Dan Schumacher, 5/7/24

Enhancing Donor Prediction Accuracy for the National Veterans' Organization through Advanced Modeling Techniques

Introduction
Non-profit organizations significantly depend on donations to fund their operations and deliver services. The National Veterans' Organization (NVO), with a vast database of over 13 million potential donors, primarily utilizes direct mail for fundraising. However, with an average donation return of $13 against a mailing cost of $13.60 for every 20 letters sent, the organization faces a net loss. This scenario underlines the critical need for a more targeted marketing strategy that reduces costs and enhances the efficiency of fundraising efforts.

Objectives
The objective of this project is twofold:
1. Utilize various predictive modeling techniques to optimize the direct mail campaign, aiming to increase the profitability of the National Veterans' Organization.
2. Apply and showcase the modeling techniques learned in Dr. Campbell's class and beyond this semester, emphasizing practical applications in a real-world context.

Data Sources
The data was made available on Canvas by Dr. Campbell.

Details of the Data
The training data contains 3,000 records and is evenly balanced between Donors and Non Donors. There are 22 variables in the dataset. These include:
- Zip code category, homeowner status, number of children, income, sex, wealth category, homevalue, median family income, average family income, number of promotions sent to the individual, lifetime number of promotions received by the individual, sum of all donations given, dollar amount of last gift, months since last donation, number of months between previous 2 gifts, average gift from donor.

Types of analysis Performed
I performed both traditional modeling and utilized an artificial neural network to generate my predictions.

Dan Schumacher, 5/7/24

Traditional Modeling

In the traditional modeling, I started by creating dummy variables of all of the predictors. Next I removed predictor variables that were highly correlated with one another. Starting with all predictors, I would calculate the Variance Inflation Factor (Vif) of each variable, then if any variable had a Vif greater than the cut off (I used 10), I would remove the variable and recalculate the Vif without that variable. This process repeated until all Vifs were under the cut off. At the end of the process I had removed five variables

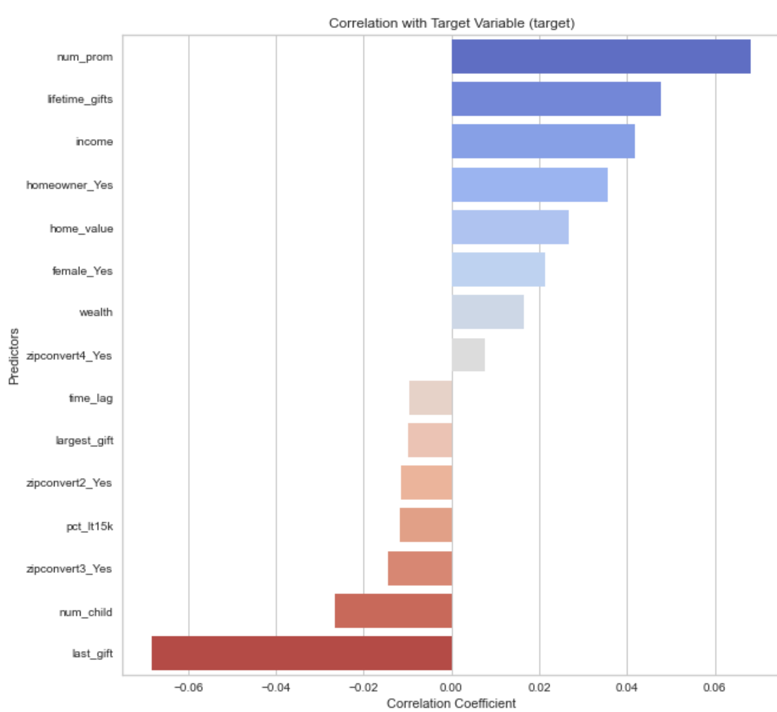
REMOVED: ['zipconvert5_Yes', 'avg_fam_inc', 'months_since_donate', 'med_fam_inc', 'avg_gift']

I fit the data to a Logistic Regression, LDA, QDA, KNN (k=1), and a Naive Bayes classifier.

The results are shown below. Our top performer on both the dev and test set was LDA. It is also noteworthy that LDA had the highest F1 score, achieving a good balance between precision and recall.

name	tp	tn	fp	fn	acc	prec	recall	f1	test_acc
Logistic Regression	167	151	145	137	0.530000	0.535256	0.549342	0.542208	0.533333
LDA	163	166	149	122	0.548333	0.522436	0.571930	0.546064	0.558333
QDA	26	268	286	20	0.490000	0.083333	0.565217	0.145251	0.525000
KNN	148	153	164	135	0.501667	0.474359	0.522968	0.497479	0.475000
Naïve Bayes	38	263	274	25	0.501667	0.121795	0.603175	0.202667	0.516667

To get a rough idea of what variables our model is using to generate predictions, we can make a correlation chart.



Here we can see that the larger the bar, the more correlated the variable is with the target. Red bars show negative correlations while blue bars show positive correlations. One thought I had, to achieve superior test accuracy, was to only use the most correlated variables on each side. I tested only using variables with an absolute correlation coefficient above .05. Unfortunately, This did not improve results.

Speaking of which, it is worth mentioning that I attempted many more models, and much more complex ones. I had a total of 79 entries on the competition's website. For the more complex models I tried transforming the continuous variables by log, squared, cubed, square root, inverse, boxcox, sigmoid, sine and cosine transformations. After Vif, in the complex modeling, I had a total of 84 variables. However, they did not help. Both the dev scores and test accuracy were several percentages shorter across each category.

ANN

Another approach that I took was making an artificial neural network for the classification task. This network had the following parameters:

1. An input layer with both categorical and continuous features processed via embedding and batch normalization respectively
2. Multiple hidden layers with configurations of [200, 100] neurons each incorporating ReLU activation and dropout regularization set at 0.4
3. Final output layer using a softmax function for binary classification
4. Robustly scaled input variables to boost performance
5. 100 epochs of training with an early stopping patience of 5
6. Cross validation with 20 splits

This approach achieved much better dev accuracy with an OOB accuracy of 59%. However, at test time it performed practically the same as LDA, and it had the downside of being much more complex, less interpretable, and doesn't always produce the same results.

Results

The models displayed varying levels of success:

- Accuracy Metrics: Each traditional model was evaluated based on accuracy, precision, recall, and F1-score. The LDA emerged as the most accurate model, as well as the one with the highest F1, showcasing a balance between precision and recall.
- Feature Importance: Looking at the correlation with target variable plot, we can see that the target variable is most strongly associated with number of promotions, and negatively associated with last gift. This loosely suggests that individuals who have donated recently are more likely to donate again, and people who receive a lot of promotions tend to donate more.

Discussion

The results indicated that, with all their bells and whistles, advanced modeling techniques such as artificial neural networks do not always have better predictive power then their traditional less complicated counterparts. However, none of the variables exhibited extremely high correlations with the target, suggesting potential limitations in the dataset's explanatory power.

Dan Schumacher, 5/7/24

Conclusion and Future Work

The study successfully demonstrated the potential predictive models to enhance donor prediction accuracy, thereby aiding the NVO in optimizing its fundraising strategies. For future work, further experimentation with ANN hyperparameters (such as learning rates, epochs, batch sizes, dropout rates) is recommended. Additionally, obtaining more comprehensive data, including demographic information, military service details, and any other variables that could potentially have a higher correlation with the target variable, would be useful. Having more variables more closely correlated to the target would significantly improve predictive power.

Appendices

The appendices contain two different notebooks. The first is the simple modeling I performed to achieve the best results. The second includes transformed variables (test accuracy was not included, as all results were in the mid 40s). Finally there is a python script that is my artificial neural network.