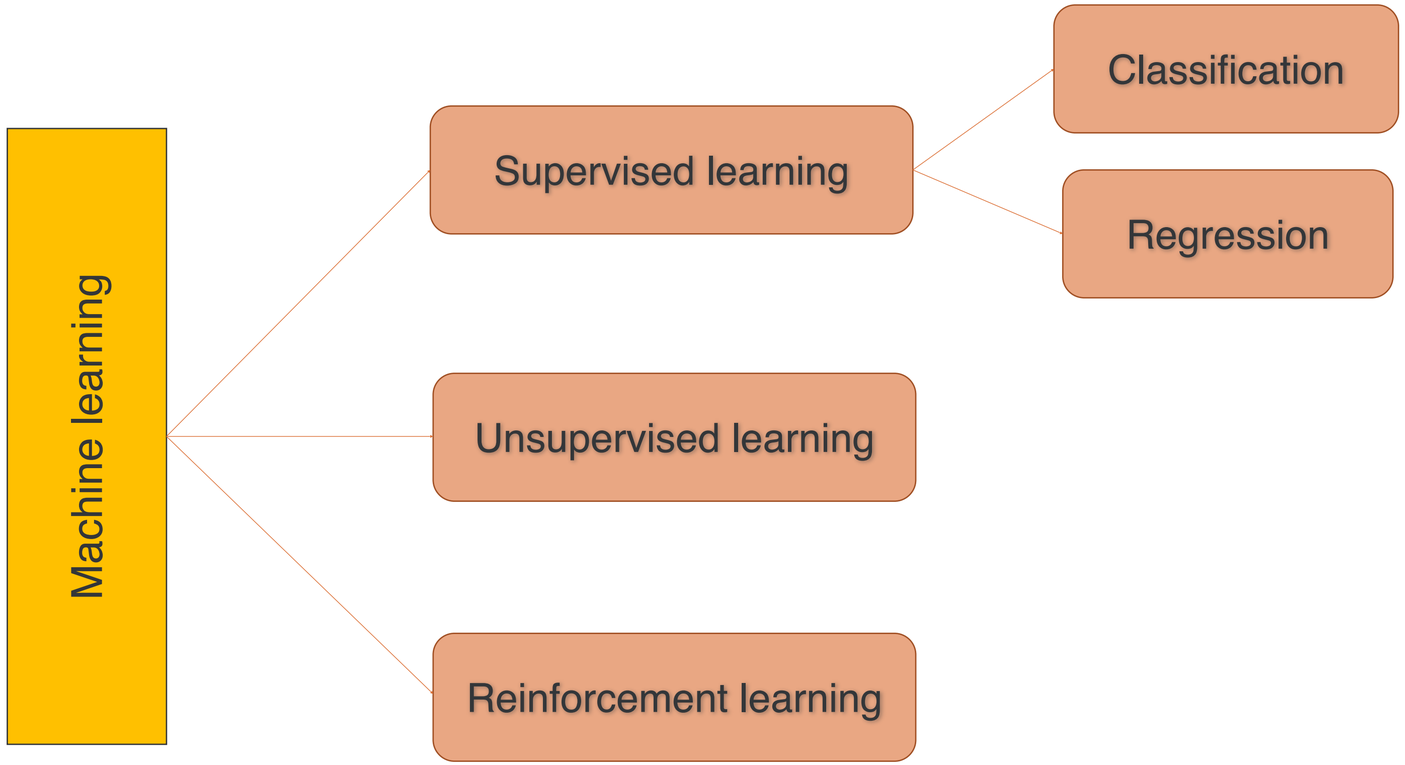
A complete analysis for predicting banking behavior by the classification method

Henry Yang (2021-6-5)

What is predictive modeling (machine learning)?

- Learn from data, identify patterns and make decisions with minimal human intervention.
- Extract business insight from the complex data.



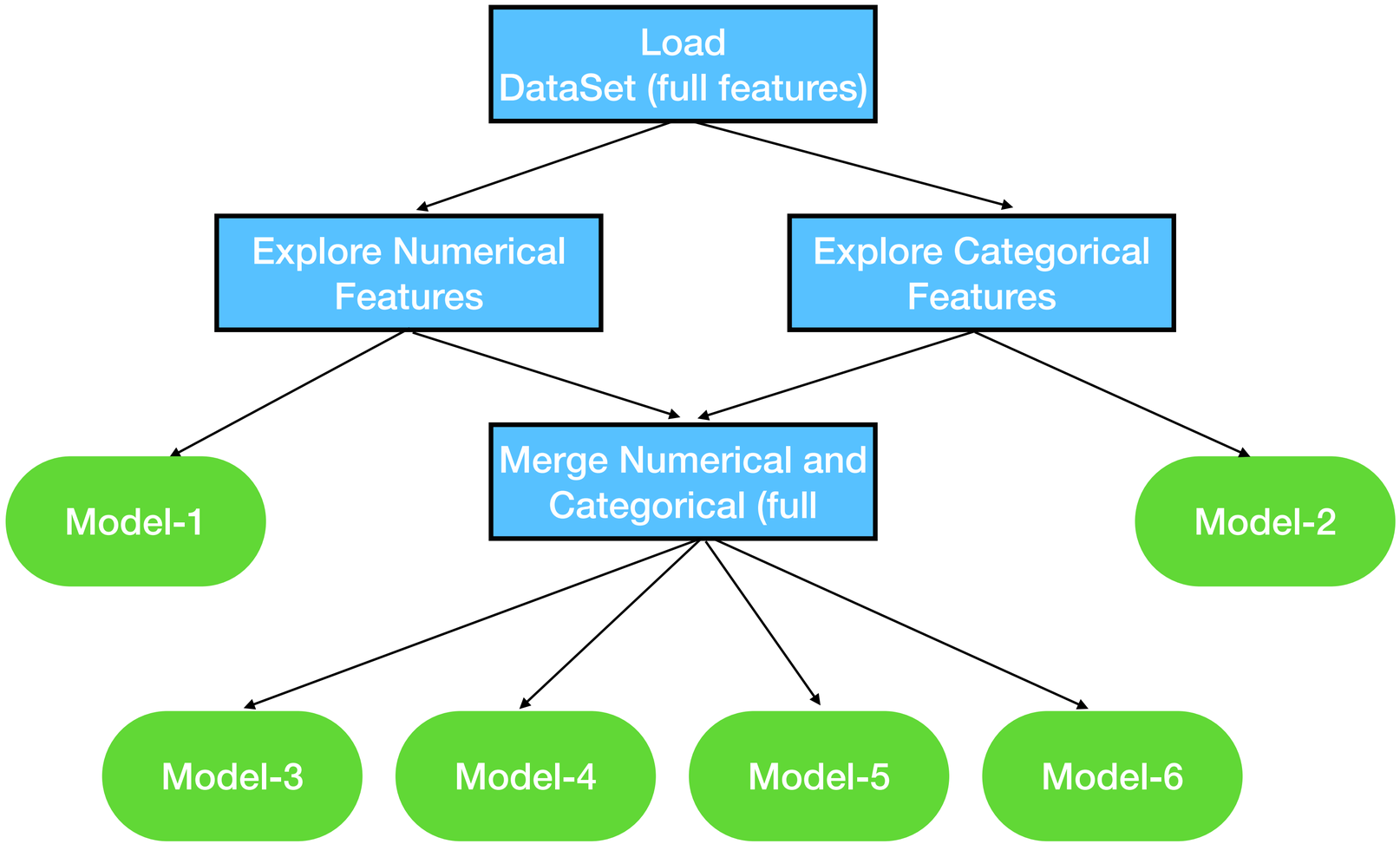
How the selected model can be used to score customer propensity to purchase?

In this study, I built 6 models based on the historical data. The best model is XGBoost (model-3 in the next slide). The data is randomly split as the training and testing part. The testing data was not seen during training. Therefore, the statistic validation metrics from the testing data should be very similar to these calculated using the new data.

- In our case, the attribute duration was removed from modeling. Therefore, prediction with the rest 19 features is realistic!

- The metrics AUC_ROC (0.80) and f1 score (0.72) is fairly good for prediction the likelihood for a new custom to purchase.

The diagram below shows the modeling procedures
(Model-3 is the best)



Which features are most predictive of purchase and what tactics or strategies these insights might suggest for the bank's marketing team?

In this study, the feature importance was calculated using the model based feature selection algorithm (XGBoost). The top three features are the followings:

feature	importance
nr.employed	0.343
poutcome_success	0.130
emp.var.rate	0.103

This suggests that number of employees (nr.employed), success from the previous marketing campaign (poutcome_success) and the employment variation rate (emp.var.rate) are the most important factors for the purchase.