Gallup

A white background with black text

AI-generated content may be incorrect.



A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.



Glint



A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.



A screenshot of a computer

AI-generated content may be incorrect.



Result under Random Forest

A computer screen shot of a program

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.Since the mean of those variables are around 60, it is difficult to predict all DVs in above 70. Thus, the predicted range is from 60-70.

A close-up of numbers

AI-generated content may be incorrect.

Also, it has shown that scores 80 is the threshold such that the predicted satisfy scores is slightly affected by those questions when their scores pass 80. Therefore, it would try to find how many scores needed to increase when we want to satisfy scores reach score 80.

A screenshot of a computer screen

AI-generated content may be incorrect.

The table orders by n\_capped ascendingly and the required increase is calculated individually. Focusing on Q21 and Q22, the required\_increase of them are 84, which means that they have met the threshold. Especially for avg\_score\_Q22, although it is the most importance variables in this model, increasing the score of it would only have slightly increased in Satisfy scores that never reach score 80. The Top 5 questions need to be improved are avg\_score\_Q30, 8, 11, 29 and 20 respectively. Noted, if for increasing with variables combinations, at least 15 variables are needed to reach score 80.

The overall plot with importance coloring(Higher importance, darker color) A graph with different colored lines

AI-generated content may be incorrect.

For knowing which questions are more important for improving the DV score, we have tested the variables importance when avg\_scores increase in 20 score

[Noted:

1. Less n\_capped means more can be improved.

2. Importance\_order based on the original data

3. ~~mtry is fixed to 1 means only 1 variables would be tried at each split~~; ntry is ntree=n\_predictors/2=11 for reducing overfitting.

4. Cross validation has done in cutting in 10 splices.

5. %IncMSE means how much prediction error increases without the variable); Higher %IncMSE → More important variable

6. For the required\_increase, it assumes that all other variables are unchanged expect the target variables.