Decision Tree Exercise

Please submit a PDF or Word document

In the scenario for this exercise, you are a data analyst at a national charitable organization. Your organization seeks to use the results of a previous postcard mail solicitation for donations to better target its next one. In particular, you want to determine which of the individuals in your mailing database have characteristics similar to those of your most profitable donors. By soliciting only these people, your organization can spend less money on the solicitation effort and more money on charitable concerns.

Analyze donor dataset, and present:

Submit the figure output of the first 4 questions.

1. Use a graph to show the distribution of LIFETIME\_GIFT\_COUNT
2. Use a graph to show relationship between MEDIAN\_HOME\_VALUE and LIFETIME\_GIFT\_COUNT
3. Build a decision tree with all variables to predict if an individual should be targeted as a donor

(target variable: TARGET\_B).

(random\_state=80). Use 80% in the training set.

1. Prune the tree with your choice of configurations.
2. Present the accuracy, sensitivity, and specificity of original and pruned trees.
3. Recommend promotion strategies (to promote donations) for the company based on the pruned tree. (write down all decision rules learned from the pruned tree, reference the code book for the meaning of variables)

|  |  |
| --- | --- |
| **Variable** | **Description** |
| CARD\_PROM\_12 | Number of card promotions sent to the individual by the charitable organization in the past 12 months |
| DONOR\_AGE | Age as of last year's mail solicitation |
| INCOME\_GROUP | one of 7 possible income level groups based on a number of demographic characteristics |
| LIFETIME\_CARD\_PROM | total number of card promotions sent to the individual by the charitable organization |
| LIFETIME\_GIFT\_COUNT | total number of donations from the individual to the charitable organization |
| MEDIAN\_HOME\_VALUE | median home value (in $100) as determined by other input variables |
| MEDIAN\_HOUSEHOLD\_INCOME | median household income (in $100) as determined by other input variables |
| MONTHS\_SINCE\_FIRST\_GIFT | number of months since the first donation from the individual to the charitable organization |
| MONTHS\_SINCE\_LAST\_GIFT | number of months since the most recent donation from the individual to the charitable organization |
| MONTHS\_SINCE\_LAST\_PROM\_RESP | number of months since the individual has responded to a promotion by the charitable organization |
| MONTHS\_SINCE\_ORIGIN | number of months that the individual has been in the charitable organization's database |
| NUMBER\_PROM\_12 | number of promotions (card or other) sent to the individual by the charitable organization in the past 12 months |
| PER\_CAPITA\_INCOME | per capita income (in $) of the neighborhood in which the individual lives |
| RECENT\_CARD\_RESPONSE\_COUNT | number of times the individual has responded to a card solicitation from the charitable organization since four years ago |
| RECENT\_RESPONSE\_COUNT | number of times the individual has responded to a promotion (card or other) from the charitable organization since four years ago |
| WEALTH\_RATING | one of 10 possible wealth rating groups based on a number of demographic characteristics |