Potential Customers Prediction to Open Term Deposits in Bank Marketing Campaign

- **Problem:** The bank faced challenges with costly and inefficient marketing campaigns, 3 years of marketing only resulted in a conversion rate of 13%.

- **Proposed Solution:** Developed targeted customer segmentation strategies and a predictive model to assess the likelihood of customer interest in term deposit offers, resulting in more efficient and effective campaigns.

- **Goals:** Quickly identify potential customers for targeted outreach, design campaign strategies to increase conversion rates, and ultimately boost company profits by acquiring new customers and increasing third-party funds.

PROJECT LINK

Approach

- Conducted **bivariate and multivariate analysis** using Python and various libraries to uncover actionable patterns in each variable.
- **Cleaned data** and **performed feature engineering** to prepare it for analysis and modelling.
- **Developed a predictive model** to estimate the likelihood of customer interest in term deposit offers, allowing the company to streamline marketing efforts and improve decision-making for campaign targeting.

Tools

- **Programming Language:** Python
- **Tools:** Jupyter Notebook
- **Libraries:** Pandas, NumPy, Matplotlib, Scikit-Learn
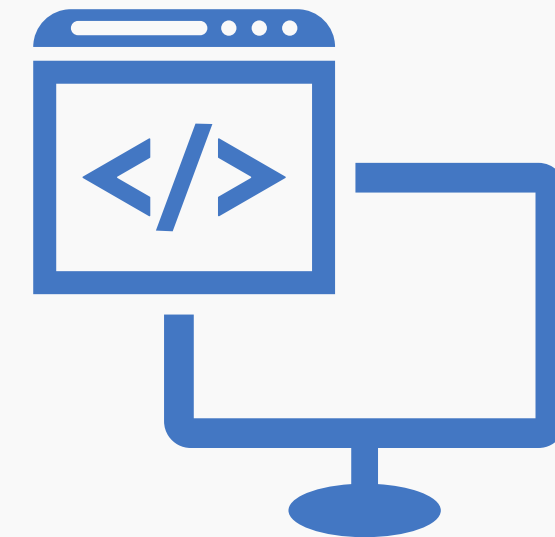
Datasets

- Variable information includes **customer data**, **last contact**, **current campaign**, and **macroeconomic** data.

- Dataset has **41288 rows** and **21 variables**.

- Data was created by Sérgio Moro (ISCTE-IUL), Paulo Cortez (Univ. Minho), and Paulo Rita (ISCTE-IUL), **collected from May 2008 to November 2010**.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 41188 entries, 0 to 41187
Data columns (total 21 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   age             41188 non-null   int64
 1   job             41188 non-null   object
 2   marital         41188 non-null   object
 3   education       41188 non-null   object
 4   default         41188 non-null   object
 5   housing         41188 non-null   object
 6   loan            41188 non-null   object
 7   contact         41188 non-null   object
 8   month           41188 non-null   object
 9   day_of_week     41188 non-null   object
 10  duration        41188 non-null   int64
 11  campaign        41188 non-null   int64
 12  pdays           41188 non-null   int64
 13  previous        41188 non-null   int64
 14  poutcome        41188 non-null   object
 15  emp.var.rate    41188 non-null   float64
 16  cons.price.idx  41188 non-null   float64
 17  cons.conf.idx   41188 non-null   float64
 18  euribor3m       41188 non-null   float64
 19  nr.employed     41188 non-null   float64
 20  y               41188 non-null   object
dtypes: float64(5), int64(5), object(11)
memory usage: 6.6+ MB
```
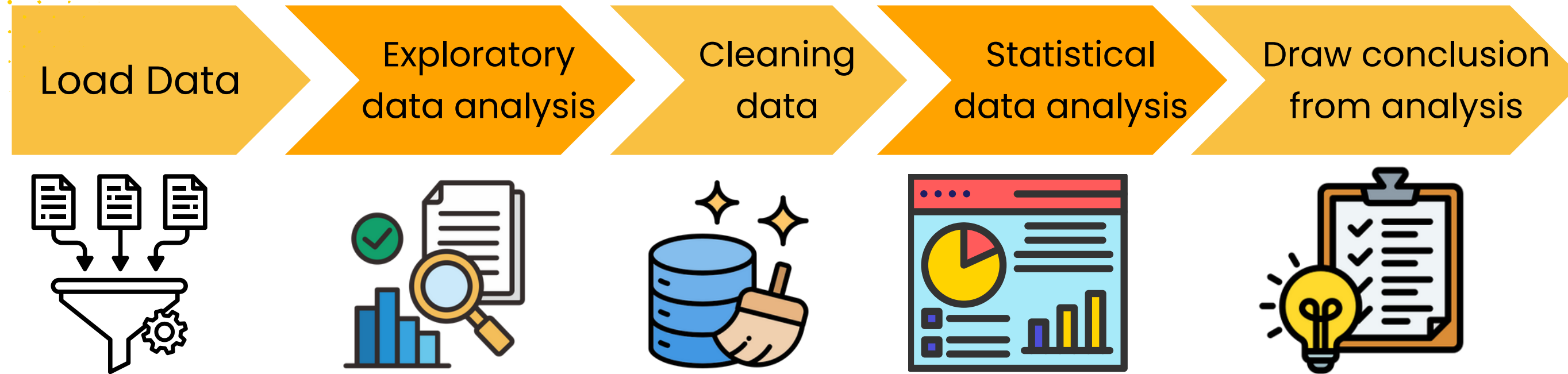
Research

- Based on statista.com, the average telemarketing cost in Portugal is €1.50 per call.
- According to dooly.ai, achieving a 70% contact rate requires sales reps to call prospects 6 times, resulting in a total cost of €9 per person.
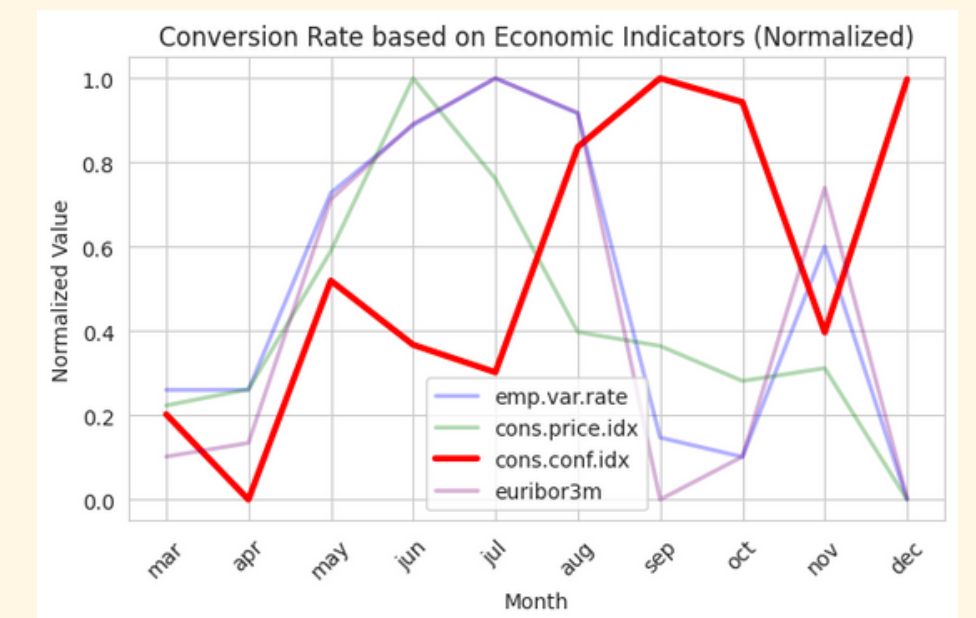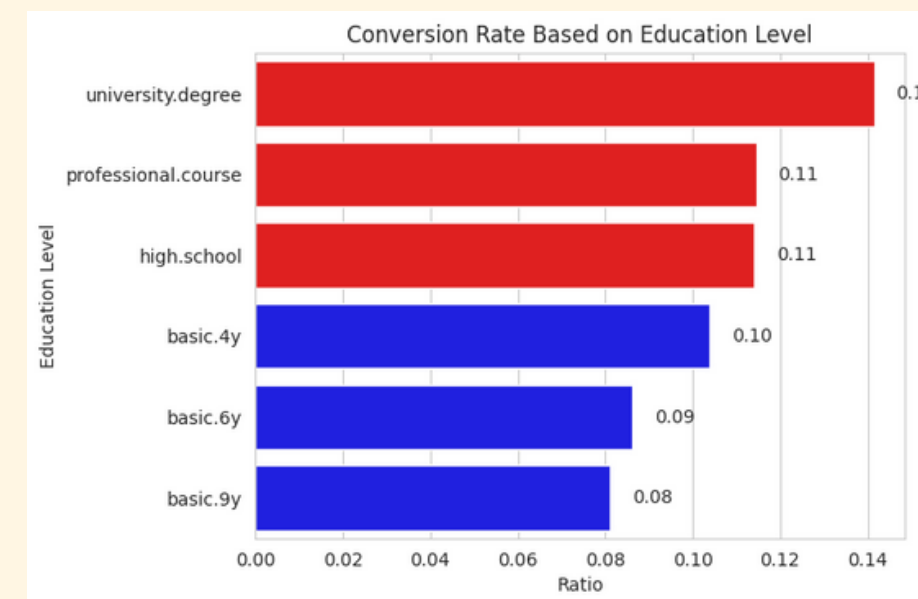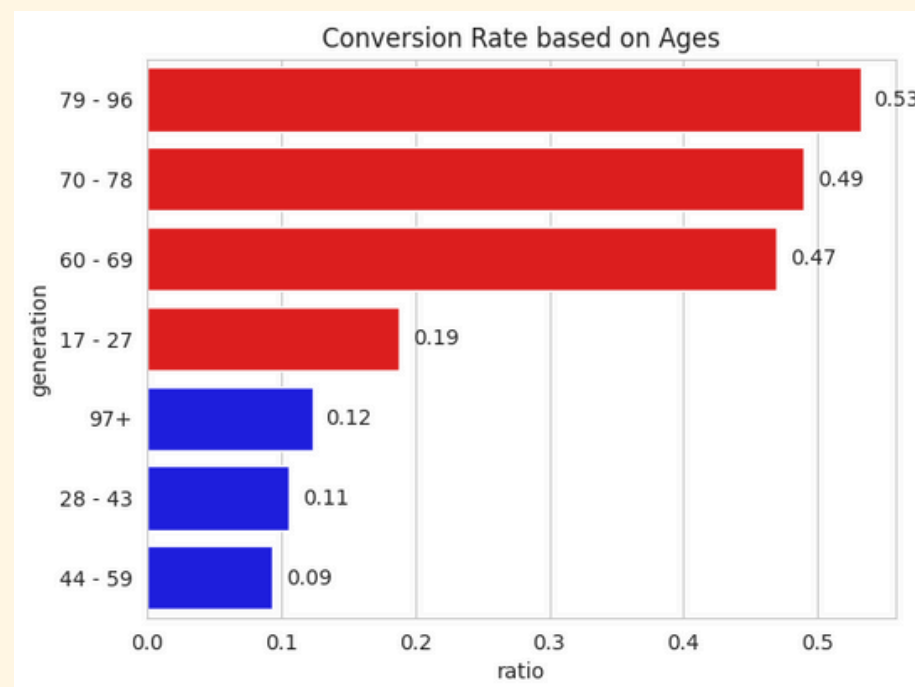
- In 2014, deposit interest was 0.15%, and lending interest was 5.59%.
- Minimum term deposit at ABANCA PT ranged from €100 to €500.
- Most job in data had an average annual gross salary of €14,940, with living costs around €8,940.
- Disposable income estimated at €6,000; 60% (€3,600) allocated to term deposits for conservative investors.

- For a €3,600 deposit held for one year, the company earns a profit of €196.

- Formula: Profit = Deposit × (Interest Income - Interest Expense)

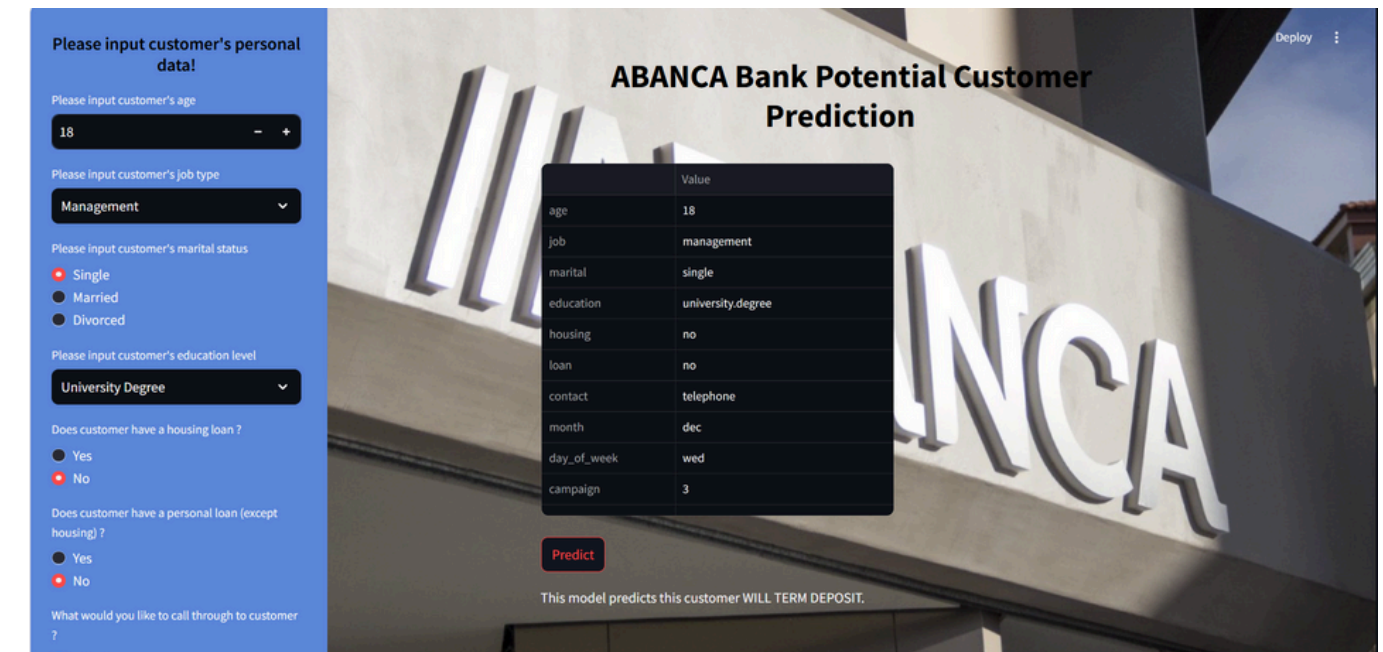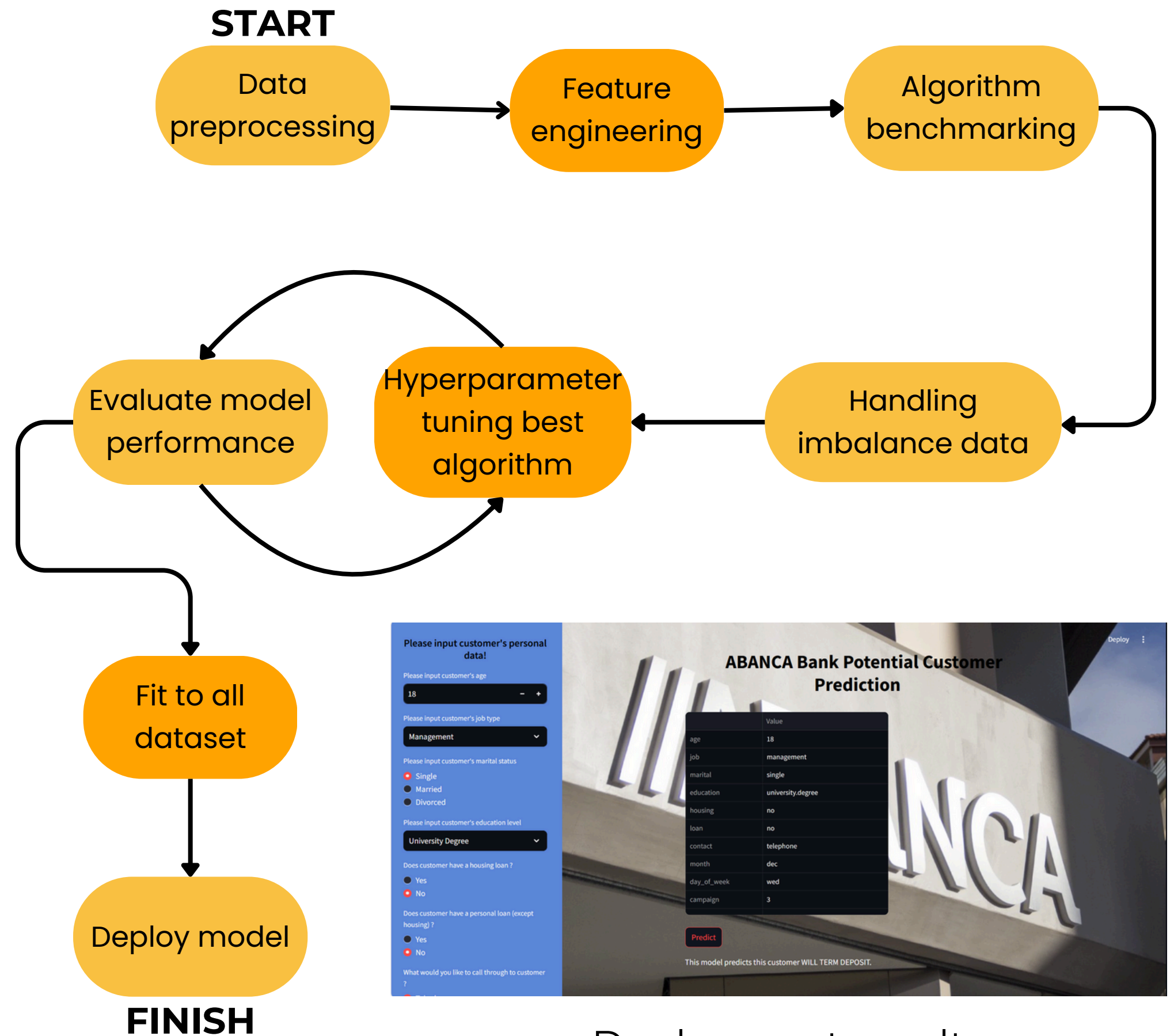- Profits: €3,600 × (5.59% – 0.15%) = €196

The Process

| Load Data | Exploratory data analysis | Cleaning data | Statistical data analysis | Draw conclusion from analysis |

Three key visualization examples

Conversion Rate based on Ages

generation	ratio
79 - 96	0.53
70 - 78	0.49
60 - 69	0.47
17 - 27	0.19
97+	0.12
28 - 43	0.11
44 - 59	0.09

Conversion Rate Based on Education Level

Education Level	Ratio
university.degree	0.14
professional.course	0.11
high.school	0.11
basic.4y	0.10
basic.6y	0.09
basic.9y	0.08

Conversion Rate based on Economic Indicators (Normalized)
- emp.var.rate
- cons.price.idx
- cons.conf.idx
- euribor3m

Building Model Process

- The **entire modeling process**, from data cleansing to algorithm selection, was **tailored to business needs** to ensure relevance and impact.

- The **model** achieved a recall score of 78%, **effectively identifying potential customers** with a high likelihood of conversion.

```
Skor Data Testing
              precision    recall   f1-score    support

          0       0.95       0.55       0.70       6413
          1       0.18       0.78       0.29        809

   accuracy                             0.58       7222
  macro avg       0.57       0.66       0.49       7222
weighted avg       0.87       0.58       0.65       7222
```

START

Data preprocessing → Feature engineering → Algorithm benchmarking

Evaluate model performance ⇄ Hyperparameter tuning best algorithm ← Handling imbalance data

Fit to all dataset

Deploy model

FINISH

Deployment result

Model Transparency

To ensure transparency and interpretability of the predictive model, I utilized SHAP (SHapley Additive exPlanations) values. This approach helps explain the influence of each feature on the model's predictions, providing clear insights for stakeholders and building confidence in how the model makes decisions.

How to Read

- **Prediction:** 68.78% likelihood of term deposit interest.
- **Key Factors:** Age (31), education (university degree), employment levels, consumer confidence, interest rates, and inflation.
- **Visual Impact:** SHAP values highlight positive (red) and negative (blue) contributions, ensuring transparency and trust in the model's predictions.
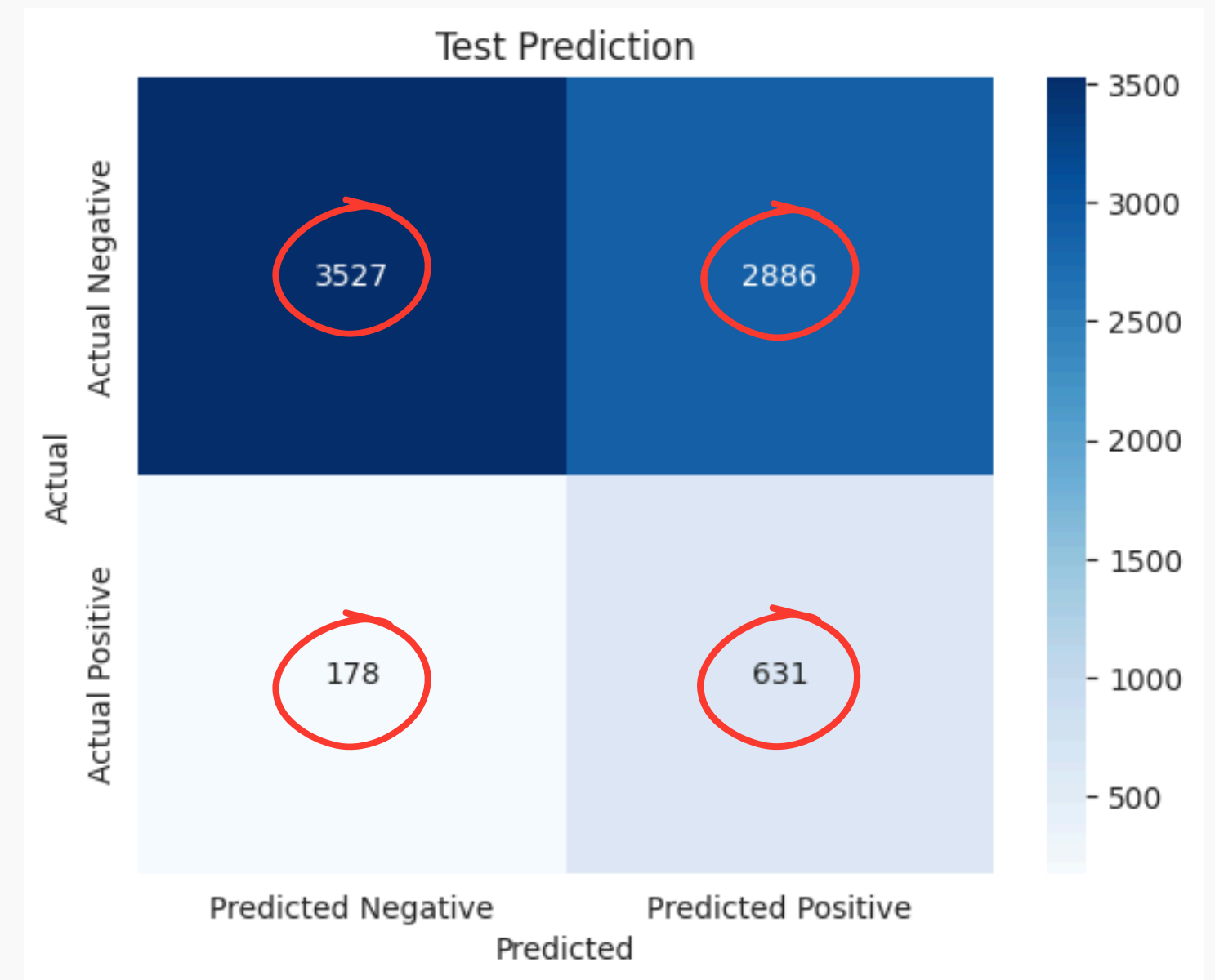
Prediksi orang ini menempatkan deposito berjangka
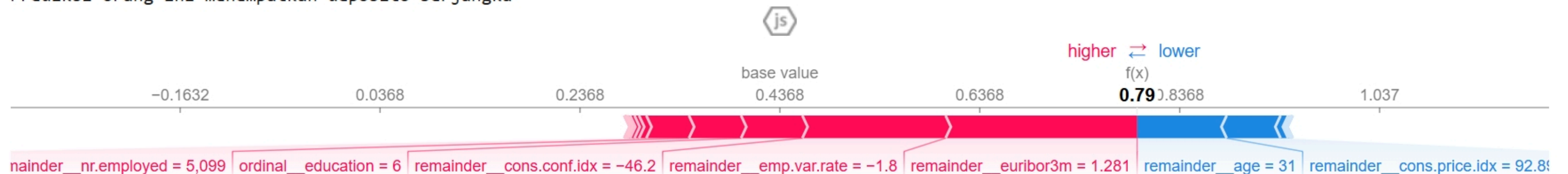
Before Using Machine Learning

Sales team will contact all prospects without filtering, which means profits are calculated by subtracting the total telemarketing costs for all contacted individuals from the funds acquired.
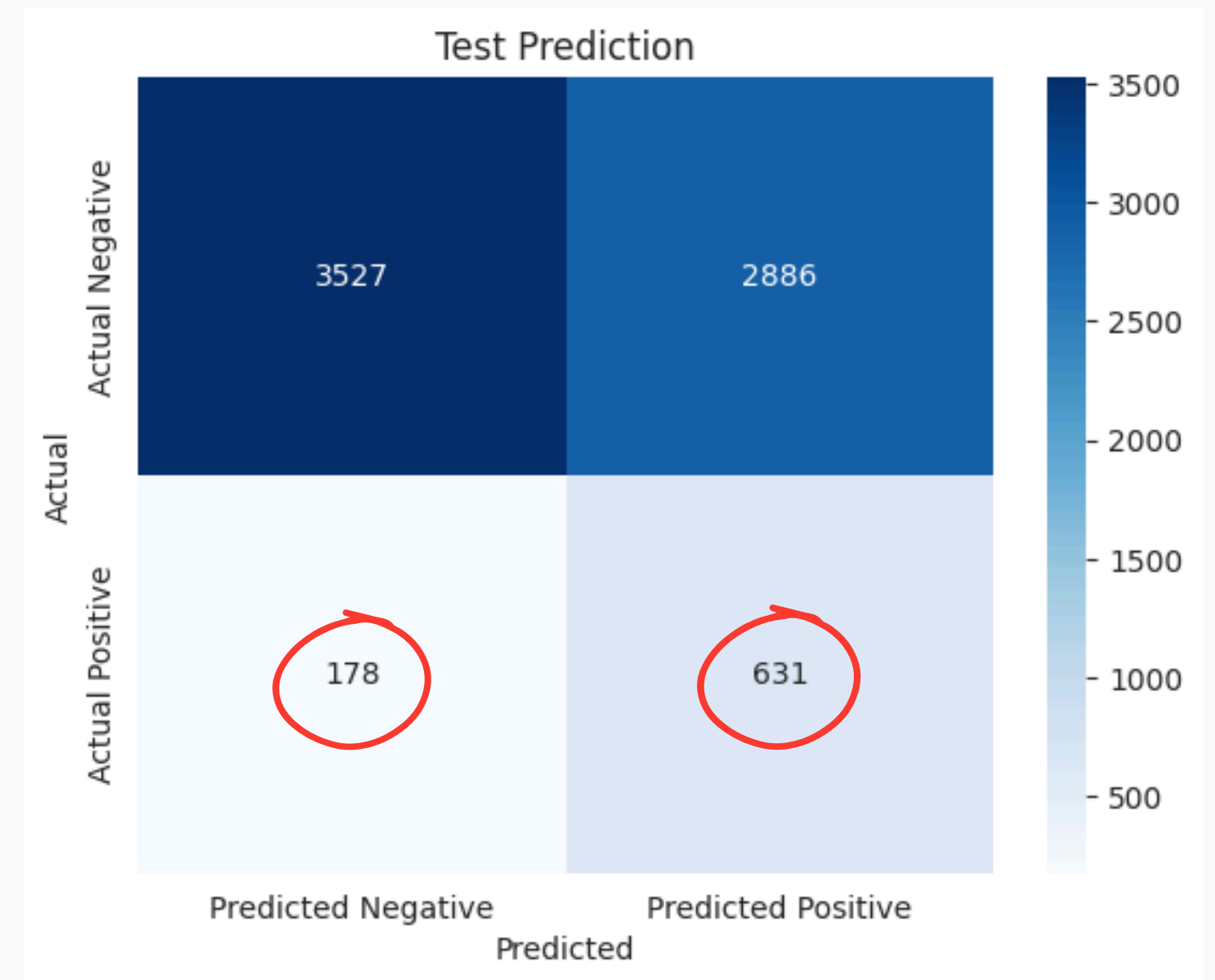
196 * (631+178) − 9 * (631+178+3527+2886) = 93,566 EUR

After Using Machine Learning

If the company uses machine learning, the sales team will only contact individuals predicted to be interested in term deposits, maximizing profits by subtracting telemarketing costs from the total funds acquired.

**196 * 631 – 9 * (631+178) = 116,395 EUR**

Test Prediction

Results

Provided specific recommendations for improved marketing strategies:

- Targeted customer to offer term deposit.
- Optimal campaign approaches.
- Timing based on macroeconomic conditions to conduct more campaigns.

The model predicted the probability of customer interest before outreach, enabling more effective targeting.

Business Impact
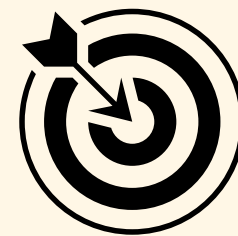
Enabled more targeted and personalized campaigns

10X
Financial losses decreased

24%
Profits Increased

reduced time spent filtering potential customers

KEY INSIGHTS

01 Customers aged 17-27 and 60+, who are single, have higher education, and are either students or retired, are more likely to accept term deposit offers.

02 Avoid targeting individuals with prior loan defaults to reduce risk.

03 Campaigns are more effective when conducted via mobile communication and scheduled midweek.

04 High consumer confidence correlates with increased customer interest in term deposits.