

MARKETING CAMPAIGNS DATA OF A PORTUGUESE BANKING INSTITUTION

April 2022

Project Team:
Syrvachev Sergey
DataScientist



1

INTRODUCTION

2

RESEARCH

3

FINALIZE

-
- Business Problem
 - Data Sources & Methods
- Current situation
 - Data Exploration
 - Modeling
- Business recommendations
 - Next Steps



INTRODUCTION

- Business problem
- Key Idea
- Data sources & Methods

BUSINESS PROBLEM

- A Portuguese banking institution hired our team **to create a model which predicts if the client will subscribe to a term product.**
- We should consider the associated revenue from each contract and associated costs to make each contact.
- We should explore the data and find valuable insights about crash safety issues.



DATA UNDERSTANDING

Sources

- **Data collected from one of the marketing campaigns**

Provided by the customer

Data covering period interval from March to December

Methods

- Exploratory data analysis (EDA)
- Feature engineering
- Major prediction models for classification problems

METRICS

Performance:

To compare models, we will focus on 2 major metrics:

- Accuracy - How well we can predict TP and TN. General metrics that will show model performance.
- Recall - We work under the assumption that the cost of contact is lower compared to the possible profit of the customer signed. So we are interested in catching as many potential signed-up customers and minimizing FN. So I will use recall as the second metric.

Economy

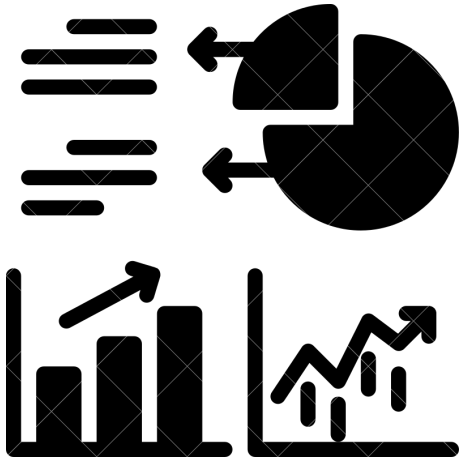
We will use the campaign balance metric:

Balance = Revenue from subscription - contact cost * number of contacts*

- Contact cost - 5 USD: The cost associated with each contact was made with the customer
- Revenue from sign-up client - 200 USD: Revenue that is associated if customer subscribed to the product

*Some customers were contacted several times.

RESEARCH



- Data exploration
- Modeling
- Data Assessment





CURRENT SITUATION

09



LAST MARKETING CAMPAIGN:

> 41 000 PEOPLE CONTACTED:



11% OF THEM RESULTED
IN SUBSCRIPTIONS



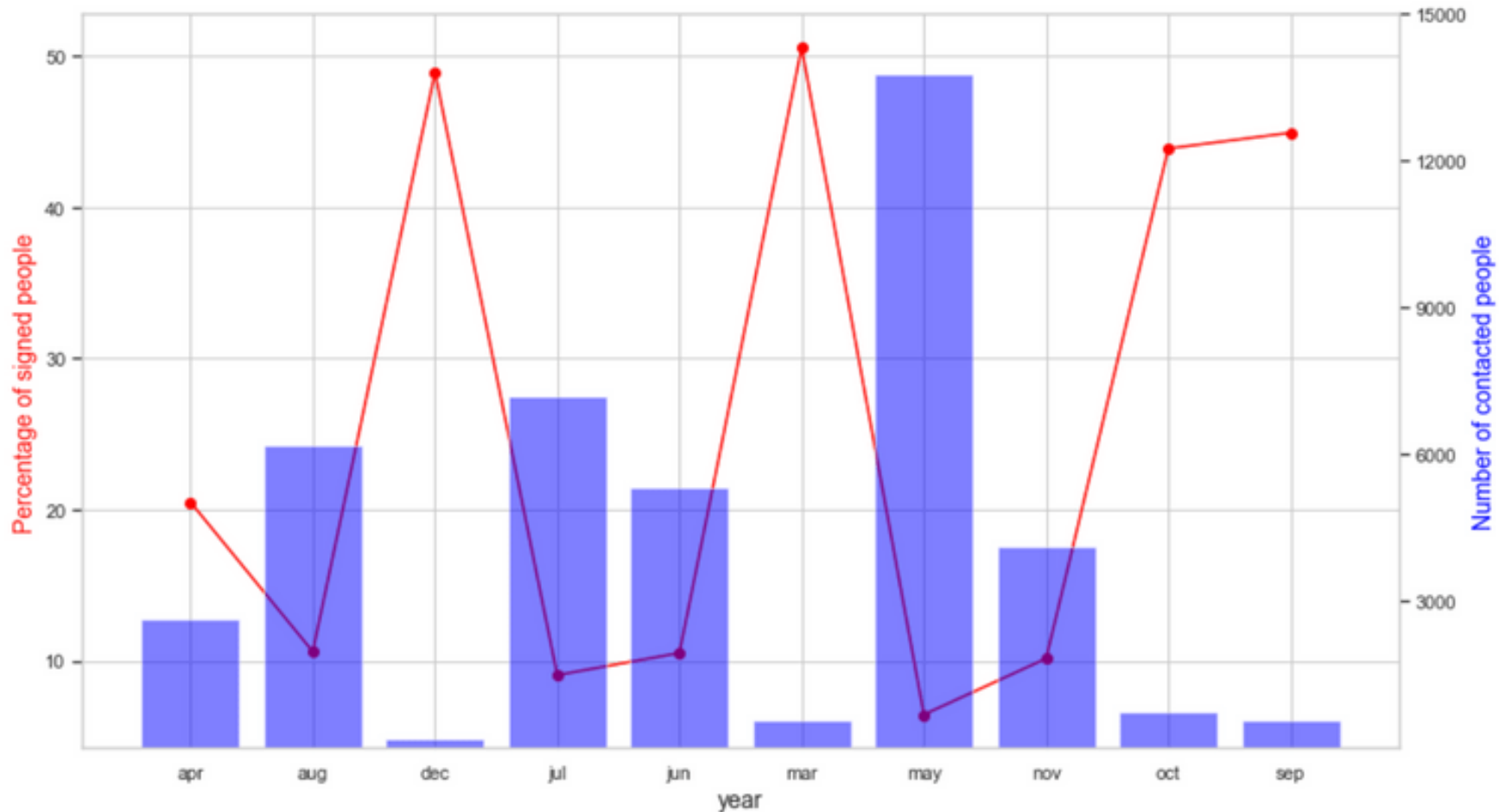
AGE BETWEEN 17 AND 98



DIFFERENT SOCIAL
GROUPS

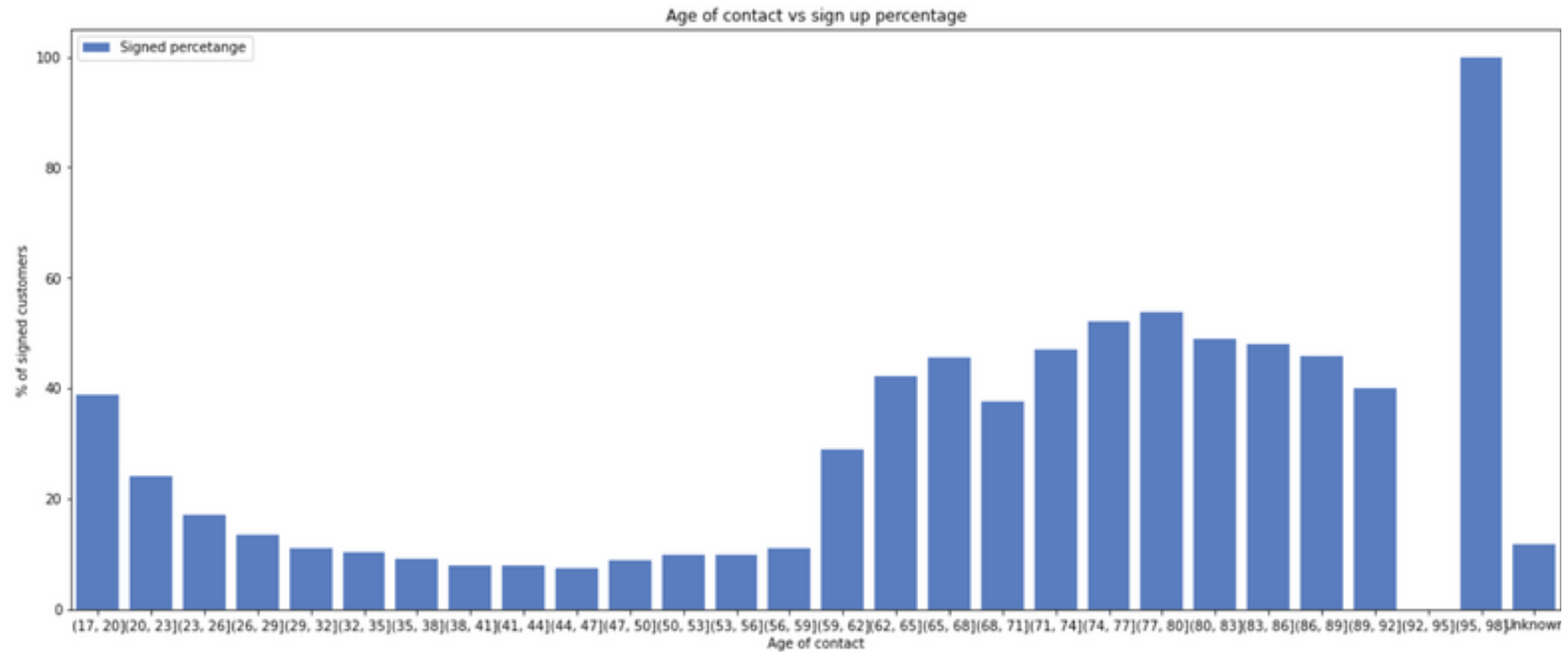
PATTERNS IN SUBSCRIPTION

10



DURING THE MONTH WITH THE LOWER NUMBER OF CONTACTS, THE PROPORTION OF PEOPLE WHO SUBSCRIBED WAS HIGHER

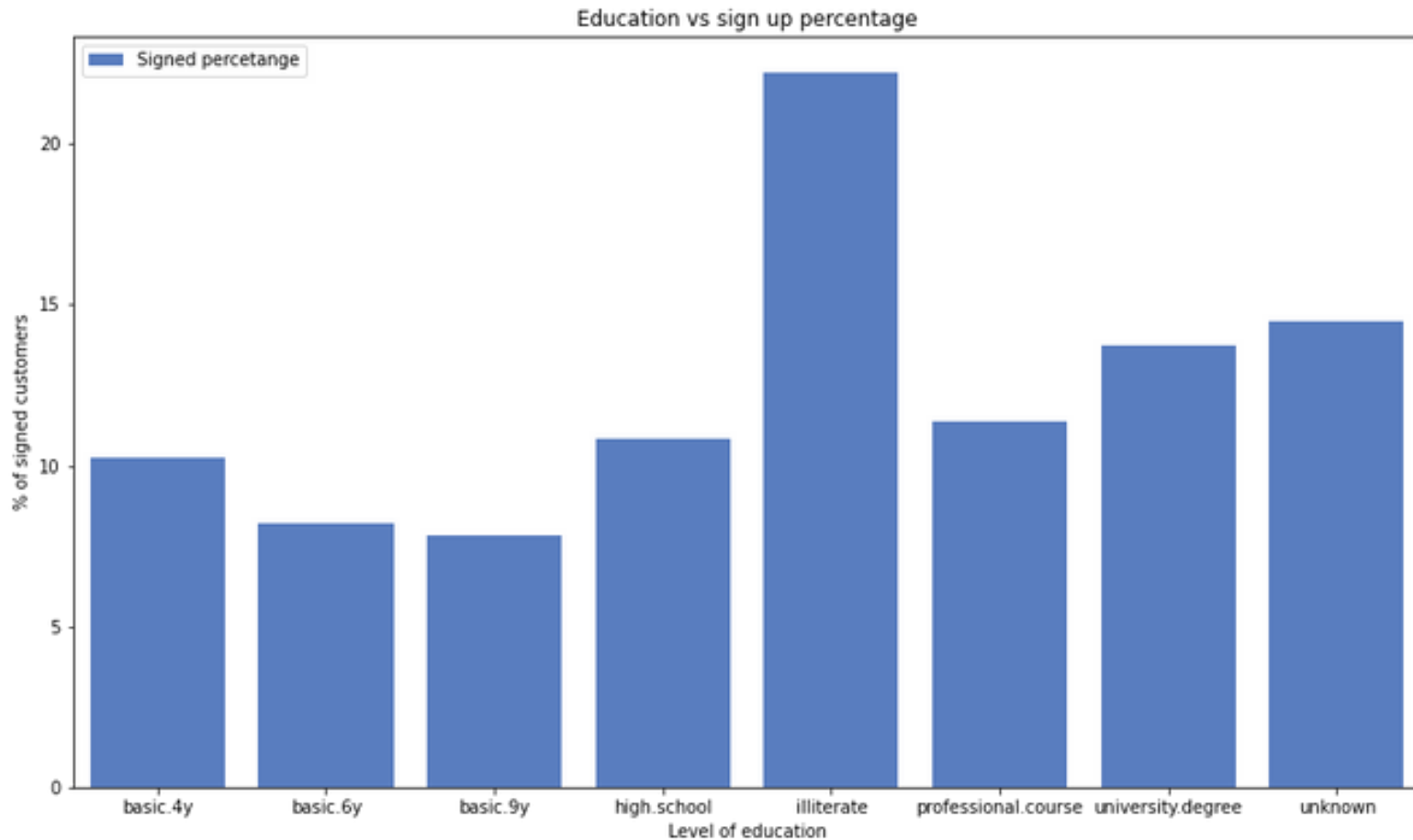
PATTERNS IN SUBSCRIPTION BY AGE



AGE GROUPS MORE PRONE TO SUBSCRIBE:

- 17-23
- 59+

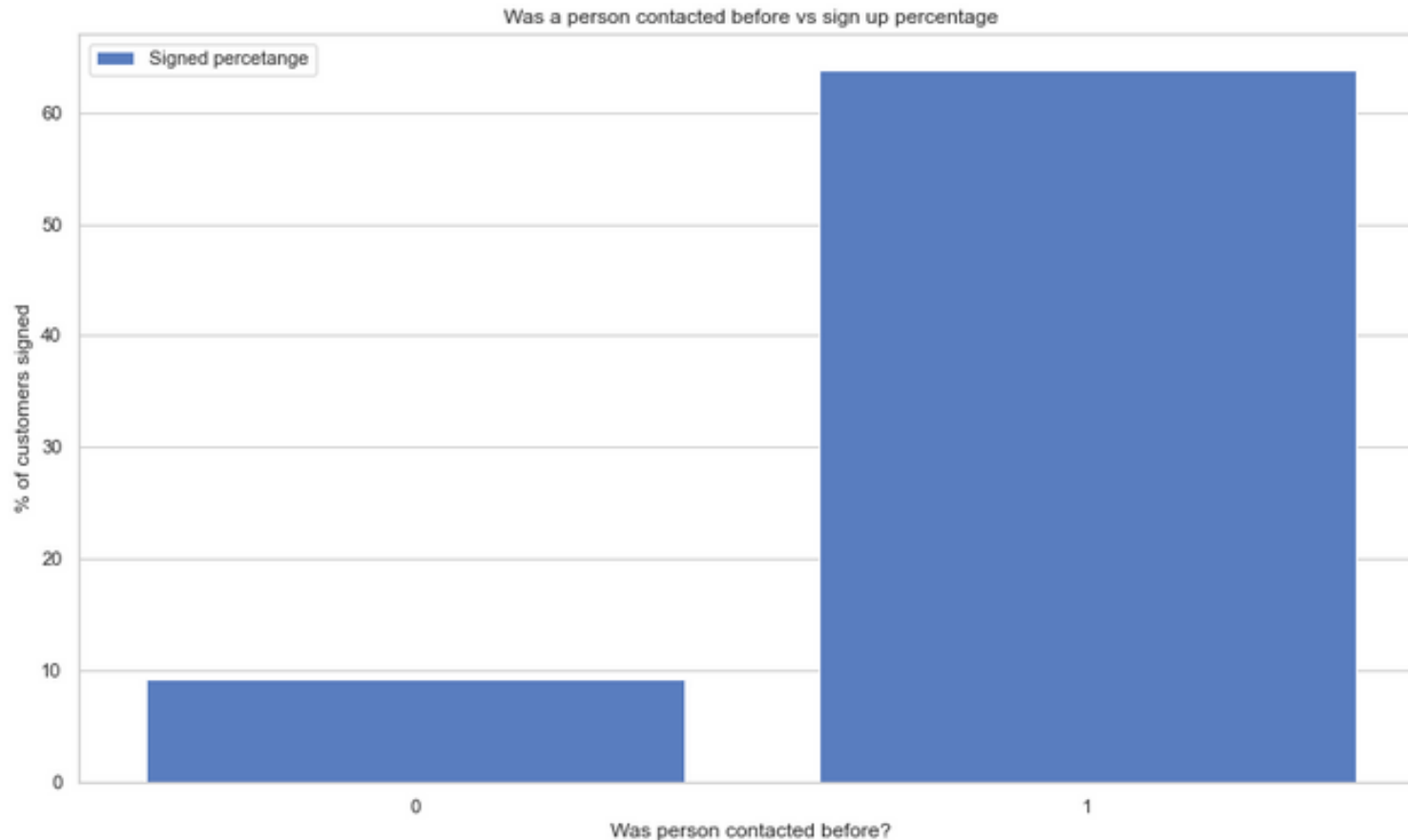
PATTERNS IN SUBSCRIPTION BY LEVEL OF EDUCATION



PREFERABLE LEVEL OF EDUCATION :

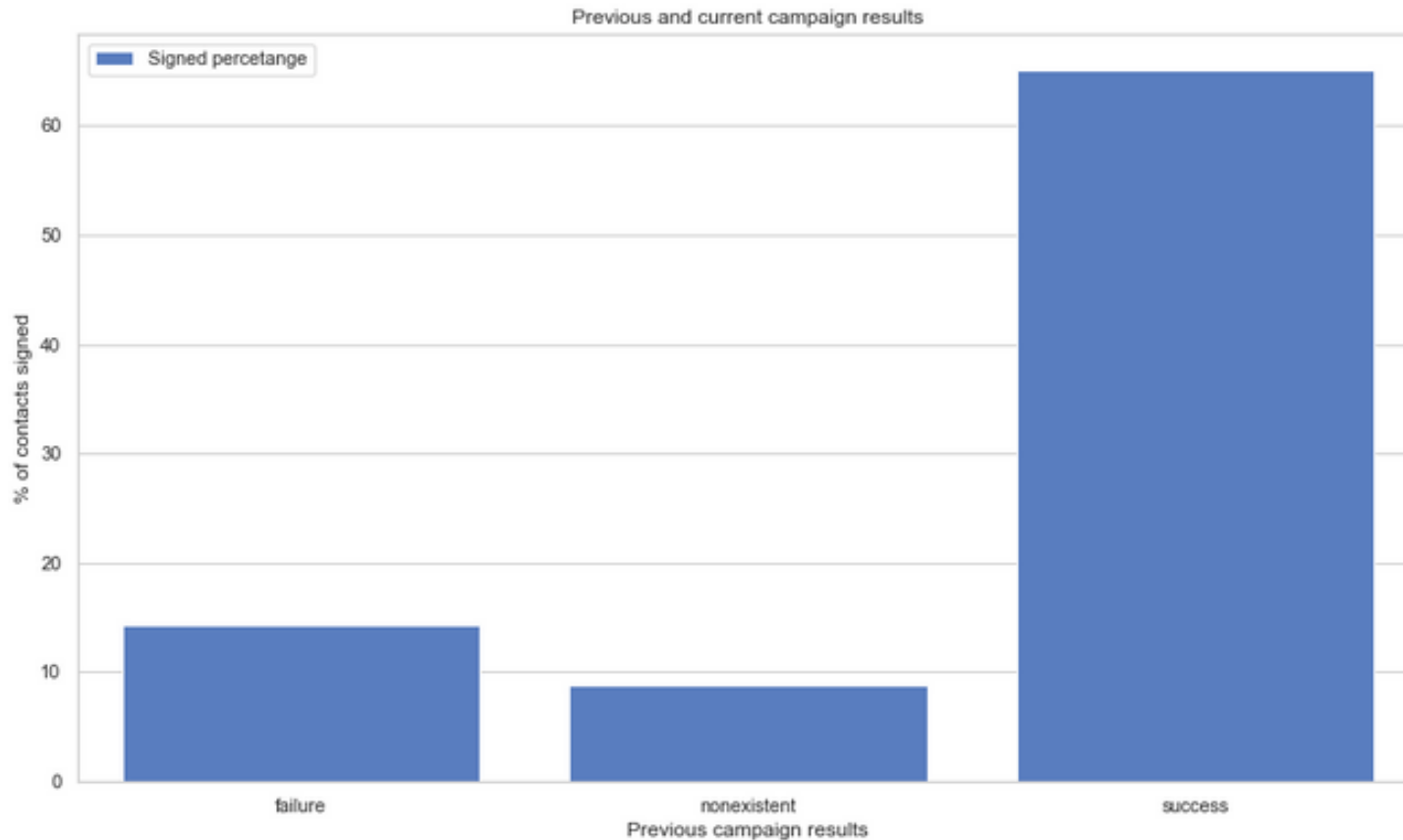
- ILLITIRATE
- UNIVERSITY DEGREE

EFFECT OF PREVIOUS CAMPAIGNS

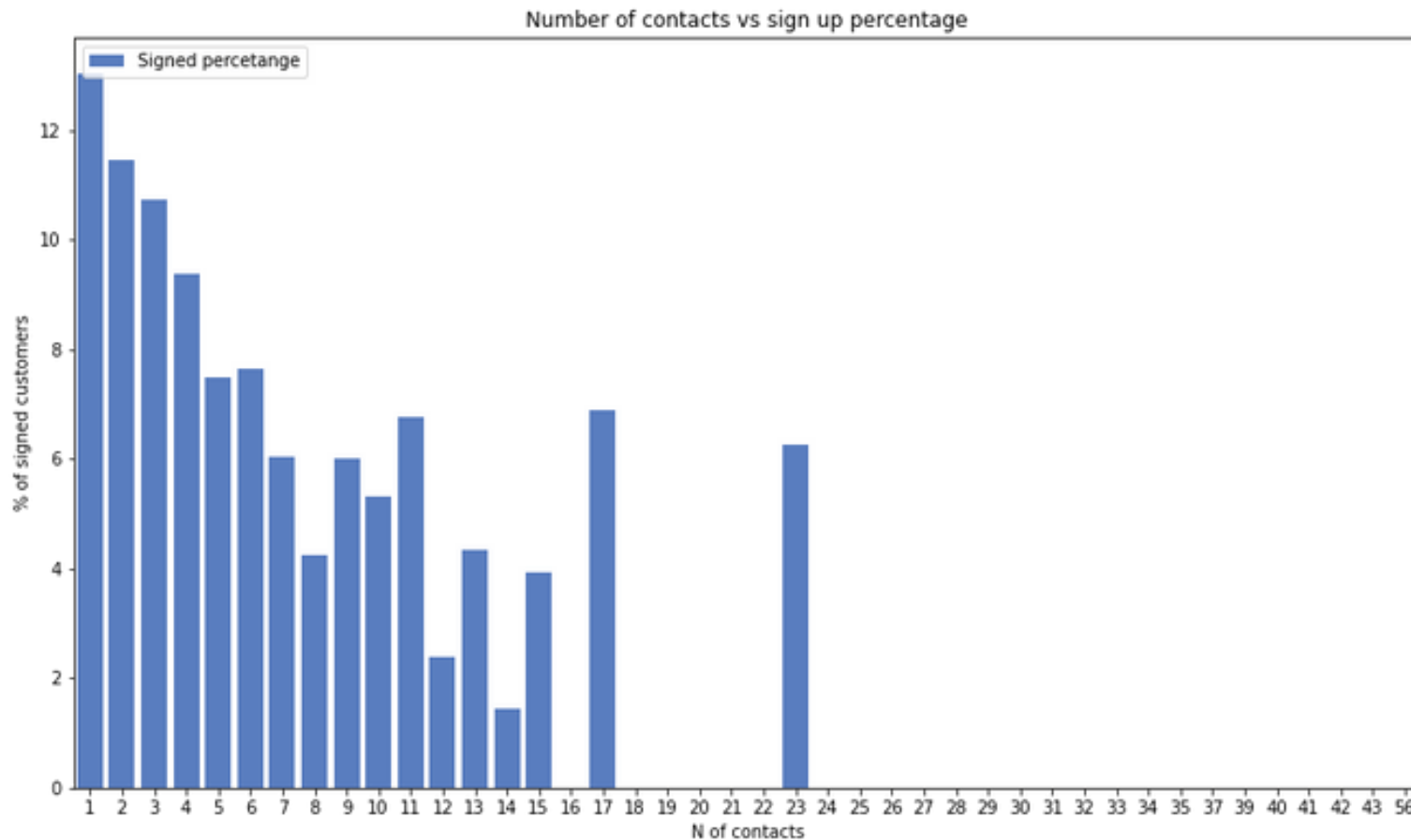


PEOPLE WHO WERE CONTACTED IN THE PREVIOUS CAMPAIGN ARE MORE LIKELY TO SUBSCRIBE TO THIS OFFER.

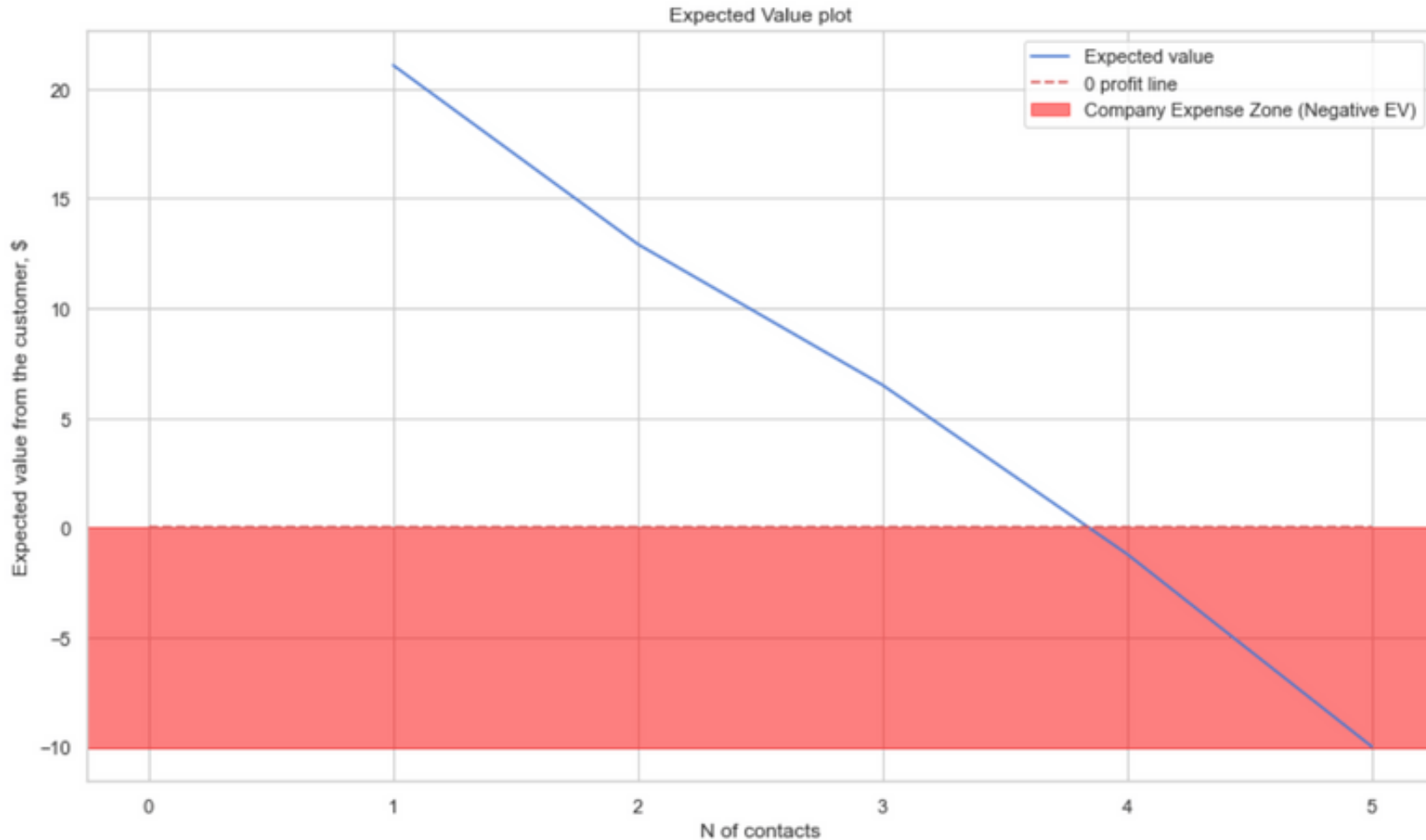
EFFECT OF PREVIOUS CAMPAIGNS 2



SUCCESS IN PREVIOUS CAMPAIGNS CONTRIBUTE TO NEW OFFER



- CUSTOMERS WERE CONTACTED UP TO 56 TIMES.
- THE MOST SIGNIFICANT PERCENTAGE OF PEOPLE SUBSCRIBED AFTER 1ST CONTACT



- THE EXPECTED VALUE FROM THE CUSTOMER WHO WAS CONTACTED MORE THAN 3 TIMES IS NEGATIVE.
- THE OPTIMAL NUMBER OF CONTACTS 1-3



MODEL HIGHLIGHTS:

- TRAINED ON **33 000 RECORDS**

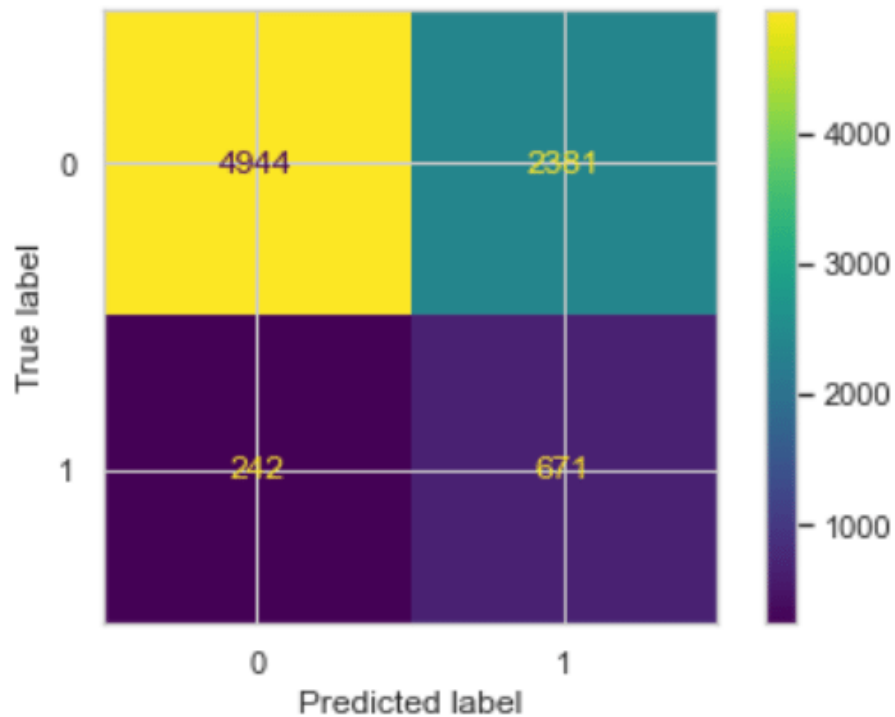
- FOCUSED ON PREDICTING CUSTOMERS WHO ARE **MORE LIKELY TO SUBSCRIBE**

- WAS CHOSEN FROM **MORE THAN 50 CANDIDATES**

- CORRECTLY PREDICT **> 70% CASES**

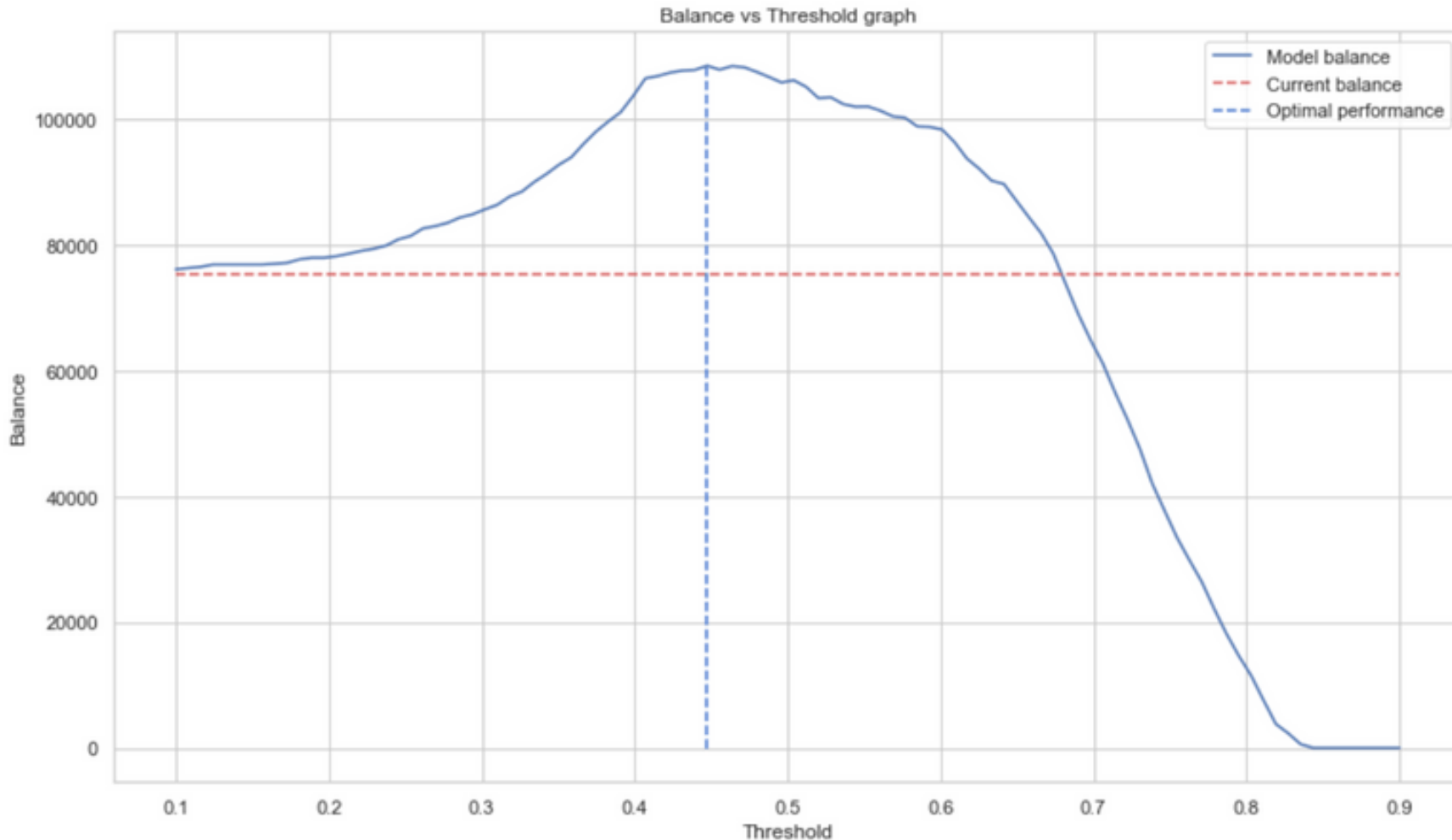
ARTIFICIAL INTELLIGENCE THAT SERVES FOR HUMANITY

Final model recall 0.7349397590361446
Final model accuracy 0.6815974751153192
Final model balance 105990
Confusion matrix:



INCOME BALANCE INCREASE FOR 40%

105 990\$ VS 75 415\$

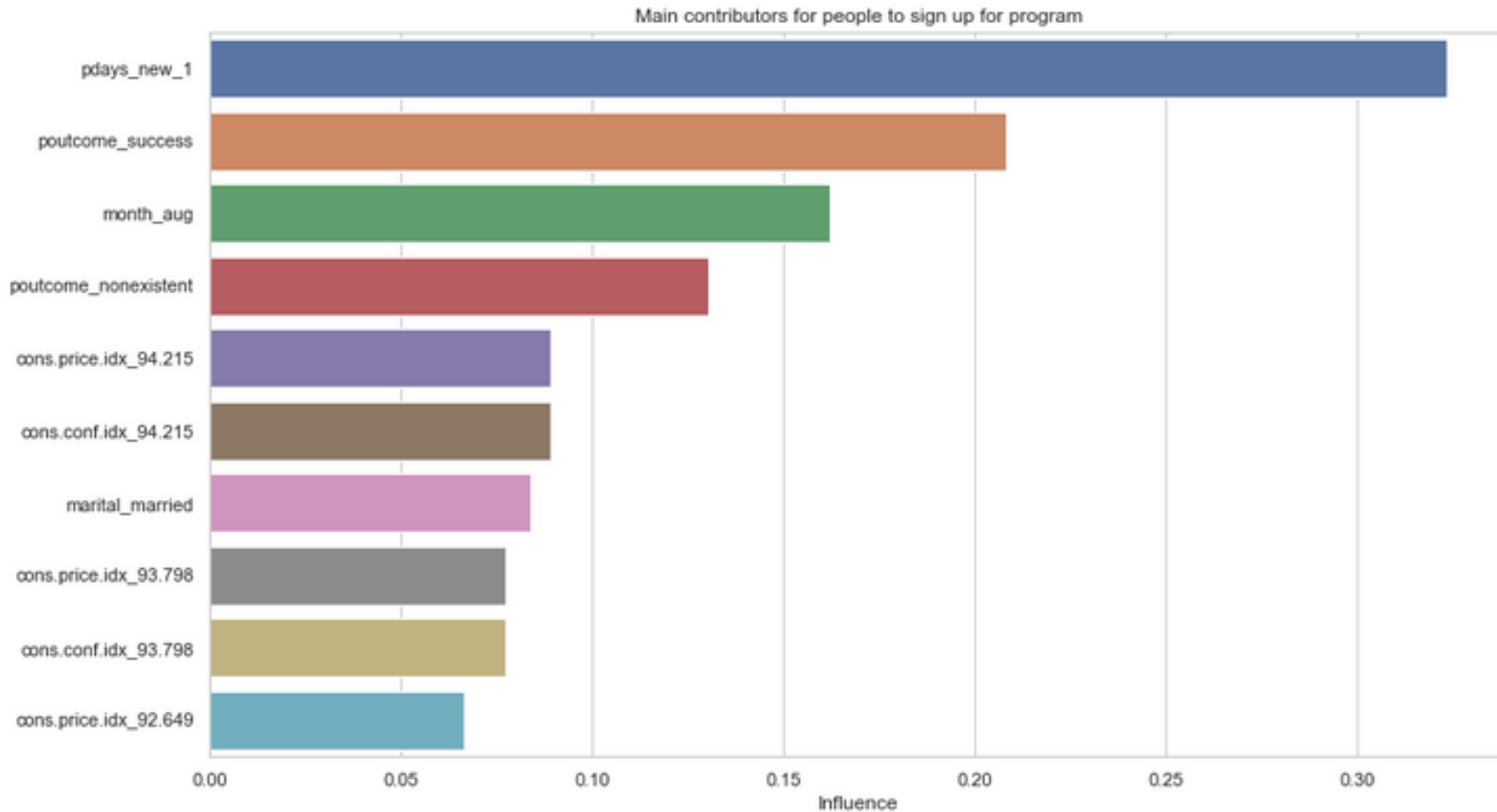


OPTIMIZED PREDICTION BOUNDARY - 0.447

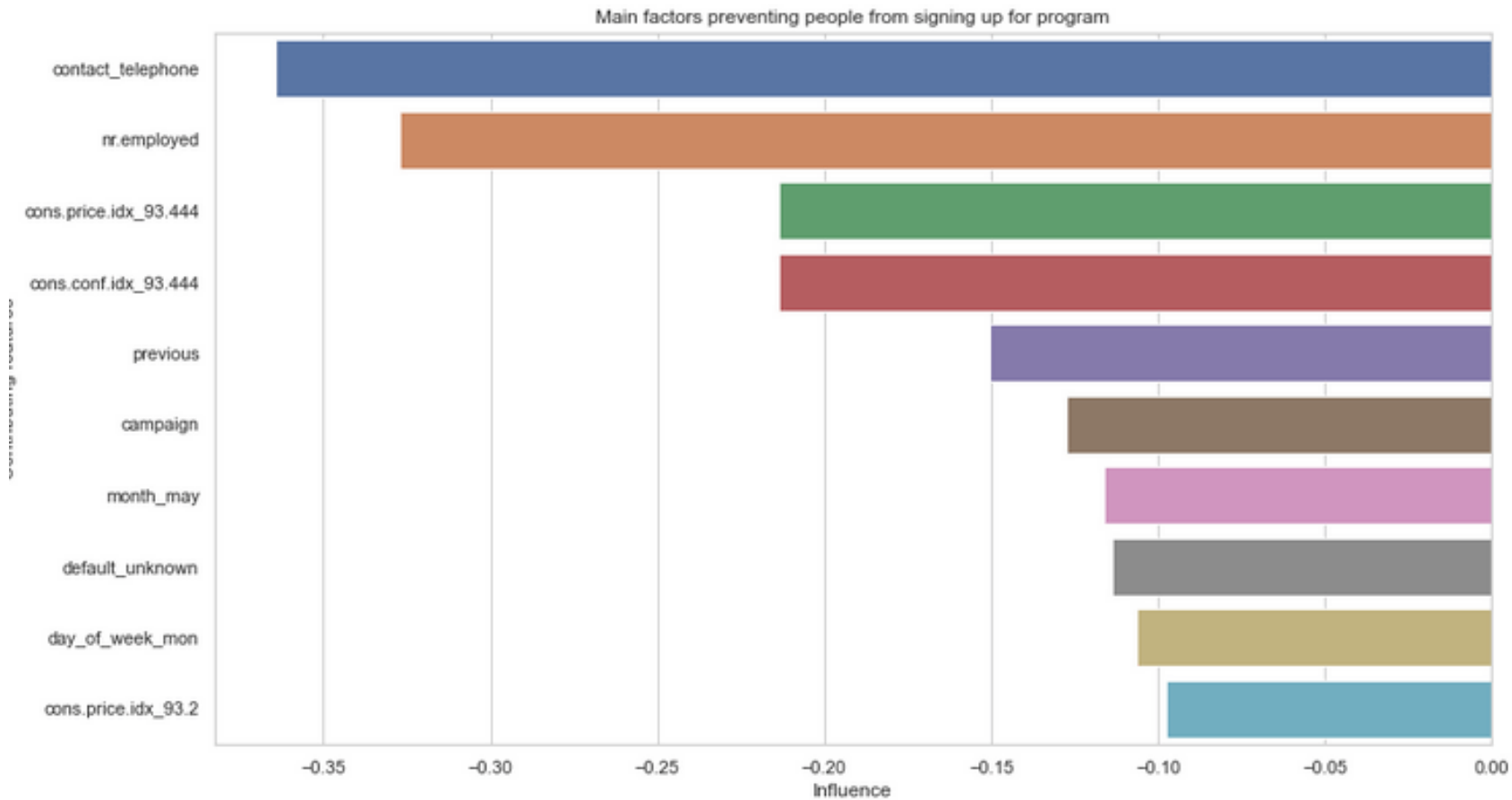
MODEL ACCURACY 0.4954

MODEL RECALL 0.8467

MODEL BALANCE 108505

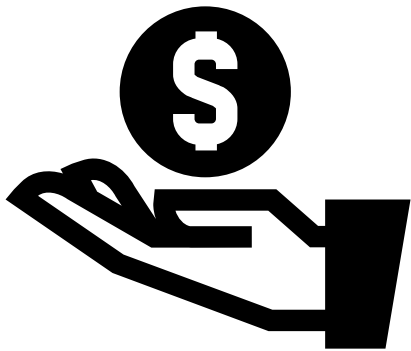


MAIN CONTRIBUTORS CORRESPOND TO EDA



SHOULD BE DOUBLE CHECKED

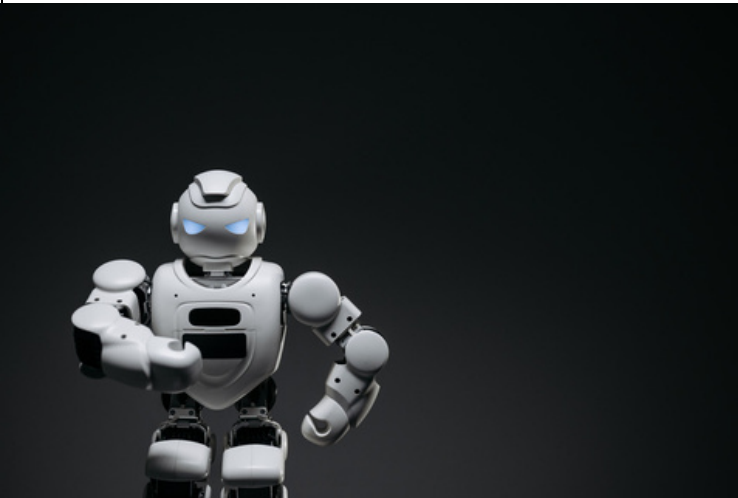
FINALIZE



- Business recommendations
- Next Steps



Recommendations



01

**REDUCE NUMBER OF
CONTACTS**

As we can see from EDA, EV of clients with more than 3 contacts is below zero

02

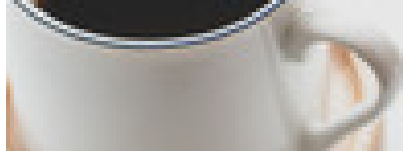
MONTHLY ANOMALIES

Some months have spikes in subscriptions, needs to be investigated

03

ML IMPLEMENTATION

Implementation of ML algorithms allows to increase in the overall revenue from marketing campaign



Next steps

- Repeat baseline calculations and modeling considering the reduced number of calls, "monthly subscriptions" investigations
- Get information about planned marketing campaigns and consider the Future Value of customers.
- Use other ML algorithms



A photograph of a two-lane asphalt road stretching into the distance, shrouded in a thick fog. The road is flanked by dense green trees and foliage. On the right side, several utility poles with wires are visible. The overall atmosphere is misty and serene.

"NEVER NEGLECT AN
OPPORTUNITY FOR
IMPROVEMENT"

SIR WILLIAM
JONES

Q & A:

**Thank you for joining
today's presentation.**



SYRVACHEV SERGEY

- DATA SCIENTIST
- DEST.STUDIO@GMAIL.COM
- LINKEDIN: /SSYRVACHEV
- GITHUB: 314KA4Y
- MEDIUM: @SERGEYSYRVACHEV