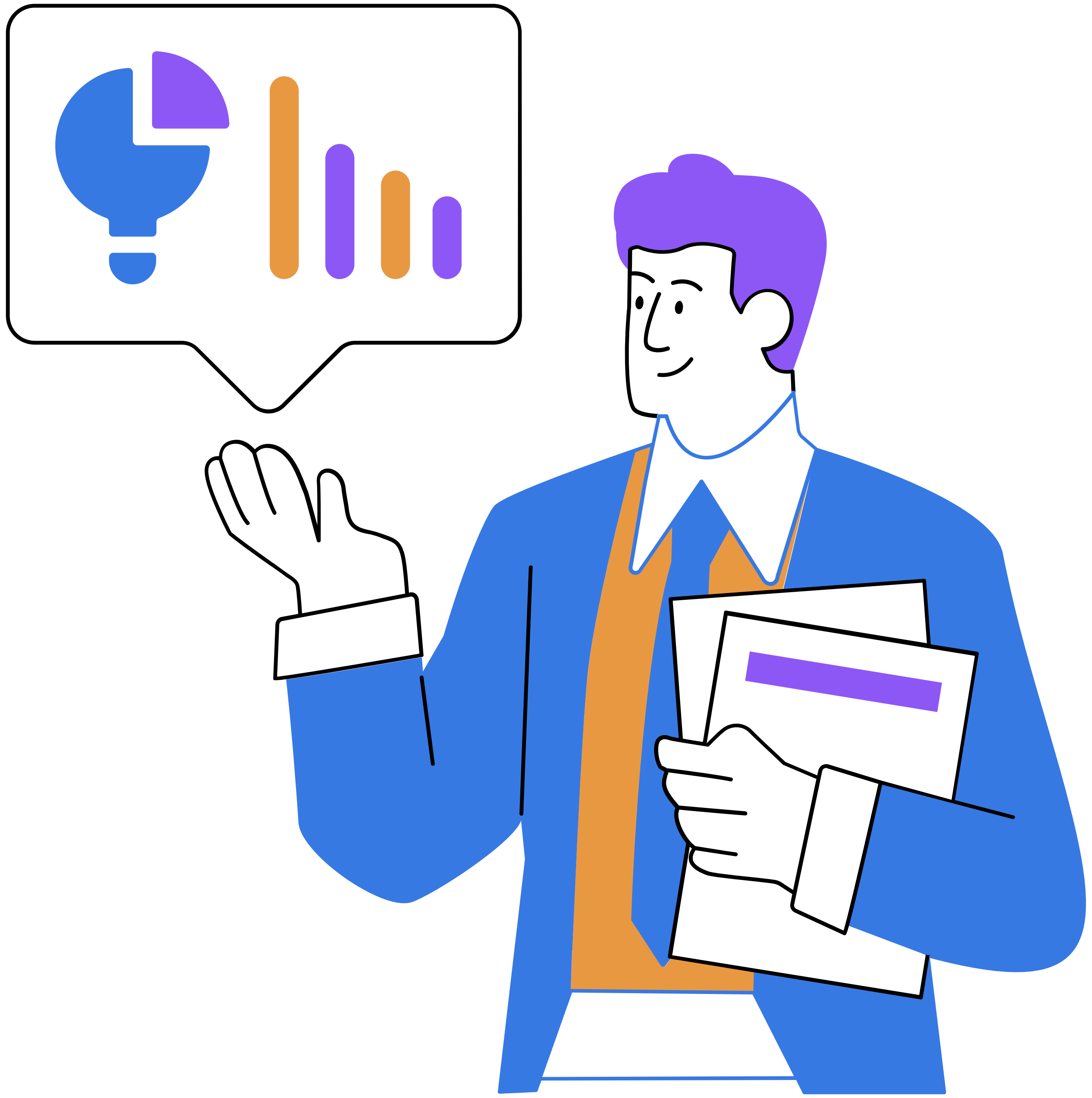
Retaining customers is **5x** cheaper than acquiring new ones, reducing marketing expenses by **up to 70%**.

Increasing Customer Lifetime Value (CLV)

Loyal customers spend 67% more in their third year;
a **5% increase in retention** can boost profits by **25% to 95%.**

Building Competitive Advantage

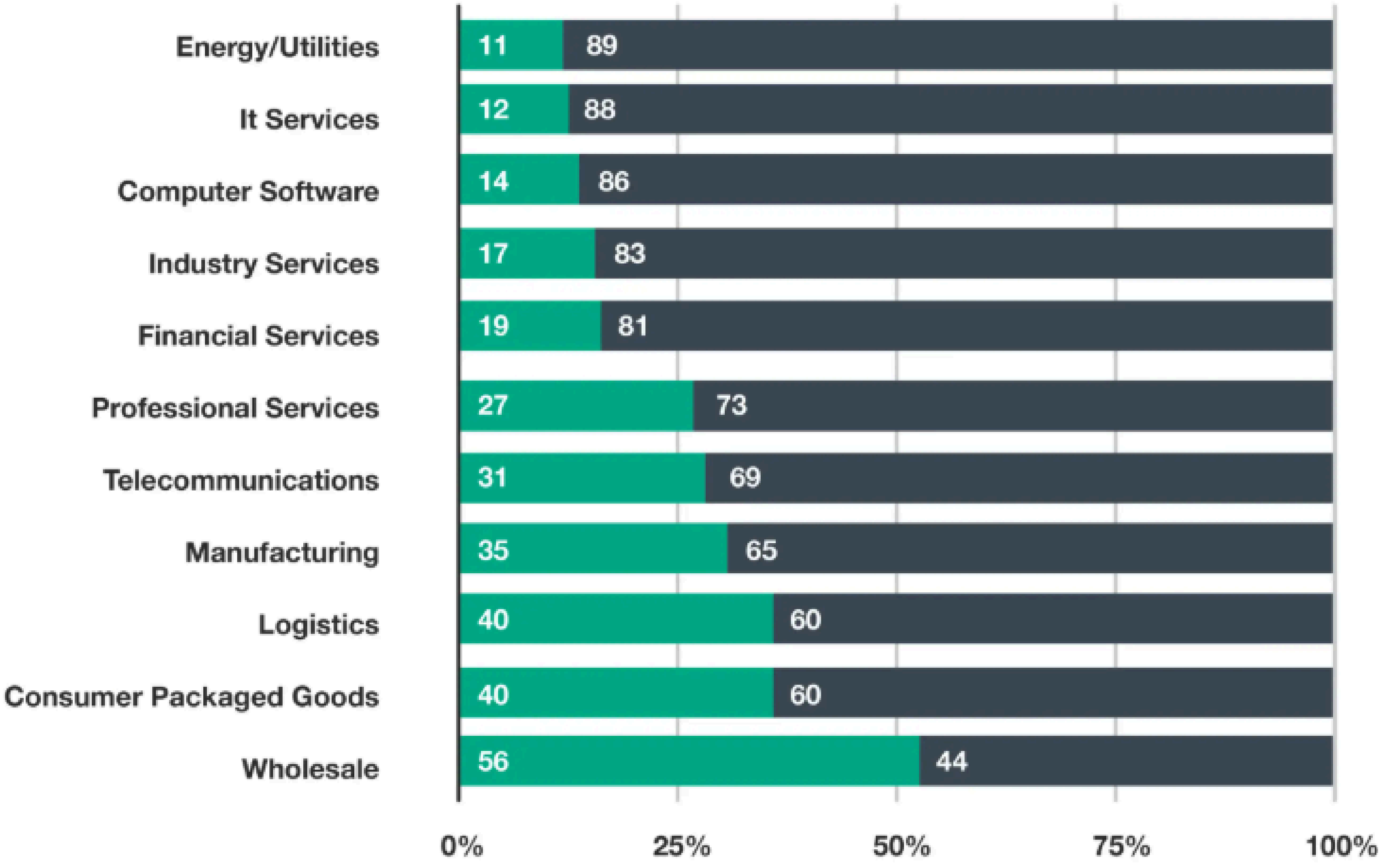
High retention rates lead to **23%** greater share price stability and reduce market share loss by **10% annually**.



Artikel **Harvard Business Review** oleh **Frederick F. Reichheld** yang juga mengungkapkan bahwa biaya retensi lebih rendah berbanding biaya akuisisi.

Studi oleh **Bain & Company** yang sering mengutip bahwa meningkatkan tingkat retensi pelanggan sebesar **5%** dapat meningkatkan keuntungan dari 25% hingga 95%.

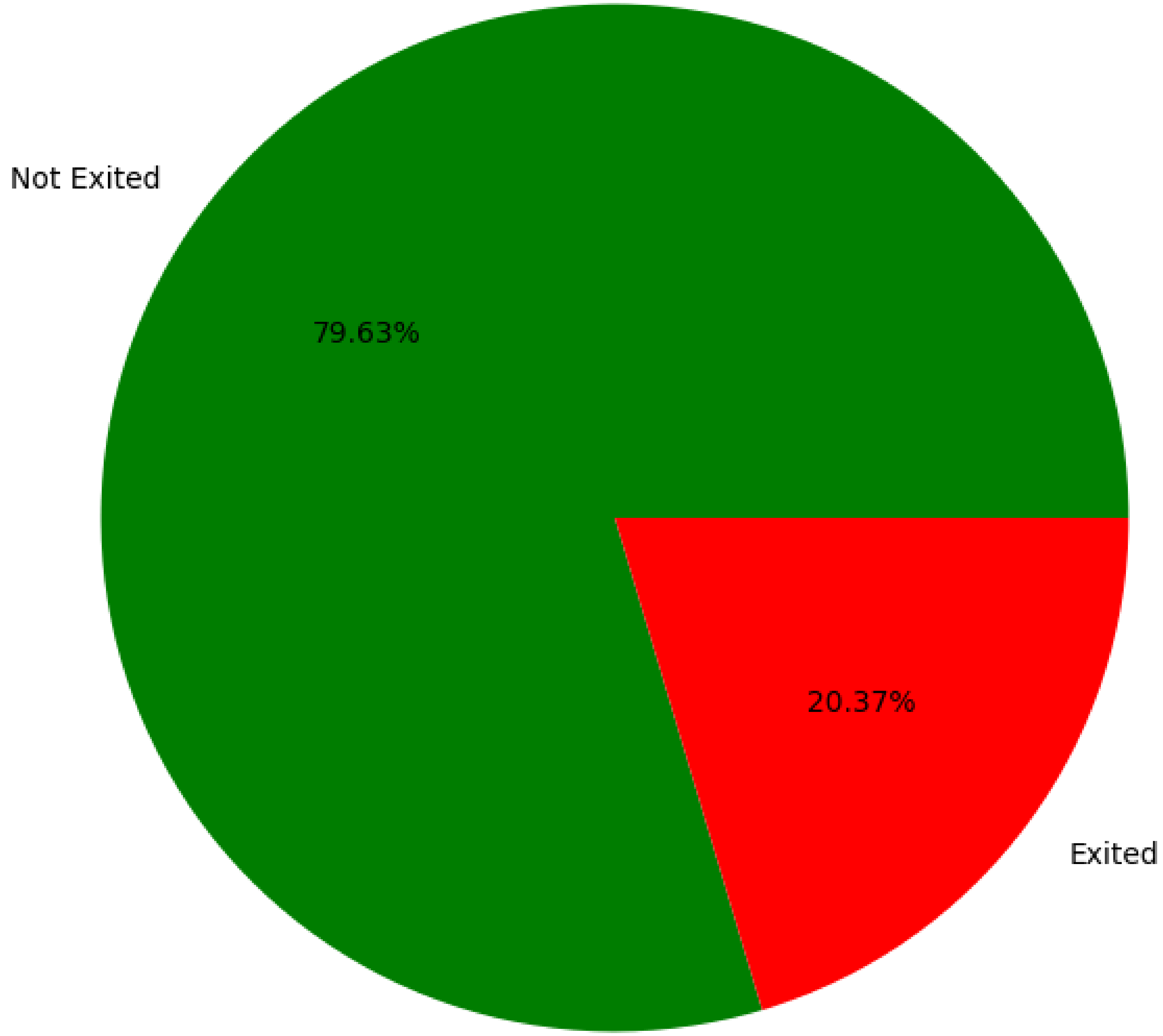
Median Customer Churn Rates by Industry 2024



Financial Services: We found a median customer retention rate of 81% for financial services businesses, and therefore a median churn rate of 19%.

(CustomerGauge, April 18, 2024)

Customer Churn Distribution



20.37%

Customer stop using our product

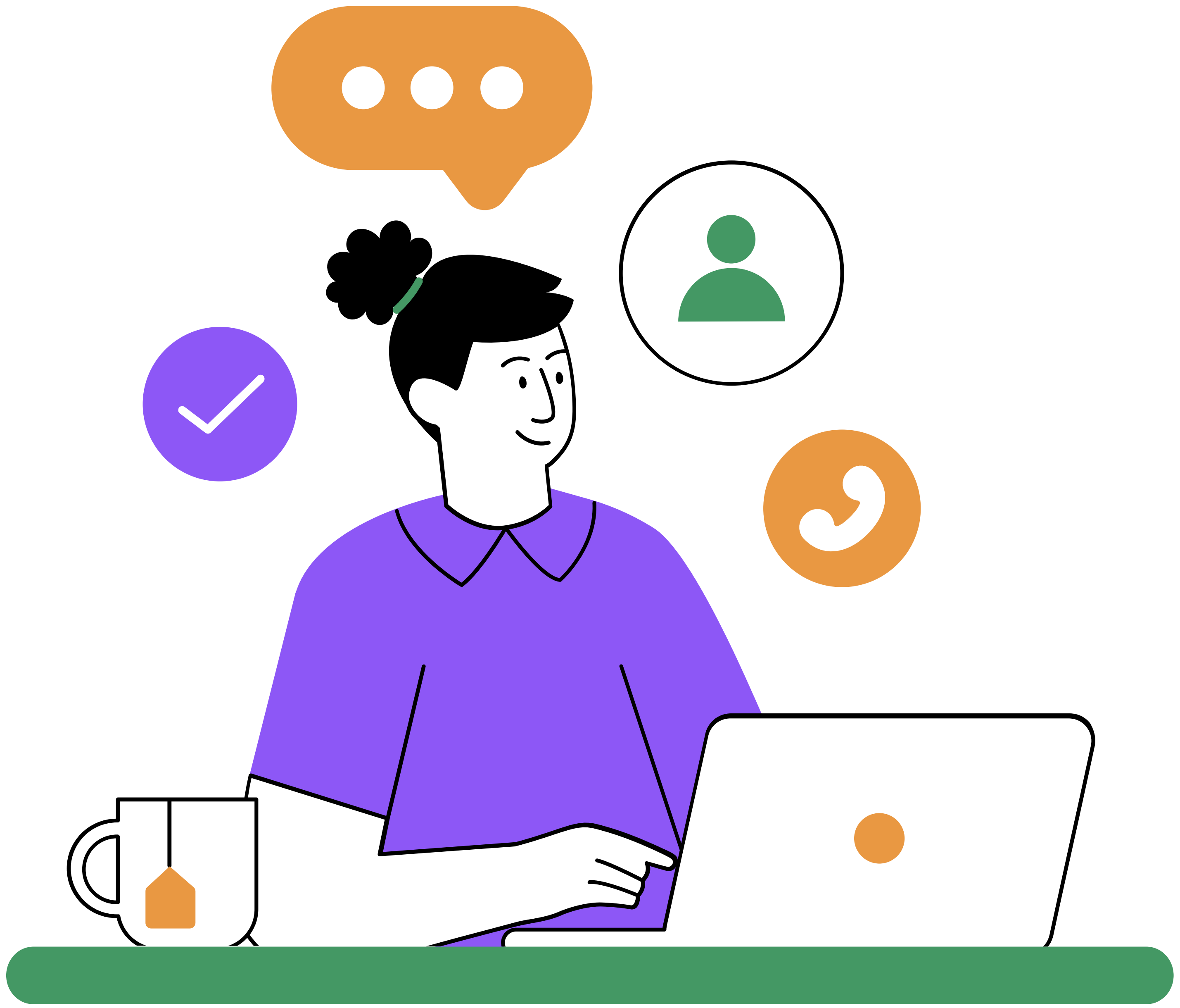
Predictive Churn Model for Banking Customers

This project focuses on developing a predictive model to identify bank customers who are likely to churn.

By predicting churn early, the bank can implement strategic interventions to improve customer retention, which is crucial for maintaining profitability and competitive advantage in a competitive banking environment.

Data Understanding

```
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   RowNumber        10000 non-null  int64
 1   CustomerId       10000 non-null  int64
 2   Surname          10000 non-null  object
 3   CreditScore      10000 non-null  int64
 4   Geography        10000 non-null  object
 5   Gender           10000 non-null  object
 6   Age              10000 non-null  int64
 7   Tenure           10000 non-null  int64
 8   Balance          10000 non-null  float64
 9   NumOfProducts    10000 non-null  int64
 10  HasCrCard        10000 non-null  int64
 11  IsActiveMember   10000 non-null  int64
 12  EstimatedSalary  10000 non-null  float64
 13  Exited           10000 non-null  int64
```



Feature Selection

```
Exited             1.000000
Age                0.285323
Balance            0.118533
EstimatedSalary    0.012097
HasCrCard         -0.007138
Tenure            -0.014001
CreditScore       -0.027094
NumOfProducts     -0.047820
IsActiveMember    -0.156128
```

Key Findings (Churn Factors)

- Age (0.285) → Older customers are more likely to churn.
- IsActiveMember (-0.156) → Inactive customers have a higher churn risk.
- Balance (0.118) → Higher balances slightly increase churn risk.
- NumOfProducts (-0.047) → Customers with more products are less likely to churn.
- CreditScore, Tenure → Minimal impact but may contribute with other factors.
- EstimatedSalary, HasCrCard → No significant correlation with churn.

Recommended Features for Modeling

✔️ Use: Age, IsActiveMember, Balance

✔️ Consider: NumOfProducts, CreditScore

❌ Drop: EstimatedSalary, HasCrCard,Tenure (low impact)

💡 Focus on customer engagement & product bundling to reduce churn! 🚀

📊 Key Takeaway: Geography plays a critical role in customer churn and must be factored into retention strategies! 🌍

```
chi_test('Geography')
✓ 0.0s

chi2 statistic:      301.255
p-value:             3.8303176053541544e-66
degrees of freedom: 2
expected frequencies:
 [[3992.6482 1021.3518]
 [1997.9167  511.0833]
 [1972.4351  504.5649]]
Reject Null Hypothesis
There is a significant association between Geography and Exit.
```

Statistical Significance:

✅ p-value = 3.83e-66 → Strong association between geography and churn

✅ Rejects the null hypothesis → Churn behavior differs by region

Expected Churn Distribution:

🔷 France: Expected Stay: 3,992 | Expected Exit: 1,021

🔷 Germany: Expected Stay: 1,998 | Expected Exit: 511

🔷 Spain: Expected Stay: 1,972 | Expected Exit: 505

🔷 Business Implications

✓ Churn rates vary by region → Localized strategies are needed

✓ Germany has the highest churn risk → Requires urgent retention efforts

✓ Include geography in churn models to improve prediction accuracy

📊 Key Takeaway: Gender is an important predictor in churn modeling! 🚀

```
chi_test('Gender')
✓ 0.0s

chi2 statistic:     112.919
p-value:            2.2482100097131755e-26
degrees of freedom: 1
expected frequencies:
 [[3617.5909  925.4091]
 [4345.4091 1111.5909]]
Reject Null Hypothesis
There is a significant association between Gender and Exit.
```

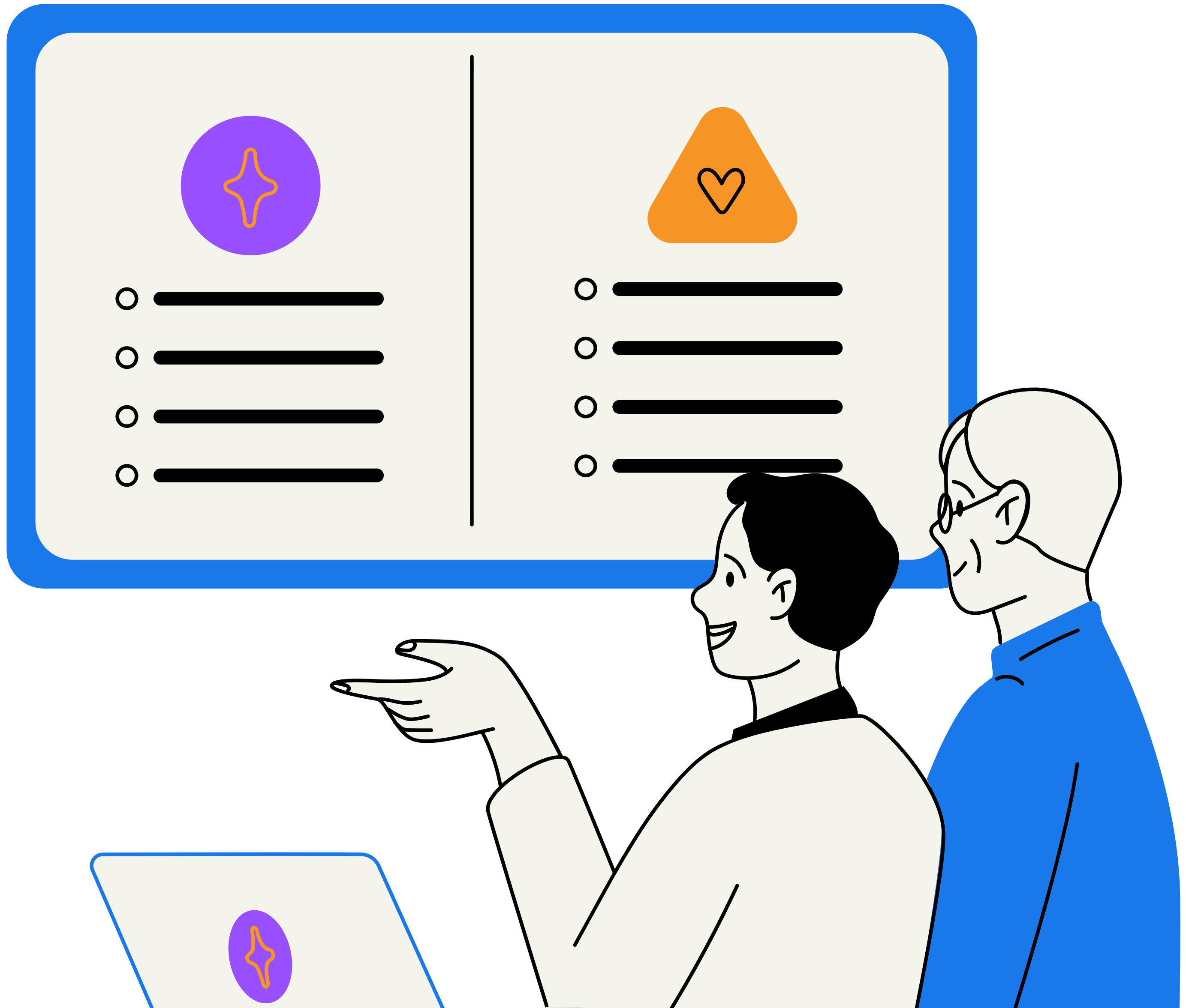
- ✅ p-value = 2.24e-26 → Significant association between gender and churn
- ✅ Rejects the null hypothesis → Gender impacts customer retention

- Expected Churn Distribution:
- 🔹 Male: Expected Stay: 4,345 | Expected Exit: 1,111
- 🔹 Female: Expected Stay: 3,617 | Expected Exit: 925

🔹 Business Implications

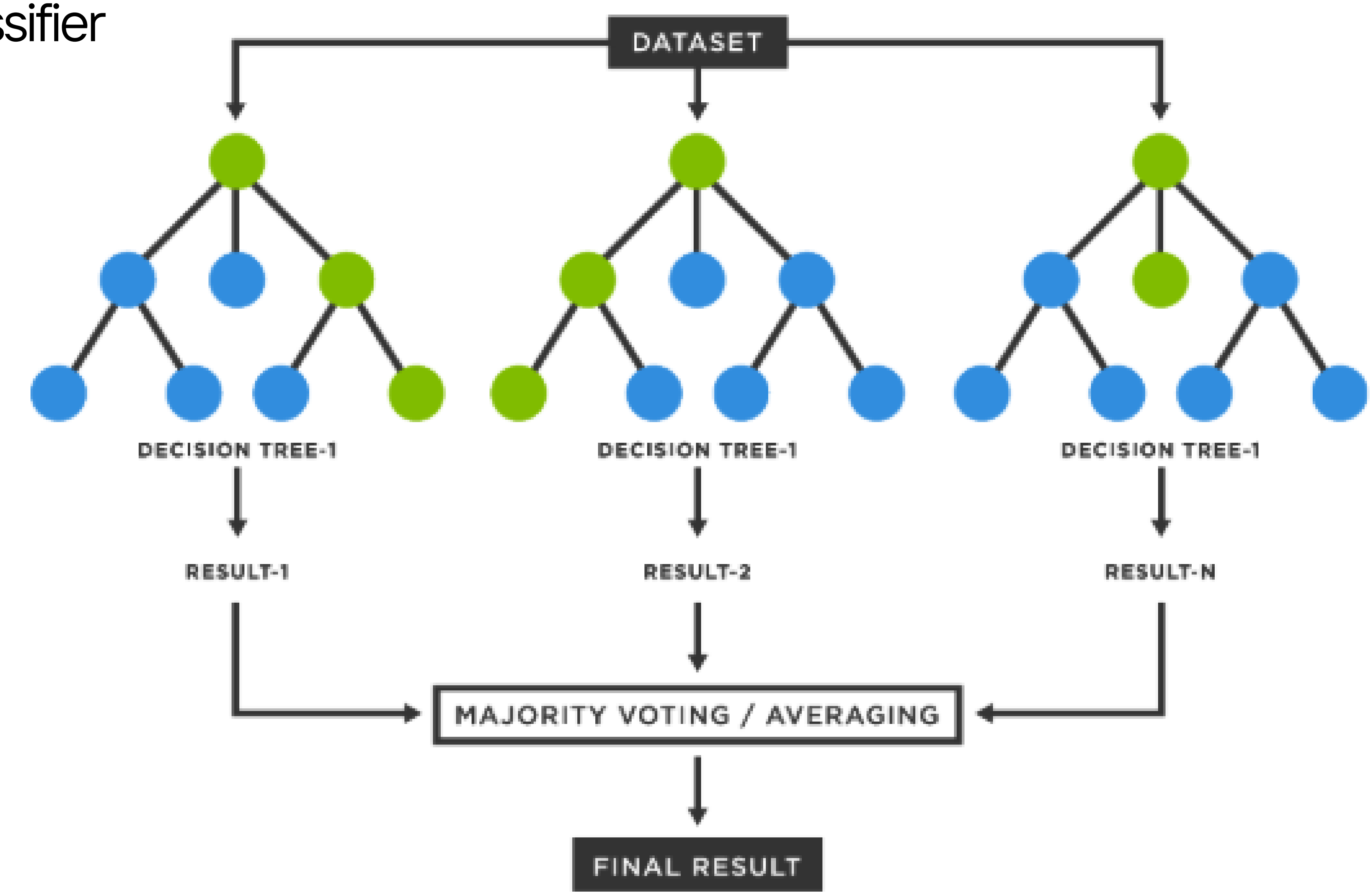
✓ Gender matters in churn prediction → Needs tailored retention strategies

✓ Include gender in churn prediction models for better targeting



Modelling

Random Forest Classifier



Hyperparameters Tuning

Explanation of RandomForestClassifier Parameters

max_depth=10 → Limits the depth of each tree to 10 levels, preventing overfitting.
max_features='sqrt' → Each tree considers √(total features) at each split, improving randomness & reducing correlation between trees.
min_samples_leaf=4 → A node must have at least 4 samples to become a leaf, helping prevent overfitting.
min_samples_split=2 → A node must have at least 2 samples to split, ensuring meaningful splits.
n_estimators=100 → The model uses 100 decision trees, balancing performance and computational efficiency.
class_weight='balanced' → Adjusts class weights inversely proportional to class frequencies, helping with imbalanced data.
random_state=42 → Ensures reproducibility, making results consistent across runs.
✅ Key Benefit: This configuration controls overfitting, improves generalization, and handles imbalanced churn data effectively. 🚀

```
# Initialize the RandomForestClassifier with the optimized number of estimators
rfc = RandomForestClassifier(max_depth= 10, max_features= 'sqrt', min_samples_leaf= 4, min_samples_split= 2, n_estimators= 100,class_weight='balanced',random_state=42)
```

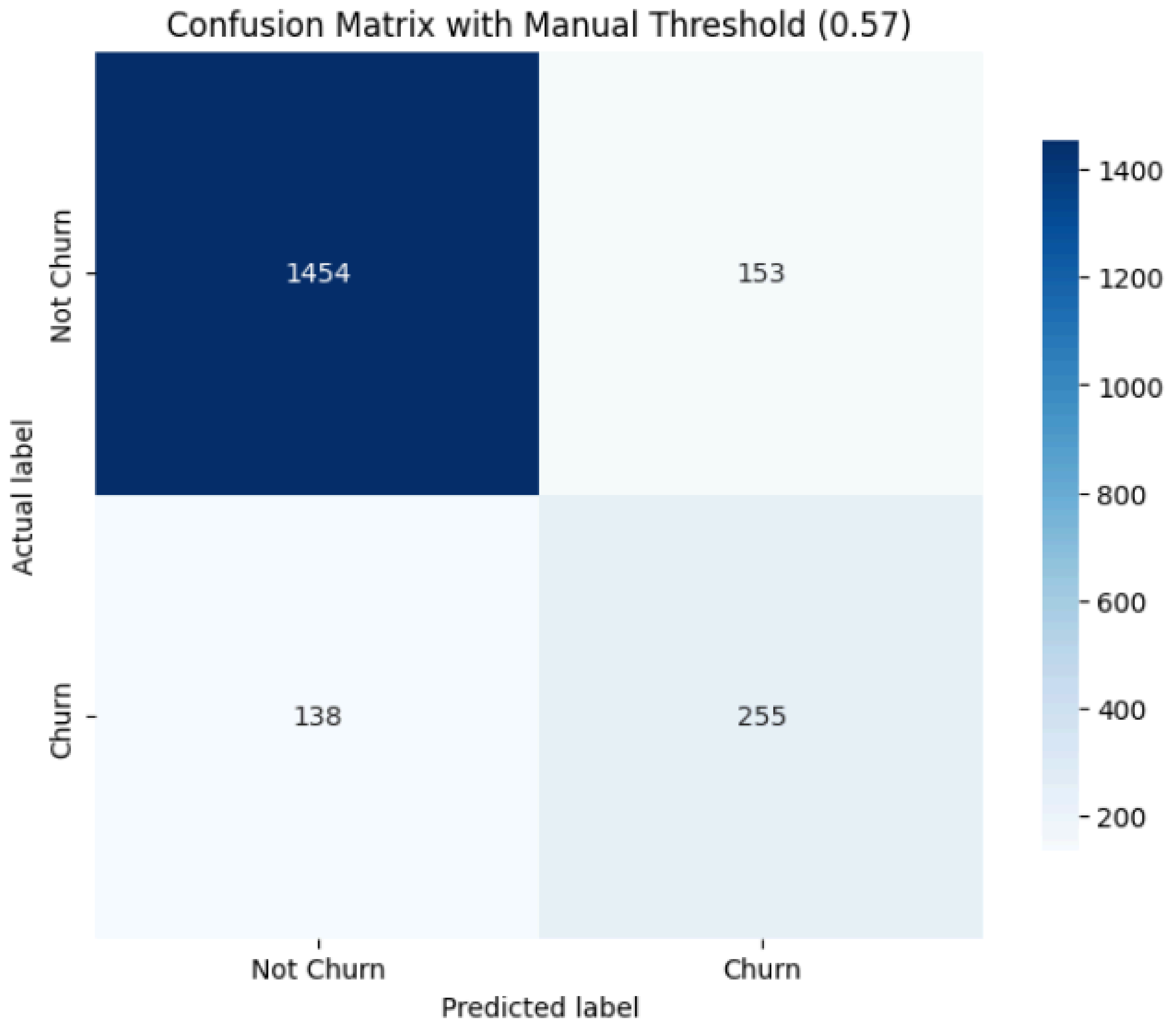
✓ 0.0s

Evaluation

Classification Report for RandomForest Classifier (Wit

	precision	recall	f1-score	support
0	0.91	0.90	0.91	1607
1	0.62	0.65	0.64	393
accuracy			0.85	2000
macro avg	0.77	0.78	0.77	2000
weighted avg	0.86	0.85	0.86	2000

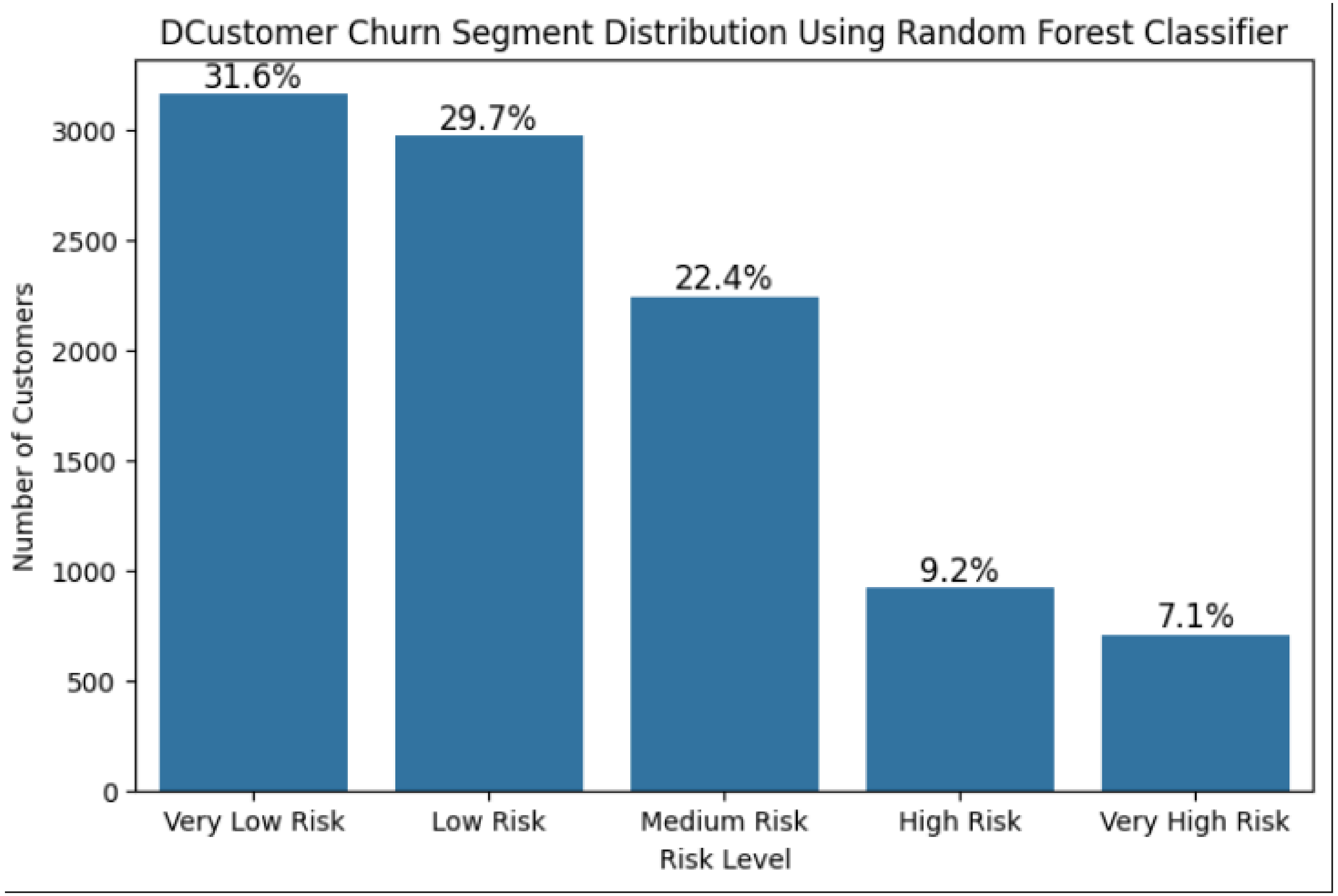
Summary:
Our model correctly predicts most churners and non-churners.
91% of non-churners were correctly predicted.
62% of predicted churners were actually correct.
We captured **65%** of actual churners but missed some.
Overall, the model is **85%** accurate but could be improved to find even more churners.



Result

🔹 Churn Segmentation Summary
📌 Customer distribution based on churn risk (Random Forest Model)

Key Takeaways:
✅ 61.3% are in Low & Very Low Risk → Maintain engagement to keep them loyal.
⚠️ 22.4% are at Medium Risk → Proactive outreach needed to prevent churn.
🚨 16.3% are in High & Very High Risk → Immediate action required to retain these 1,628 customers.



🚀 Retention & Engagement Action Plan

✅ Low-Risk Customers (61.3%) → Keep Engaged & Satisfied

✔ Loyalty Perks – Discounts, cashback, premium services.

✔ Personalized Banking – AI-driven savings & investment suggestions.

✔ Referral & Upselling – Encourage account upgrades & referrals.

✅ Medium-Risk Customers (22.4%) → Prevent Churn Early

✔ Proactive Support – Identify concerns & engage before issues escalate.

✔ Targeted Offers – Fee waivers, adjusted loan rates, flexible plans.

✔ Better Communication – Personalized messages & dedicated managers.

✅ High-Risk Customers (16.3%) → Urgent Retention Measures

✔ Immediate Outreach – Direct calls/emails with retention incentives.

✔ Special Packages – Fee reductions, better loan rates, priority support.

✔ Exit Survey – Capture churn reasons to improve future retention.

📊 Business Impact

✔ 10-15% churn reduction → Protects revenue & customer base.

✔ Higher satisfaction & loyalty → Increases customer lifetime value.

✔ Lower acquisition costs → Retaining is 5x cheaper than acquiring new customers.

“If your retention is poor then nothing else matters.”

– Brian Balfour,
Founder/CEO of Reforge, former VP Growth at HubSpot

Thank You

@abiltio