

PREDICTIVE ANALYSIS

PROSPECTIVE DONOR RECOMMENDATION
BASED ON PREDICTIVE ANALYSIS DONE ON THE HISTORICAL DATA
BY IYIOLA OYEWUMI

PROBLEM STATEMENT

- The client seeks to use the results of a previous postcard mail solicitation for donations to improve outcome in the next campaign.
- We want to determine which of the individuals in their mailing database have characteristics similar to those of their most profitable donors.
- By soliciting only these people, our client can spend less money on the solicitation effort and more money on charitable concerns.

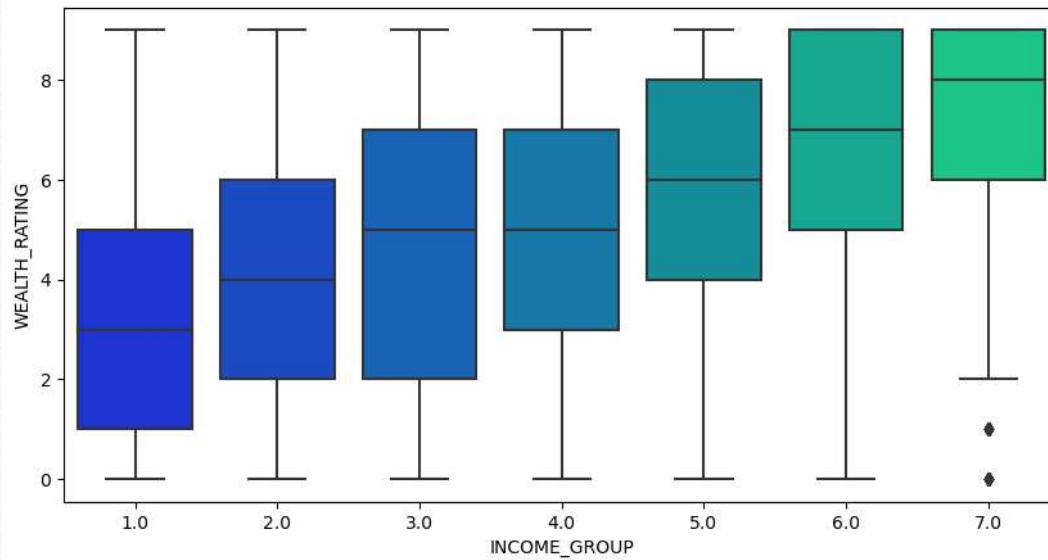
AVAILABLE INFORMATION

- Donor Raw Data: this is a historical data, containing previous donor details. Target B column, provides information whether they have donated in past campaign or not.
- Prospective Donors: this is a list of new contact; your client is interested in reaching out to in the next campaign they intend to carry out soon.

METHODOLOGY

- Data Preprocessing : This was done to investigate the provided raw data and ensure that the data is clean from inaccuracies that might affect the result of the model.
- Some of the data in the columns have missing values e.g Wealth rating, Income group, Donor age, Month since the last response.
- The median value was used to fill in these missing values and through evaluation, I noticed the wealth rating is dependent on the income group.

Box plot evaluation shows the higher the income group, the higher the wealth rating. This was used to evaluate the missing values for the wealth rating.



Models used to test and train the data

- Logistic Regression
- Linear Discriminant Analysis
- K neighbors Classifier
- Decision Tree Classifier
- Random Forest Classifier
- Gaussian NB
- SVC

RESULTS OF THE TEST

- LR: 0.748459 (0.007114)
- LDA: 0.749556 (0.007145)
- KNN: 0.700984 (0.007141)
- CART: 0.631130 (0.008449)
- RF: 0.748093 (0.008503)
- NB: 0.691416 (0.012570)
- SVM: 0.748825 (0.006903)
- The RF model accuracy is = 74.81% which is the model I adopted for my prediction

Result of the model on the prospective donor data: The control is the identification number of the donors. The prediction score of 1 depicts a likely donor. Result below.

- CONTROL_NUMBER prediction
- 1 142 1
- 7 665 1
- 9 743 1
- 10 788 1
- 14 968 1
-
- 2127 190018 1
- 2128 190099 1
- 2134 190412 1
- 2142 190840 1
- 2147 191710 1
- [448 rows x 2 columns]

Model Evaluation

- Evaluated model accuracy = 92.3873261528809%
- The prospective donors have been evaluated with the control number as their identity. This surely will help the client to focus more attention on the predicted prospective donors for the next campaign.

Thanks
