PREDICTING TERM DEPOSIT SUBSCRIPTION

CENTENNIAL COLLEGE

GLEN GEORGE BAUTISTA
MARIA PAULA SANCHEZ GALINDO

AGENDA

Executive Summary

Introduction, Objective, Approach

Business Understanding

Definitions, Scenario

Data Understanding

Initial Observations

Data Preparation and Modeling

Data Preparation, Model Exploration, Model Selection

Actions

Conclusions and Recommendations

EXECUTIVE SUMMARY

Introduction

- Direct product introduction is prevalent in numerous industries, including the banking industry.
- By evaluating bank marketing data, it is possible to select the style of marketing to implement.
- The effectiveness of marketing campaigns should be addressed in order to reduce labor costs and enhance revenues.

Objective

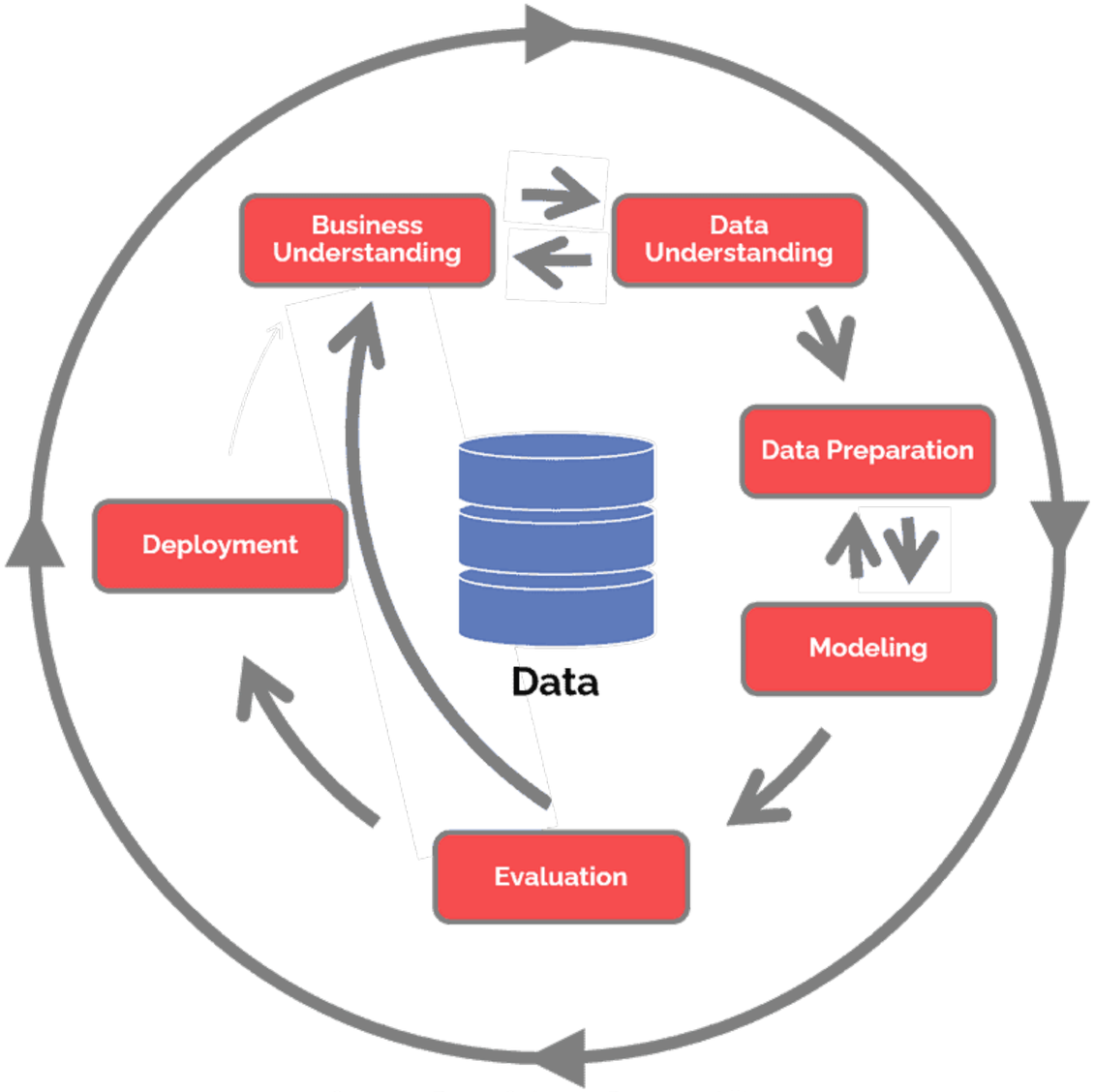
- Examine clients' information and behavior to identify which factors influence their decision to open a term deposit.
- Conduct statistical analysis and evaluate various models.
- Increase campaign effectiveness.

Approach

- CRISP-DM
- Classification Models

APPROACH

Cross Industry Standard Process for Data Mining



Source: Data Science Process Alliance

BUSINESS UNDERSTANDING

THE INDUSTRY

TERM DEPOSIT

A deposit held at a financial institution that has a fixed term.

STRATEGY

Communicating straight to the customer via telephone calls.

BUSINESS UNDERSTANDING

CLASSIFICATION MODEL

Client Data

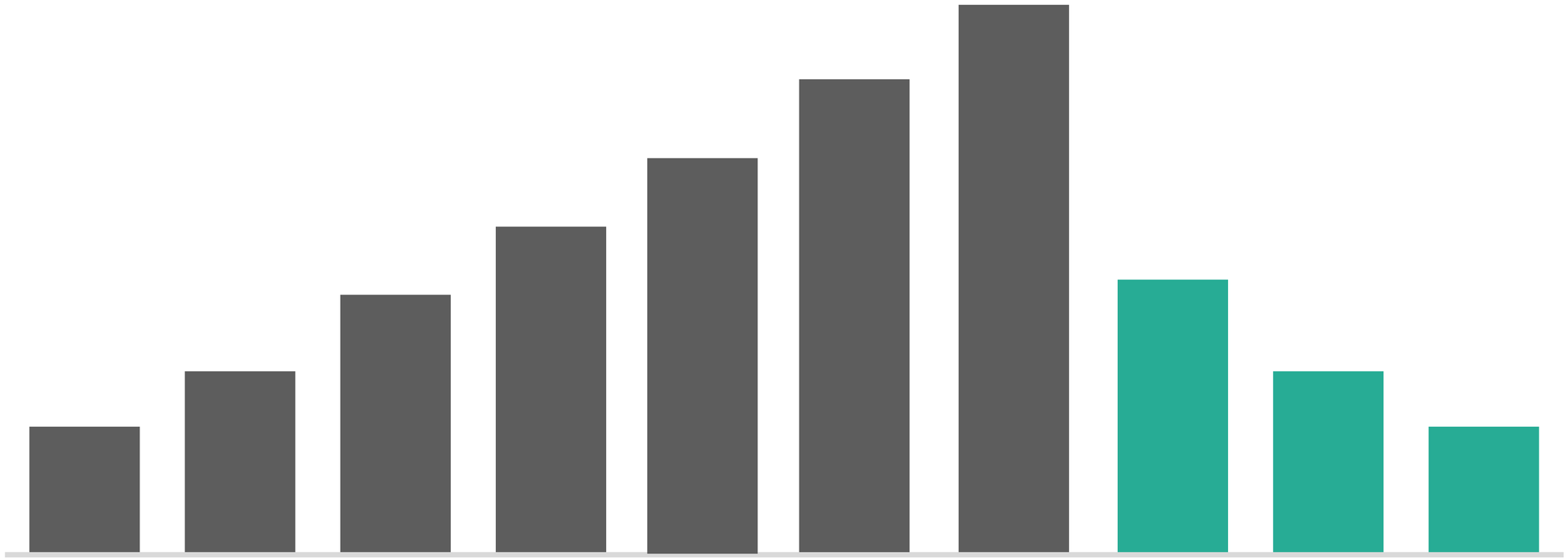
age, marital status, loan status, etc.

Campaign Contact Information

last contact, preferred contact type, etc.

INITIAL OBSERVATION

OVERALL RESPONSE RATE



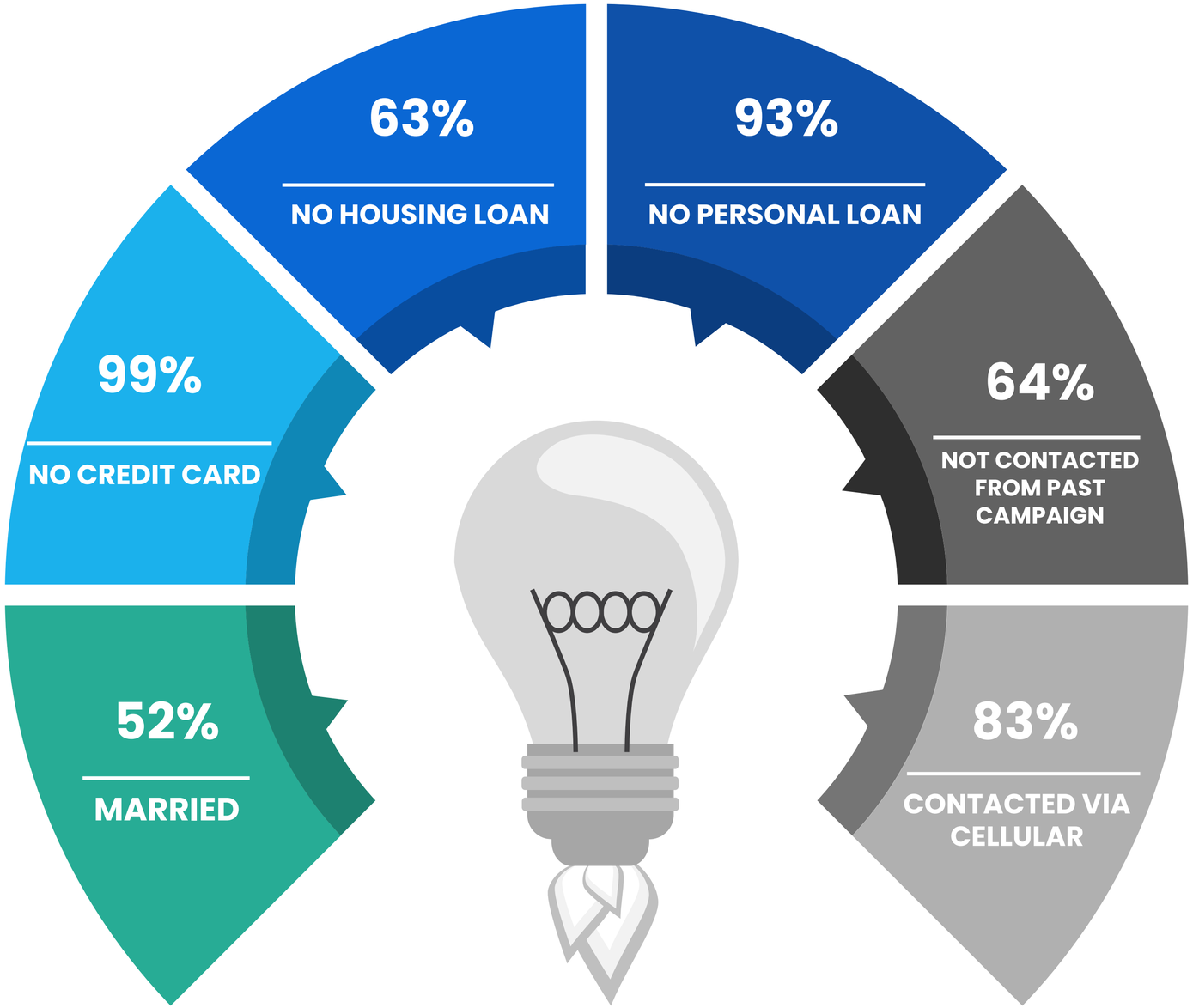
89%

Client non-subscription of term deposit

11%

Client subscription of term deposit

DATA UNDERSTANDING



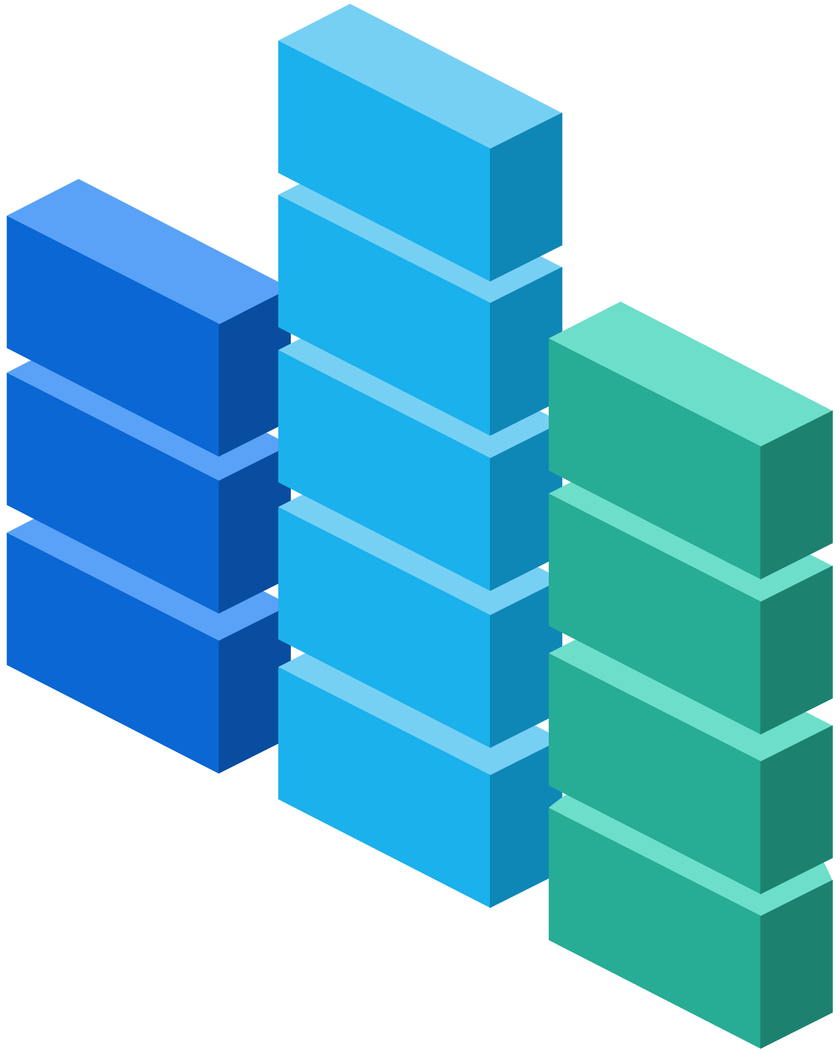
DATA PREPARATION

Variables Investigation

- Minimum values
- Maximum values
- ‘Unknown’ data

Data Cleansing

- Outliers
- Missing values
- Unnecessary columns



MODELLING

Model Exploration

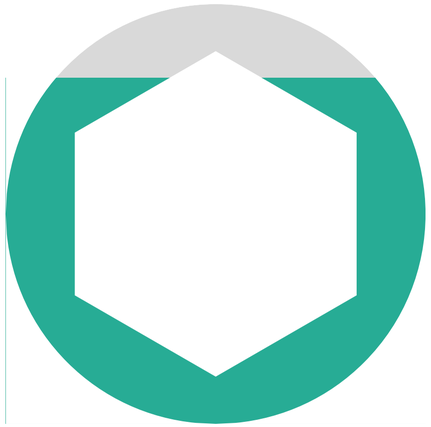
- Modeling Approach
- Metrics
- Default Classifiers
- Hyperparameter Tuning

Model Recommendation

- Comparison Table
- Model Selection
- Model Recommendation

The approach to modeling in this investigation was divided in two groups, for both of them the set of techniques utilized were the same (Random Forest Classifier, Decision Tree Classifier, Logistic Regression Classifier, and Gradient Boost Classifier), with the difference that in the second group, we did a hyperparameter tuning.

MODEL EXPLORATION



1

Modeling Approach

Divided into two groups (default classifiers, classification with hyperparameter tuning)

2

Metrics

Accuracy, Recall, Specificity, Precision, Balanced Accuracy, F1 score, AUC-ROC Curve

3

Default Classifiers

Random Forest Classifier
Decision Tree Classifier
Logistic Regression Classifier
Gradient Boost Classifier

4

Hyperparameter Tuning

Hyperparameter Tuning using GridSearch

MODEL RESULTS

1 Comparison for Default Classifiers

	Model	Accuracy	Recall	Specificity	Precision	Balanced Accuracy	F1 score
0	RandomForestClassifier()	0.665886	0.567807	0.678445	0.184412	0.623126	0.278404
1	DecisionTreeClassifier()	0.568503	0.426297	0.586712	0.116669	0.506505	0.183200
2	LogisticRegression(max_iter=5000)	0.127836	0.926887	0.025519	0.108571	0.476203	0.194374
3	GradientBoostingClassifier()	0.772840	0.504717	0.807173	0.251026	0.655945	0.335292

Table 2: Comparison table results of classifiers models with default parameters

2 Comparison after Hyperparameter Tuning

	classifier	accuracy	recall	specificity	precision	f1-score	balanced
0	RandomForest_h	0.583027	0.688679	0.569498	0.170015	0.272706	0.629089
0	DecisionTree_h	0.886085	0.063090	0.991468	0.486364	0.111691	0.527279
0	LogisticRegression_h	0.886554	0.075472	0.990411	0.501961	0.131215	0.532942
0	GradientBoosting_h	0.891440	0.175708	0.983088	0.570881	0.268711	0.579398

Table 3: Comparison table for the result of the models – Hyperparameter Tuning

MODELING SELECTION

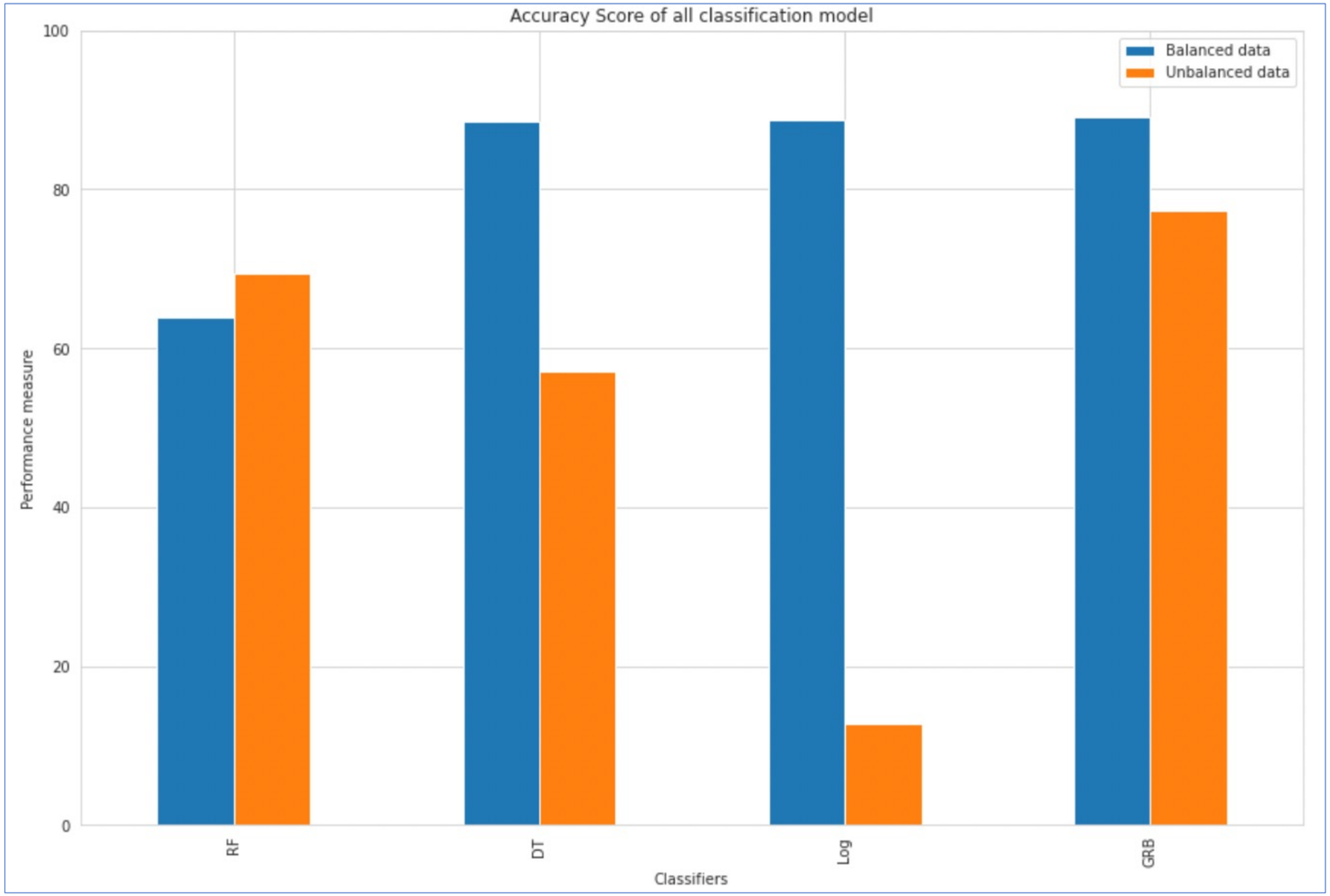
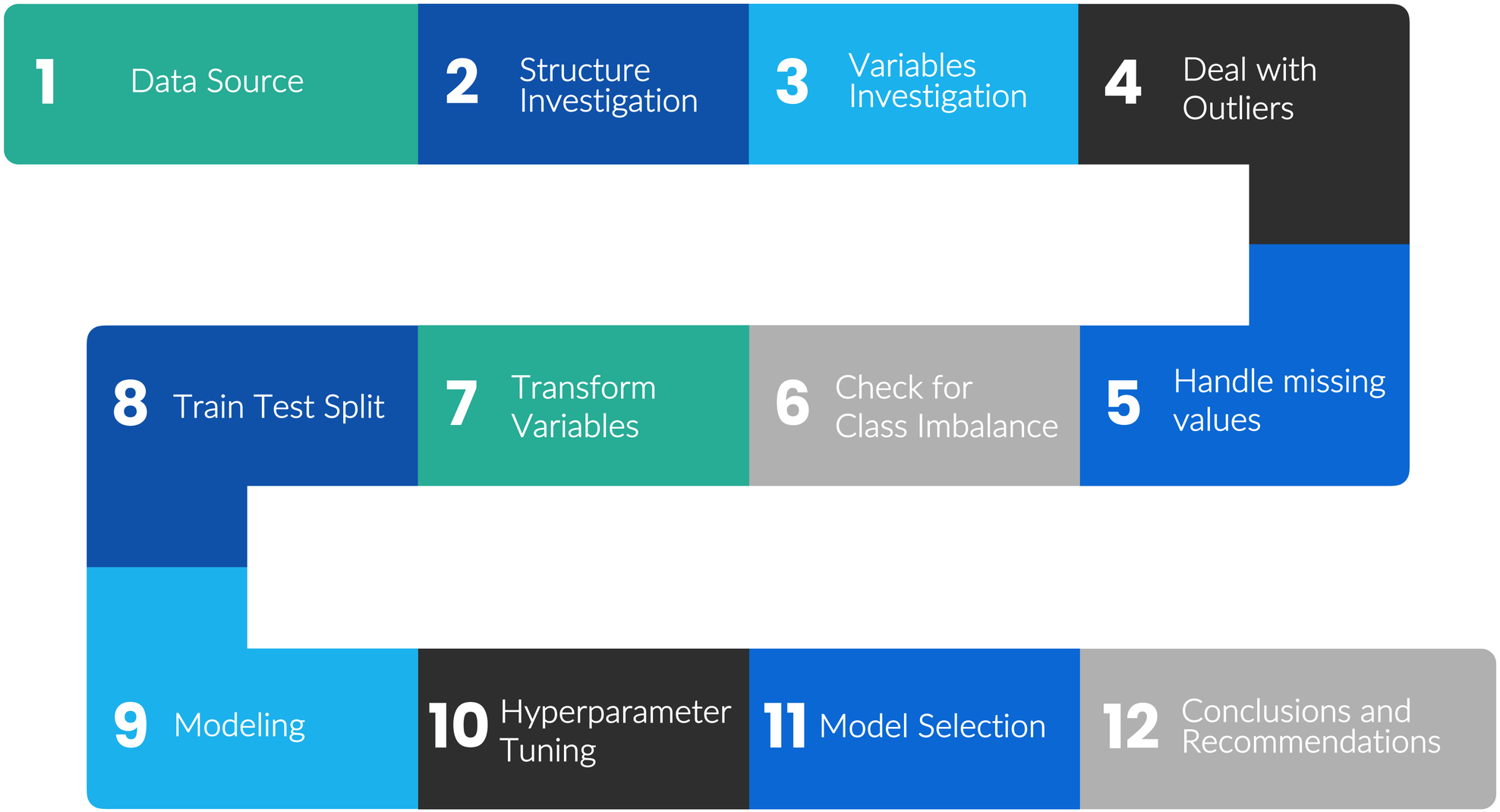


Figure 1: Plot - Accuracy score of all classification models.

PROJECT FLOW



CONCLUSIONS

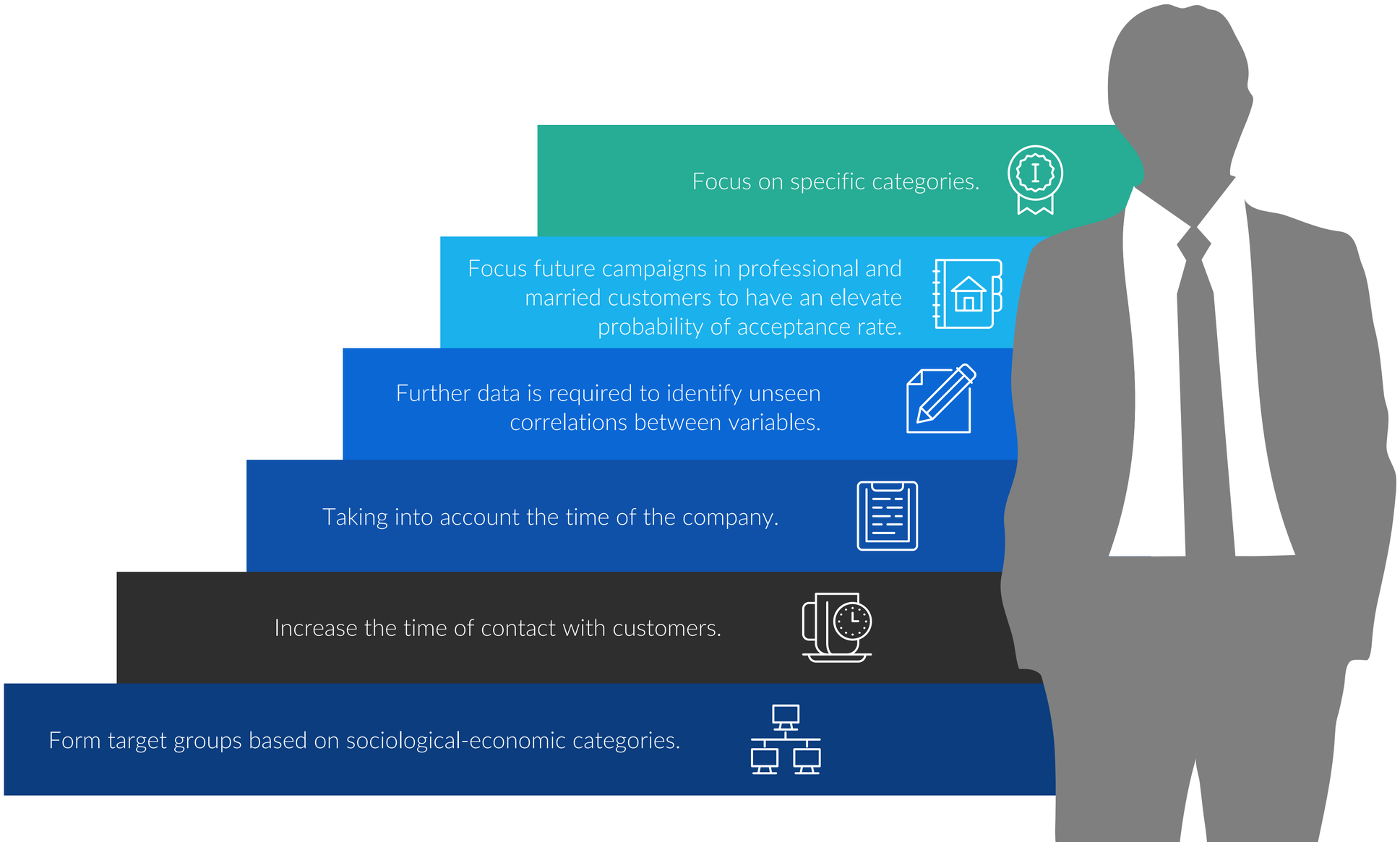
Administrative and blue-collar professionals are more receptive to accept the term deposit.

Higher educated customers are more knowledgeable of considering the term deposit offer.

Married customers are likewise more inclined to consent to the term deposit.

May is the most effective month to conduct a bank marketing campaign.

RECOMMENDATIONS



THANK YOU!

MARIA PAULA SANCHEZ GALINDO

Business Analytics and Insights
Centennial College

GLEN GEORGE BAUTISTA

Business Analytics and Insights
Centennial College