

PA Workshop

**Marketing**

1. **What are the important determinants of a positive subscription (Y/N)?**

job, education, balance, housing, loan, contact, day, month, poutcome, duration, campaign

1. **How you tuned the model?**
2. 1st model: Regression on all independent variables and check for multi-collinearity. No obvious multi-collinearity among the independent variables is found.
3. 2nd model: Use the step() function as a variable selection technique, which returns accuracy rates of

* 0.901 on train set
* 0.903 on test set

1. 3rd model: Base on P values of the 2nd model, further manually delete variables that do not significantly add to the fit, which returns accuracy rates of

* 0.901 on train set
* 0.904 on test set

1. P value for the 3rd model closes to 0, indicating excellent overall fit.
2. Choose the 3rd model with the highest accuracy rate.

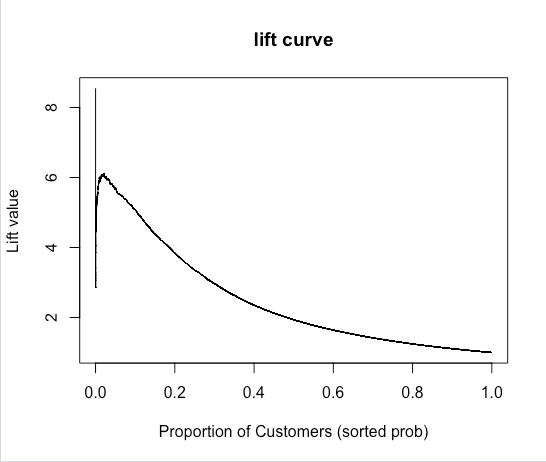
predict

0 1

0 11714 263

1 1033 554

1. **Use of lift chart for better targeting?**



By contacting 20% of customers, randomly and using no model, we should get 20% of responders. The lift value is 4 for 20% of customers, which means by using this model we could get 4 times as many respondents as if we use no model, amount to 80% of customers.