

The background of the slide is a collage of financial and business-related items. In the top left, a portion of a black calculator is visible, showing keys for '3', '6', '+', and '='. Below the calculator, there are several charts: a bar chart at the top with months from May to December on the x-axis, a pie chart in the center, and a line graph at the bottom left with data points connected by lines. To the right of the pie chart, there is a legend with colored squares corresponding to months from January to July. In the top right corner, a stack of Euro banknotes is fanned out. Below the banknotes, a silver compass is shown, pointing towards the top right. At the bottom, a black pen lies diagonally across a table with numerical data. The text 'Bank Marketing Success Classification problem' and 'Rui Yuan' is centered in a white box with a thin black border.

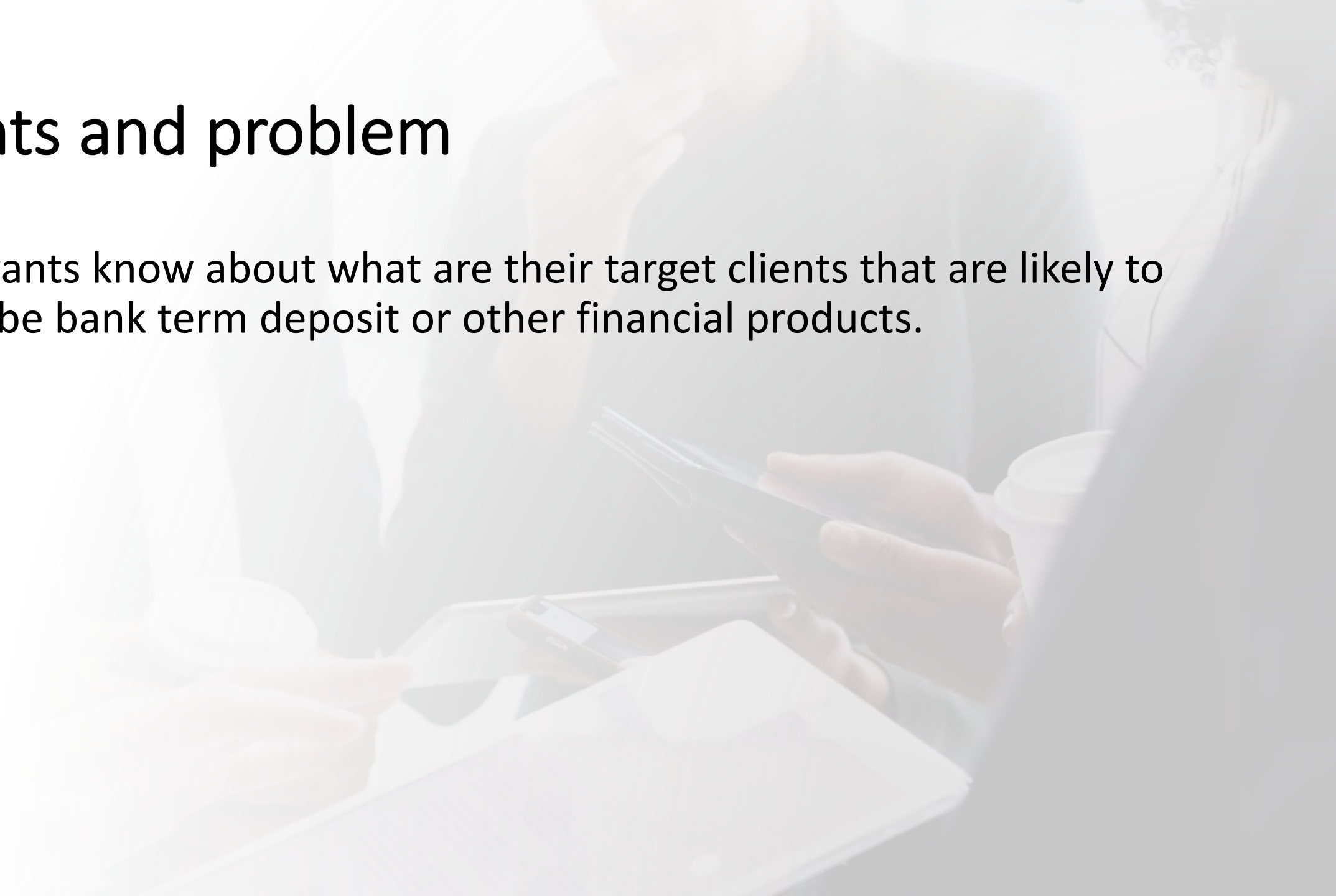
Bank Marketing Success Classification problem

Rui Yuan

125,058	154,568	95,054	124,500
125,487	56,845	97,511	125,000
124,000	110,000	99,011	154,000
1450	150,000	99,216	95,000
	35,000	101,090	154,200
		101,684	110,000
		101,962	89,000
			50,000
			10,700

Clients and problem


Bank wants know about what are their target clients that are likely to subscribe bank term deposit or other financial products.



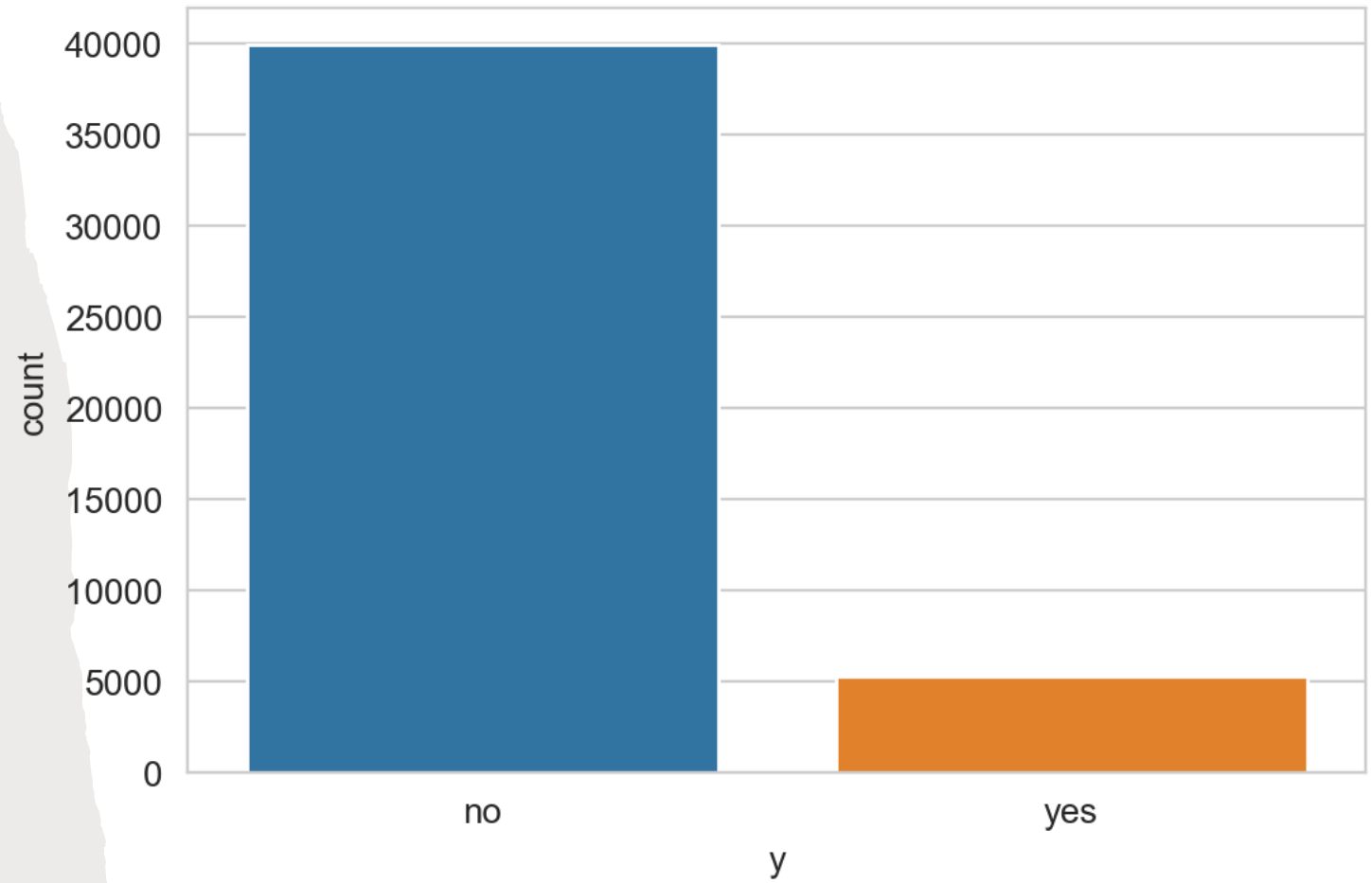
Data

- The data is about bank telemarketing which consists of clients' information, communication frequency and marketing outcome.
- 45211 rows
- Target: Whether client subscribe a term deposit (1 = yes, 0 = no)
- 15 Features: clients: age, bank balance, loan....

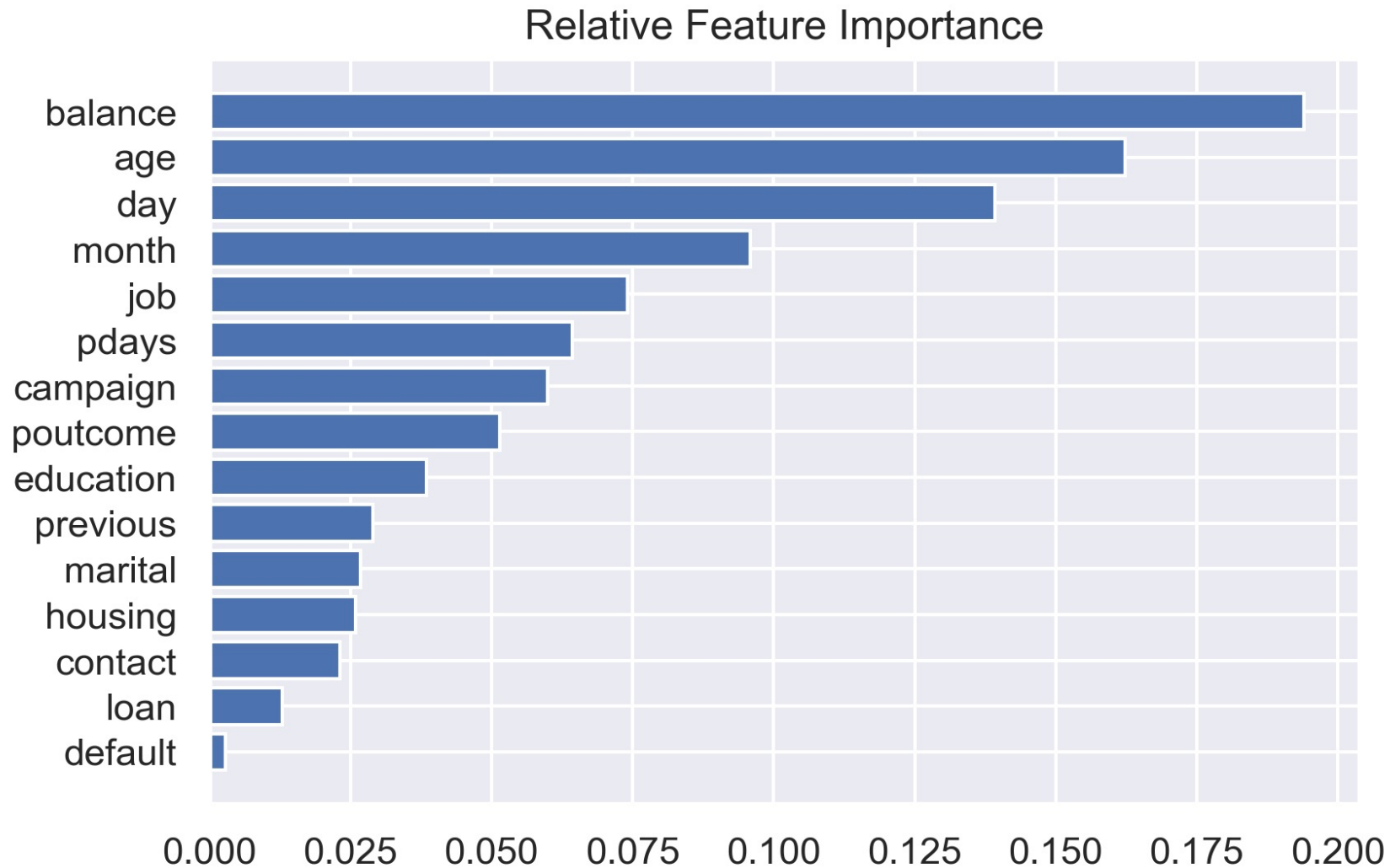
Classification Modeling Goal

- Optimization Goal:  precision score on positive class (subscribe) & number of successful deals
- In Business Sense: Adjustable model depending on the business capability. I.e., how many clients bank can reach to or how many phone calls bank can do in a given period (month).

Class Distribution



Feature Importance from RF model

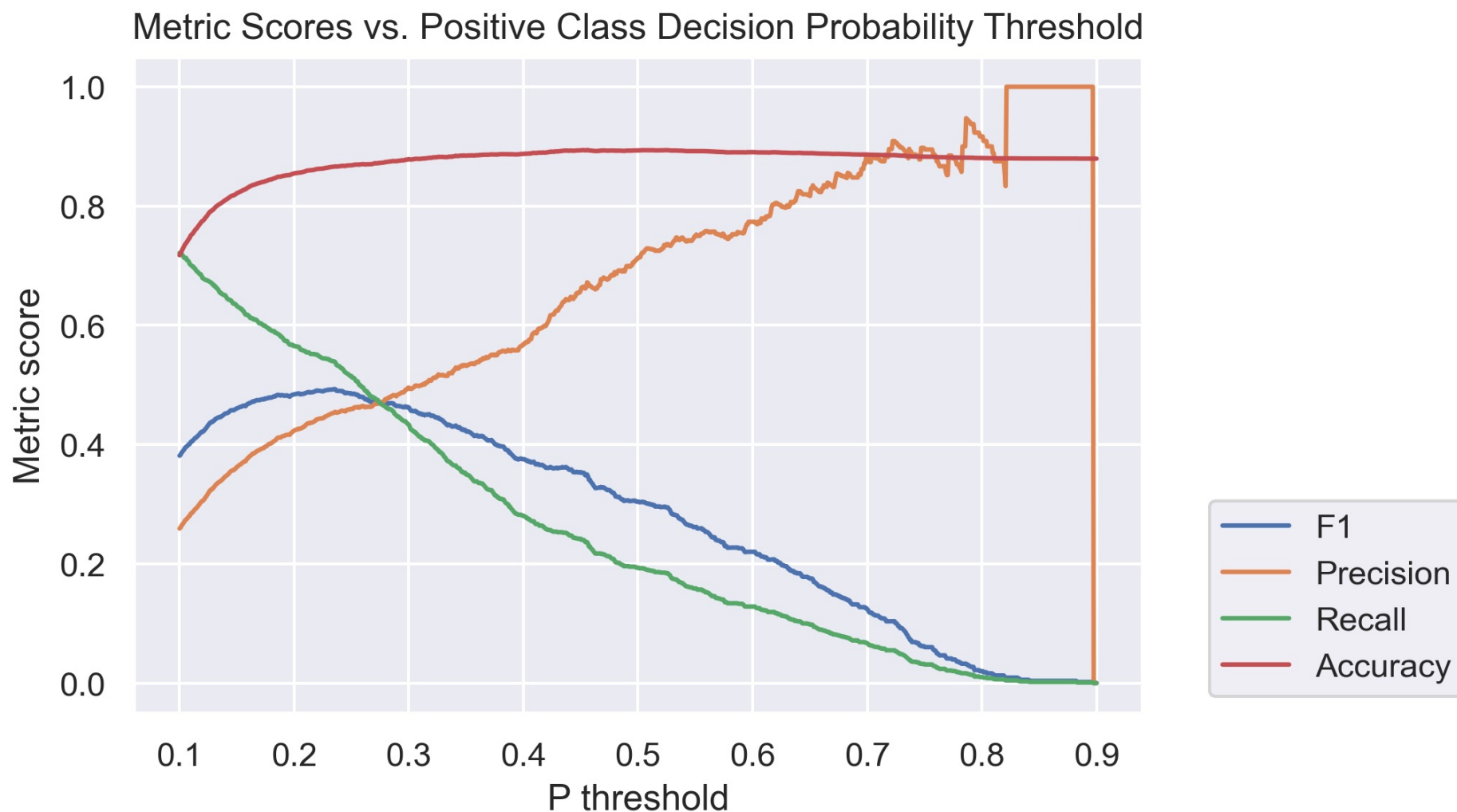


Closer look at important features

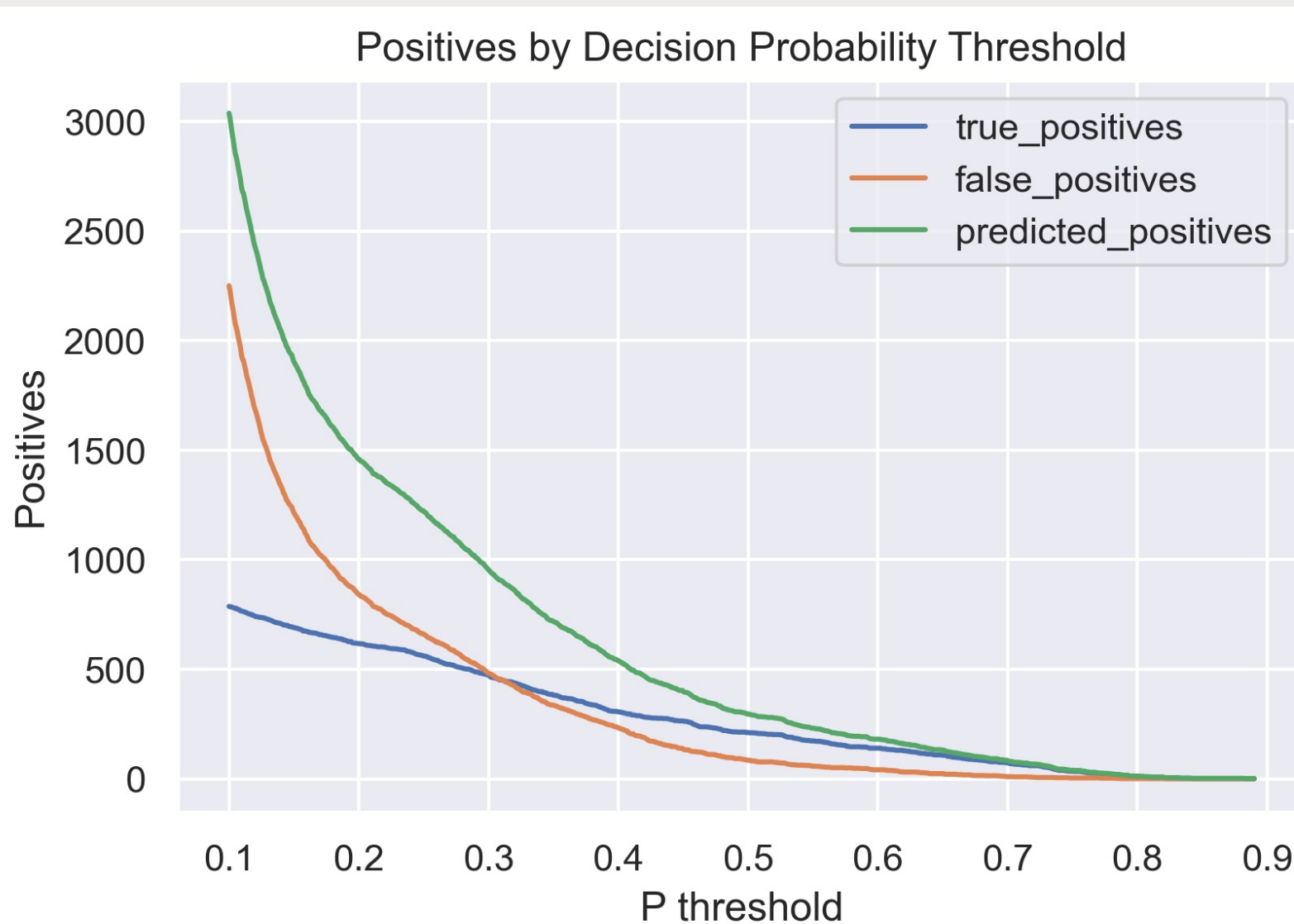
Graph idea

- Balance mean for negative and positive class ---- matplotlib bargrah
- Age (further classified to 20s, 30s, 40s--
- Job ---sns.scountplot
- Campaign(number of contacts during the campaign) ----

Solution: RF model w/ Probability Threshold Controlling



Solution: RF model w/ Probability Threshold Controlling



Appendix: Model Performance

MODEL	PRECISION SCORE ON POSITIVE CLASS	
Baseline KNN		
Tuned KNN		
Logistic Model		
Baseline Random Forest		