Investment plan to nurture client relationships.

5% increase in retention

> 25% increase in profits

The need

Increase customer retention rates for EuroBank.

The data

Year’s closing clients data, including:

Balance, number of products, salary, tenure, country of residency, is active member.

Solution Path

- Explore and understand the data.
- Create baseline models for solution, compare, improve.
- Provide recommendations to increase customer retention by 5% (next two years).

Some considerations

- Client leaves — EuroBank - 180 € / year.
- Keeping a client — EuroBank + 90 € / year.
- Misclassifying a client — EuroBank - 90 € / year.

Actual scenario

From a random sample of 2000 clients:

393 Left

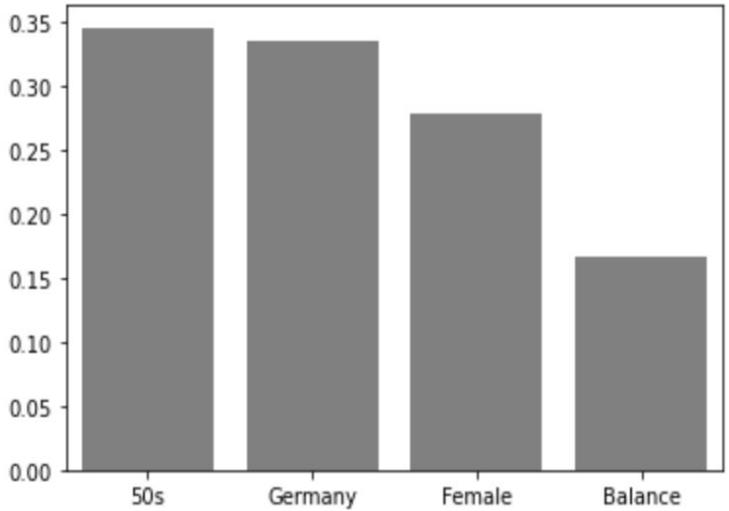
>70K €

By using the best baseline model

We are able to:

- Identify 161 at risk clients.
- Save 34k €

Results



Most interpretable model.

Saves > 15% compared to base model.

Complicated but better model.

Saves > 22 % compared to base model.

Recommendations

- Invest in surveys and special deals for customers in their 40s and 50s.
- Study the difference between services offered in Germany, compared to the rest of countries.
- Offering VIP services for clients with balance above the third quartile.
- Choose and implement one of the solutions.

Future Work

- Study customer satisfaction.
- Include data from more than a year.
- Obtain monthly ending balances.
- Generate a more robust ensembled model for predictions.

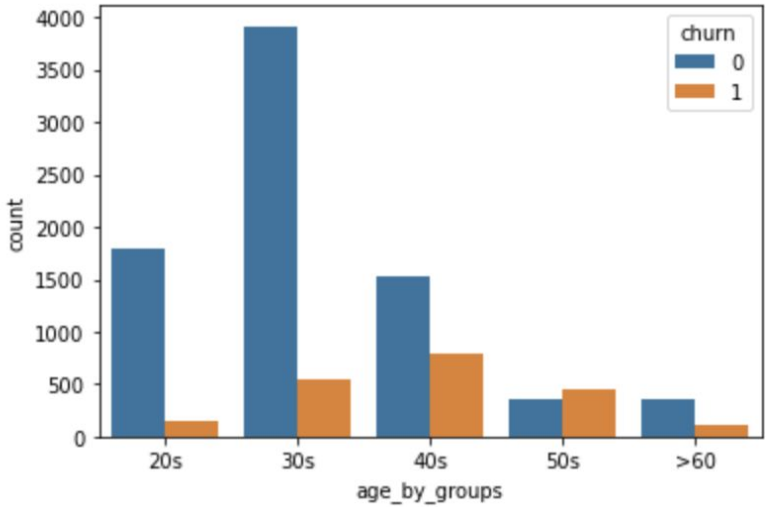
Sources:

https://www.forbes.com/sites/jiawertz/2018/09/12/dont-spend-5-times-more-attracting-new-customers-nurture-the-existing-ones/?sh=3417ab615a8e

https://unsplash.com/

Appendix

Churn based on age groups.



KNN baseline Confusion Matrix.

```
              precision    recall  f1-score   support

           0       0.87      0.94      0.90      1607
           1       0.63      0.41      0.50       393

    accuracy                           0.84      2000
   macro avg       0.75      0.67      0.70      2000
weighted avg       0.82      0.84      0.82      2000

[[1511   96]
 [ 232  161]]
```

Final Logistic Regression Confusion Matrix.

	precision	recall	f1-score	support
0	0.89	0.88	0.89	1607
1	0.53	0.56	0.55	393
accuracy			0.82	2000
macro avg	0.71	0.72	0.72	2000
weighted avg	0.82	0.82	0.82	2000

```
array([[1414,  193],
       [ 172,  221]])
```

Final RF Confusion Matrix.

	precision	recall	f1-score	support
0	0.88	0.97	0.92	1607
1	0.77	0.45	0.57	393
accuracy			0.87	2000
macro avg	0.82	0.71	0.74	2000
weighted avg	0.86	0.87	0.85	2000

```
[[1553   54]
 [ 215  178]]
```