Predicting Marketing Campaign Response

GROUP 4

Aditee Bhattarai
Lindiwe Mukurazita
Rajadurga Ganesan
Utsav Pradhan

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

Executive Summary

Our insurance company is planning to introduce a new product and would like to engage with the customers through a new marketing campaign. The company would like to identify which customers to target so that the campaign will be a success and its objectives realized. We proceeded to understand and evaluate our existing customer dataset through exploration, analysis, and interpretation. We used for analytical tools in this process: Cluster Analysis, Decision Trees Analysis, Naïve Bayes Method, and Artificial Neural Networks (ANN).

Through Cluster Analysis, we were able to identify the consumer base that will yield us the best results. Along with Decision Trees Analysis, we can discern the most important variables for our analysis. Naïve Bayes and ANN are classification methods that allow us to train a model to learn the trends in the dataset, and deliver a prediction of how the customers will react. Among the two methods, ANN resulted in a better model with better accuracy and predictive power.

From our findings, we have realized the most important variables, and the ANN model has given us a relatively accurate model. By working in conjunction with Cluster Analysis, our model will be able to concentrate the results on the desired consumer base. With this strategy, we can accurately predict a majority of the customers' responses, limit the number of incorrect predictions and minimize the potential losses incurred, and maximize the positive responses towards our marketing campaign for the new product.

INTRODUCTION

Marketing campaign management is complex and requires a lot of time to monitor. Often due to the vast volumes of data, it is difficult for the organizations to capture the entire picture of the campaign performance. Therefore, proactive marketing management cannot be achieved without predictions.

With a campaign prediction model, we will get insights into the future marketing performance trends that will help us adjust and edit our campaign to reach better results, save costs and improve our company's revenues by investing in the best possible campaigns as well as reduce the chances that money and marketing budgets are wasted on poor performing campaigns.

Business Problem & Objective:

An insurance company is planning a marketing campaign for a new insurance product and would like to strategically choose which customers to market to. The team has identified that the response rate for the previous marketing campaign was low. Out of 9134 customers that the company reached out to, only 1308 customers responded positively, getting us a response rate of only 14.3%. Our team wants to identify and explain the variables that affect a customer's response to such campaigns, to make recommendations to increase the response rate for the company in future campaigns. We also want to identify customers who would bring the most value to the company, so that customer acquisition and customer retention plans can be made accordingly.

The variables in our data that we believe are most important to our analysis are "Response", "Renew Offer Type", and "Total Claim Amount". The variables that we believe are most important in predicting response are "Employment Status", "Marital Status", "Location Code", and "Sales Channel". Variables important in determining a customer's profitability are "Customer Lifetime Value", "Total Claim Amount" and "Monthly Premium Auto".

DATA

Description:

The dataset has a total of 9134 entries and 22 columns. The dataset does not have missing values. The types of variables are as follows:

Nominal Variables: State, Employment Status, Gender, Location Code, Marital Status, Policy Type, Policy, Renew Offer Type, Sales Channel, Vehicle Class, Response

Ordinal Variables: Education, Coverage, Vehicle Size

Numerical Variables: Customer Lifetime Value, Income, Monthly Premium Auto, Months Since Last Claim, Number of Policies, Total Claim Amount

Below is the statistics summary of numerical variables:

	Customer Lifetime Value	Income	Monthly Premium Auto	Months Since Last Claim	Number of Policies	Total Claim Amount
min	1898.008	0.00	61.00000	0.000	1.00000	0.099007
1st Qu.	3994.252	0.00	68.00000	6.000	1.00000	272.258244
Median	5780.182	33889.50	83.00000	14.000	2.00000	383.945434
Mean	8004.940	37657.38	93.21929	15.097	2.96617	434.088794
3rd Qu.	8962.167	62320.00	109.00000	23.000	4.00000	547.514839
Max	83325.381	99981.00	298.00000	35.000	9.00000	2893.239678

Table 1: Statistic summary of numeric variables

The variable "Customer" is a unique identifier and is not included in the initial data exploration process. We have also not used the variable "Effective to Date" and treated it as a character variable.

Exploratory Data Analysis:

The variables we are interested in exploring more are "Response", "Customer Lifetime Value", and "Total Claim Amount". After an initial data exploration, we believe that "Location Code", "Employment Status", "Marital Status" and "Vehicle Class" are important for our analysis. Through our initial data exploration, we have found that we have more "Female" customers than "Male". Most of the customers are "Employed" and/or "Married" and/or live in a "Suburban" location.

We also found out that most customers have taken the "Basic Coverage". The mean "Monthly Premium" amount paid by customers is $93, and the mean "Total Claim Amount" is $434. Among the states, "California" and "Oregon" have recorded higher responses. The most responsive group for the last marketing campaign were people living in the "Suburban" Location Code (*Figure 1*). We found that this group of customers had the highest claim amount, but they also pay a slightly higher monthly premium.

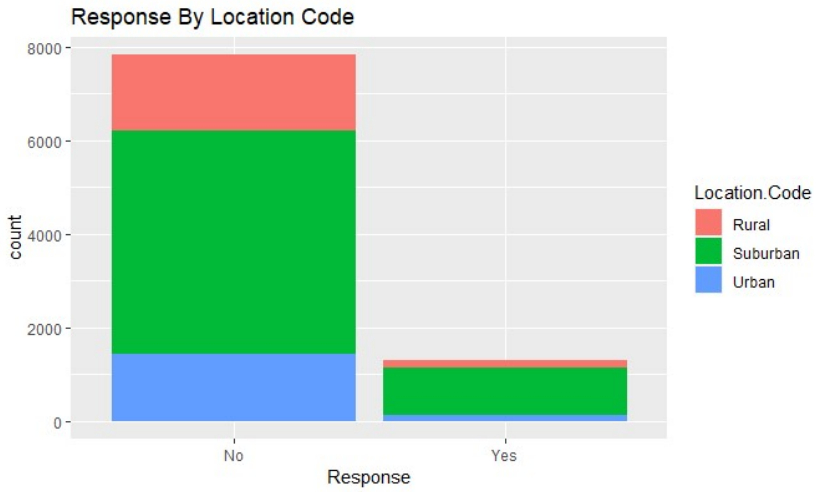

Figure 1: Response by Location code

We found that there were absolutely no responses for "Offer 4" on our previous Renewal Offer, and very little responses for "Offer 3" (*Figure 2)*.

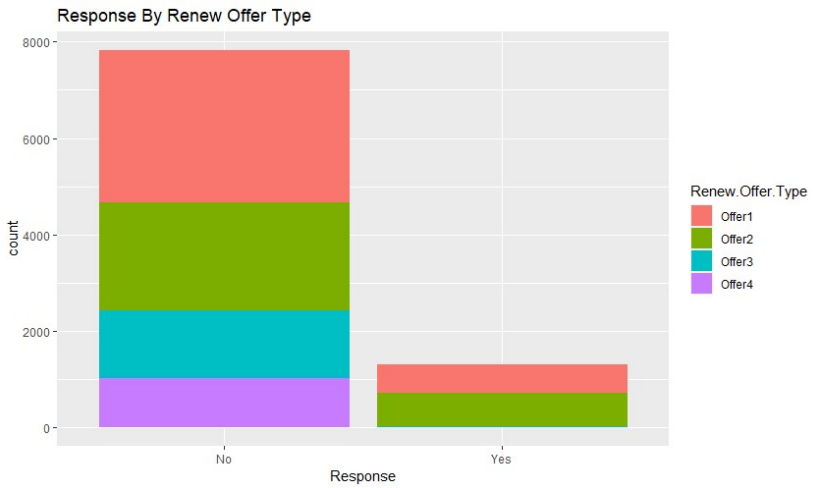

Figure 2: Response by Renew offer type

We also found that people with "Luxury Cars" and "Luxury SUVs" have a higher "Customer Lifetime Value" and pay a higher "Monthly Premium" Amount. They also have higher "Total Claim Amount" than people with other Vehicle Class.

Our initial exploration shows that while "Single" under "Marital Status" have lower "Customer Lifetime Value" (*Figure 3)* and pay the average premium amount, their total claim amount is relatively higher than other groups (*Figure 4)*.

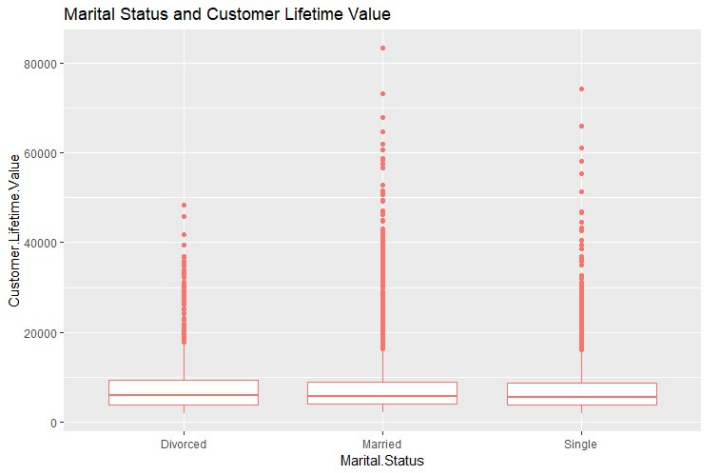

Figure 3: Marital Status vs Customer Lifetime Value

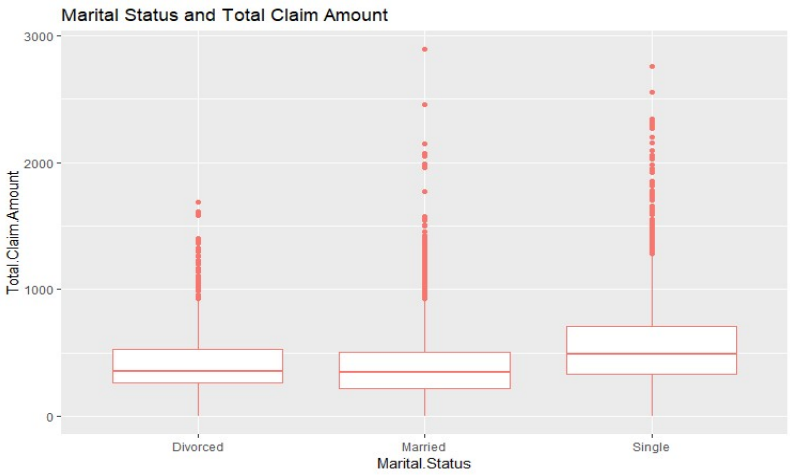

Figure 4: Marital Status vs Total Claim Amount

We also found that response rates are highest when the "Sales Channel" is "Agent".

Pre-Processing of the Data:

Before carrying out any analyses, the data was split into vectors for convenience. The nominal categorial variables were grouped as facs, ordinal categorical variables were grouped as ords and numerical variables were grouped as nums. Variables under ords were also ordered.

Binarization:

Bins were created in the original data frame for Response variable and Marital Status variable by coding them as 0 and 1. Other factor variables under facs and ords were also binarized to create dummy variables. These were the combined with the other variables in a new data frame ms_dum.

Variable Transformation:

The numerical variables were standardized and normalized using Center Scale and Yeo-Johnson method. We can observe the difference in the distribution of data under variable Total Claim Amount in *Figure 5.* And Customer Lifetime Value in *Figure 6.*

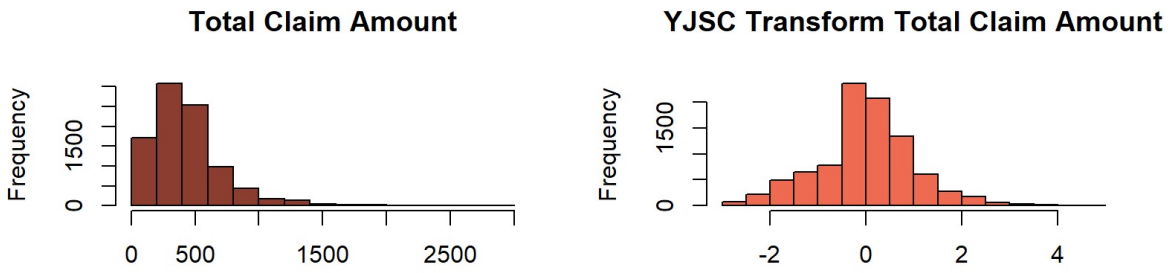

Figure 5: Distribution of data under 'Total Claim Amount'

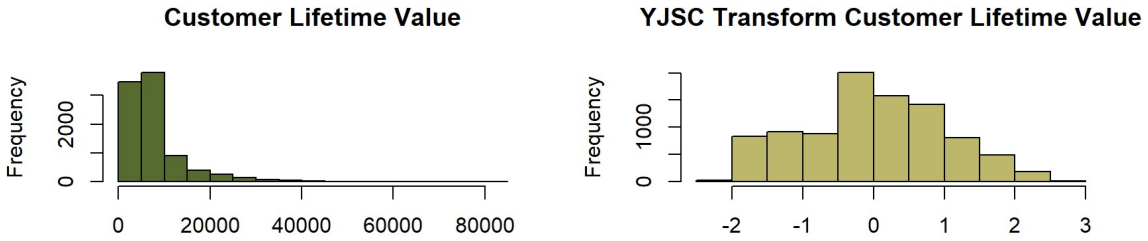

Figure 6: Distribution of data under 'Customer Lifetime Value'

Class Imbalance:

There exists a class imbalance in our dataset. Our target variable "Response" has disproportionate numbers of "No" than "Yes".

For Decision Tree Analysis, we oversampled our training data, as we did not want to lose any possible nodes. When we ran the analysis on the imbalanced training data, we were able to get only 3 nodes, which did not provide much insight into the characteristics of our data. By oversampling, we were able to get explainable results.

For Naïve Bayes, we again over sampled our data. For Artificial Neural Networks, we under sampled our data.

Data Quality:

The data frame was checked for duplicate values. There were none. The dataset also has no missing values. To confirm our finding, we also checked to see if there were any missing values that were present in our data as "0". We found that all the 0 values present in our data were under variable Income, and these 0 values were the incomes of unemployed people. Thus, we confirmed that there are no missing values in the form of 0.

We used the z-score method to detect outliers. We had 45 outliers, all under Total Claim Amount. Since our data consists of over 9000 records, we decided to get rid of the outliers as removing 45 entries will not affect our analysis a lot.

ANALYSIS

CLUSTER ANALYSIS:

We chose to begin our analysis with Cluster Analysis as we wanted to check for any pattern of demographics or behavior among the customers, and their response to our marketing campaign. We also believe that cluster analysis will serve as a basis for further analysis of the data. Before beginning cluster analysis, we wanted to check for correlation among our numerical variables to identify variables that could be merged into one, if any.

The significant correlations were among Customer Lifetime Value and Monthly Premium Auto of 0.39, between Income and Total Claim Amount of -0.36 and between Monthly Premium Auto with Total Claim Amount of 0.60. To make sure that the relationship between these variables is not too strong, we visualized them in a scatter plot and used linear regression, as seen in *Figure 7*

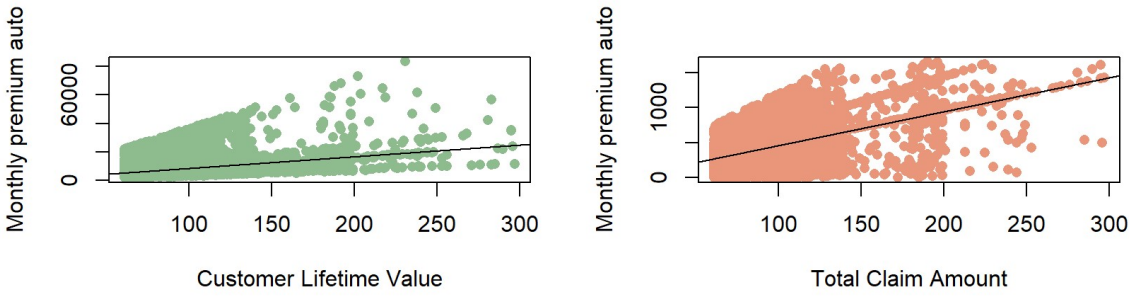

Figure 7: Scatterplot (Customer Lifetime Value vs Monthly premium auto)

Since neither of the relationships seem too strong, we decided to move forward with the variables as they were.

Hierarchical Cluster Analysis:

Due to the mixed nature of our data, we decided to use Gower Distance to calculate the distance between data points. For clustering, we used Wards Method. The Cophenetic Correlation value for this method is 0.53, which is higher than Complete Linkage Method.

In the Dendrogram seen in *Figure 8,* we can spot 4 clusters. To confirm our observation, we validated our cluster solution using both Within Cluster sum of squares and Silhouette method, as shown in *Figure 9.* Both methods showed that the optimal value for k is indeed 4.

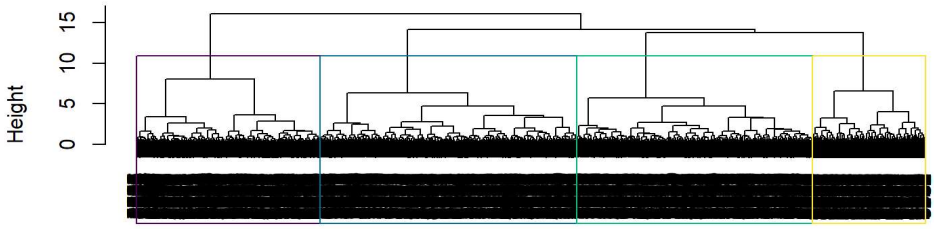

Figure 8: Dendrogram – Wards Method

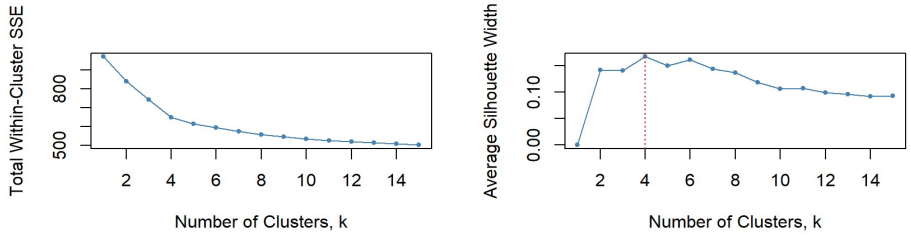

Figure 9: Validating Cluster solution

Explanation of our Cluster Solution:

Cluster	N	Gender	Location	Marital Status	Renew Offer	Average Lifetime Value	Average Monthly Premium	Average Total Claim Amount	Sales Channel	Response (yes)
1	2999	2995 Females, 4 Males	1335 Suburban	Married 2095	Offer 1- 1134, Offer2 – 877	$8259.80	$91.77	$344.17	Agent 1029	0
2	2108	1021 Female, 1087 Male	1917 Suburban	Single 1267	Offer 1- 1094, Offer 2 – 472	$7501.56	$92.6	$587.74	Agent 795	0
3	2719	9 Female, 2710 Male	1473 Suburban	Married 1826	Offer 1 – 919, Offer 2- 878	$8061.94	$92.4	$379.94	Agent 977	6
4	1338	642 Female, 696 Male	1009 Suburban	696 Married	Offer 1- 589, Offer 2 - 684	$7857.63	$94.1	$447.42	Agent 660	1302

Table 2: Explanation of Cluster solution

As we can see from above table, Cluster 4 is the only cluster that has responded to our previous campaign. It consists of 1338 customers. The positive response rate is 97.3%, which is exceptional. The cluster consists of both males and females, proportionately. More than half of the group is married, and around 75% of them live in Suburban areas. Most of them were offered Renew Offer Type 2, closely followed by renew Offer Type1. They pay the highest average monthly premium of $94.1.

Although most people in this cluster are employed, 661 of them are unemployed, which is the highest number of unemployed people among all clusters. They do not bring in most value to the company, and their total claim amount is the second highest among our clusters. It appears the company has got the marketing campaign for this cluster right and should continue the activities they did in the last campaign.

However, if we look at Cluster 1, we see that this group of customers bring in the highest values to the company. They also have the lowest claim amount among all the clusters. There were 0 responses from this cluster during our last campaign. This cluster was offered Renew offer 1 and Renew offer 2. This cluster is formed by mostly females, where 69.85% of them are married. 44.5% of them live in a suburban location. The company should investigate the offers that is being sent to them, as those offers are yielding no response.

Another cluster with a high lifetime value and low claim amount is Cluster 3. Most of them are men and most of them are married. They too mostly live in the Suburban locations and were offered Offer 1 and Offer 2. These offers yielded only a 0.002% response from the cluster. We would recommend that the company survey a small sample each of these two clusters to identify their preferences, in order to be able to offer them the right deal. If our company can get more responses from these clusters, it would be most beneficial.

External Cluster Validation:

We used Adjusted Rand Index to validate our cluster solution, externally. Our solution only has a score of 0.2141676, which is poor recovery. Even though this solution is not the most reliable, we believe the cluster information found from this analysis will prove useful to the company.

DECISION TREES:

We decided to use Decision Tree to determine what would be our most important predictors for "Response". Our target variable Response had a class imbalance, with only 20% "Yes" in our dataset, which we corrected with oversampling.

Using an 85/15 split for the training and test, we ran the initial base model.

cp	Accuracy	Kappa
0.003251245	0.8905005	0.4981629

The optimal model had accuracy of 89% and a cp value of 0.00325. The Kappa value is 49.81% therefore, it is a moderate model.

Using random oversampling, we get the following resampling results:

cp	Accuracy	Kappa
0.008769277	0.7501001	0.5001999

Here the optimal model has a cp value of 0.00877 and an accuracy of 75%. The Kappa value is 50.02%.

After running our training model with an oversampled dataset, we saw that the most important variables are:

- Renew Offer Type
- Employment status
- Total Claim Amount
- Monthly Premium Auto
- Location Code
- Sales Channel

Below is our Decision Tree:

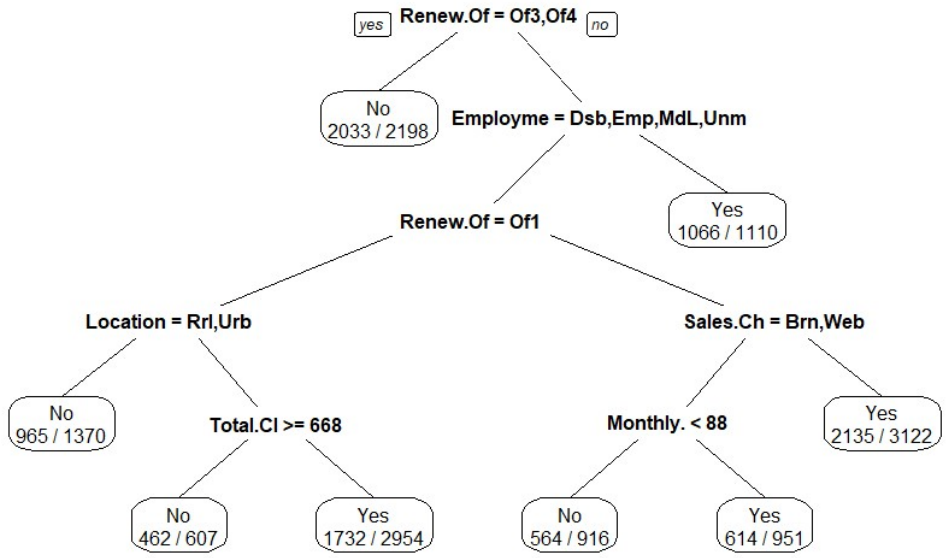

Figure 10: Decision Tree model

Analysis of Decision tree nodes
1) root 13228 6614 No (0.50000000 0.50000000)
2) Renew.Offer.Type=Offer3,Offer4 2198 165 No (0.92493176 0.07506824) *
3) Renew.Offer.Type=Offer1,Offer2 11030 4581 Yes (0.41532185 0.58467815)
6) EmploymentStatus=Disabled,Employed,Medical Leave,Unemployed 9920 4537 Yes (0.45735887 0.54264113)
12) Renew.Offer.Type=Offer1 4931 2282 No (0.53721355 0.46278645)
24) Location.Code=Rural,Urban 1370 405 No (0.70437956 0.29562044) *
25) Location.Code=Suburban 3561 1684 Yes (0.47290087 0.52709913)
50) Total.Claim.Amount>=667.6793 607 145 No (0.76112026 0.23887974) *
51) Total.Claim.Amount< 667.6793 2954 1222 Yes (0.41367637 0.58632363) *
13) Renew.Offer.Type=Offer2 4989 1888 Yes (0.37843255 0.62156745)
26) Sales.Channel=Branch,Web 1867 901 Yes (0.48259239 0.51740761)
52) Monthly.Premium.Auto< 87.5 916 352 No (0.61572052 0.38427948) *
53) Monthly.Premium.Auto>=87.5 951 337 Yes (0.35436383 0.64563617) *
27) Sales.Channel=Agent,Call Center 3122 987 Yes (0.31614350 0.68385650) *
7) EmploymentStatus=Retired 1110 44 Yes (0.03963964 0.96036036) *

Table 3: Analysis of Decision tree nodes

Our decision tree has 53 nodes; therefore, the output is too complex to be visualized. However, we can see that our root node is balanced with 6614 'Yes' responses and 6614 'No' responses.
From 2198 observations for Renew type offer3 or 4, we had 2033 'No' responses. Some of the other terminal nodes had the following observations:

- From 1370 observations for Rural /Urban Location had 965 'No' responses
- Total claim amount >=667, 1732 'Yes' responses and 462 'No' responses
- Total claim amount<667,1222 'Yes' responses
- Employment Status=Retired 1066 'Yes' responses and 44 'No' responses

	Training	**Testing**
Sensitivity	0.8386755	0.8367347
Specificity	0.6084064	0.6238218
Pos Pred value	0.6817009	0.2719735
Neg Pred Value	0.7904145	0.9578947
Precision	0.6817009	0.2719735
Recall	0.8386755	0.8367347
F1	0.8436553	0.7104623
Prevalence	0.5000000	0.1438004
Detection Rate	0.4193378	0.1203228
Detection Prevalence	0.6151346	0.4424065
Balanced Accuracy	0.7235410	0.7302782

Table 4: Decision tree- training & testing performance

As an assessment of the goodness of fit for the model, we can see that the sensitivity and specificity is almost the same for the training and testing models. The F1 value for the testing model is 71% which is less than the 84.4% for the training model. We can conclude that the model is a moderate fit.

Setting a ‘ctrl’ model with 10-fold cross validation, we train our Decision Tree model using 5-fold cross validation with three repetitions and get these re-sampling results:

cp	Accuracy	Kappa
0.0000	0.9624289	0.9248578

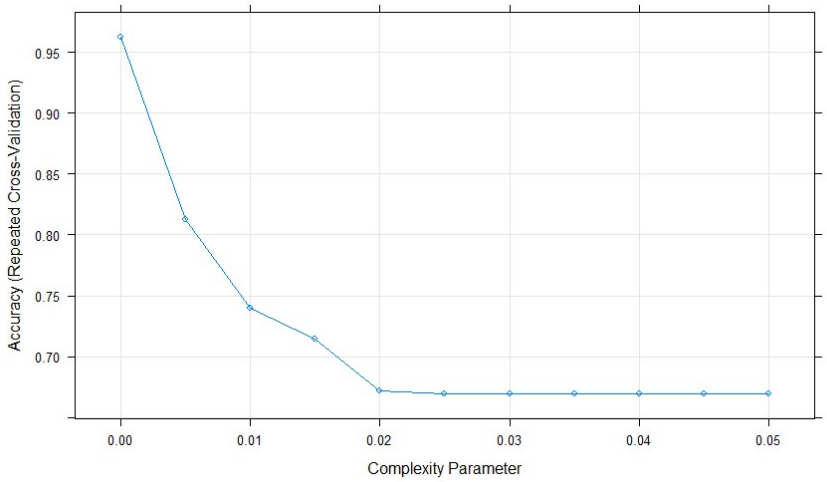

Figure 11: Decision tree fit

At a cp value of zero, we have a 96% accuracy score, and a Kappa value of 92.49% which indicates a very good model.

The most important values after tuning are:

- Customer Lifetime Value
- Total Claim Amount
- Monthly Premium Auto
- Months Since Last Claim
- Employment Status Retired
- Renew Offer Type Offer 2

Comparing the training and testing performance after tuning:

Training	**Testing**
Accuracy: 0.9798	Accuracy: 0.9347
95% CI: (0.9773, 0.9821)	95% CI: (0.9203, 0.9472)
No Information Rate: 0.5	No Information Rate: 0.8562
P-Value [Acc > NIR]: < 2.2e-16	P-Value [Acc > NIR]: < 2.2e-16
Kappa: 0.9596	Kappa: 0.7721

	Training	**Testing**
Sensitivity	0.9975809	0.9693878
Specificity	0.9620502	0.9288775
Pos Pred Value	0.9633523	0.6959707
Neg Pred Value	0.9974918	0.9944954
Precision	0.9633523	0.6959707
Recall	0.9975809	0.9693878
F1	0.9801679	0.8102345

Table 5: Comparison of training & testing performance after tuning

After tuning, the accuracy of the training model is 97.98% and that of the testing model is 77.21%. The Kappa value of the testing model is 77.21% .This indicates that the model has good accuracy, because the Kappa value adjusts accuracy by accounting for the possibility of a correct prediction by chance alone.

The precision for the testing model is 69.6% and the precision for the training model is 96.33%. The F1 value for the training model is 98.02% and the F1 value for the testing model is 81.02%.The sensitivity values for the training and testing models are 99.76% and 96.94% respectively. Hence, we can conclude that the decision tree has good performance that it classifies well. The tuned model fits well and does not show either underfitting or overfitting.

Naïve Bayes:

Naïve Bayes is a classification method which is quick to develop and use. For Naïve Bayes, we have created a data partition with "train" and "test datasets to train our model and test it for applicability. The split for the two datasets is 85:15 with 85% of the data being used for training the model. YeoJohnson has been applied to our data set to ensure proper standardization is conducted. We could not find any zero probabilities variables. Due to this, no smoothing is applied to the model data. The training data has been sampled upwards to meet the negative class. It does not hinder run time as much, and thus we can use it without an issue.

	Training	**Testing**
Accuracy	0.720	0.697
Kappa	0.441	0.260
Sensitivity	0.751	0.745
Specificity	0.690	0.689
Precision	0.708	0.286
Recall	0.751	0.745
F1-Measure	0.729	0.413

Table 6:Naive Bayes – model performance

In our model, Accuracy for Naïve Bayes is 69.7% but the Kappa Value is only at 26%. It is possible that numerous classifications are being classifying correctly due to chance, rather than our model classifying it. While we have high values for Sensitivity, Specificity, and Recall, with similar performance in both datasets, Precision values contrast highly. It is possible that the model is overfitting for the "train" dataset. Precision values is also very low, resulting in more inaccurate predictions. From the Confidence Matrix for testing data, we can see that the model incorrectly predicted the positive class in 26% of observations.

Training		
	Reference	
Prediction	No	Yes
No	4591	1655
Yes	2062	4998

Testing		
	Reference	
Prediction	No	Yes
No	808	50
Yes	365	146

Table 7: Naïve Bayes – confidence matrix

Artificial Neural Network:

The second classification model we used is the Artificial Neural Network (ANN). ANN is a machine learning model which will help make predictions through model training and testing. For ANN, we have created a data partition with "train" and "test" datasets to properly train our model and apply the learnings to the test dataset. The dataset contains the converted dummy variables and will be the basis of our partition. The split for the two datasets is 85:15 with 85% of the data being used for training the model.

For hyperparameter tuning, we have used the "grids" function. The minimum and maximum number of nodes are 3 and 19 respectively, with an increment of 4. The weight decay settings are set from 0 to 0.1 with an increment of 0.01. These settings allow to avoid overfitting. For our control object, we are using the 5-fold Cross validation method, while searching through the grid we had set up previously. The final values used for the model were size =19 and decay = 0.09.
For the ANN model, we are looking to target the "Response" variable as it our desired prediction variable. The model has been center-scaled in order to standardize the data. We are using the "nnet" method, which is used for predicting and fitting with neural networks.

	Training	**Testing**
Accuracy	0.874	0.839
Kappa	0.624	0.539
Sensitivity	1.000	0.949
Specificity	0.853	0.820

Precision	0.532	0.469
Recall	1.000	0.949
F1-Measure	0.694	0.627

Table 8: ANN results

By looking at the results from the "Testing" column from the exhibit above, the model returns a a result with high accuracy in its correct predictions, and the Kappa statistic also relatively supports that claim. While simulating the model, we find large values for Sensitivity and Specificity. This indicates that the model is accurately classifying the observations that are either positive or negative. This ensures that we can expect where most of the observations would be correctly classified. Precision, which denotes the accurate classification of predicted positive observations, is at 47%. While the rates are not exactly high, the model moderately supports correct prediction of positive values almost 50% of the time. Recall at 95% ensure that the predicted negative values are correctly classified. The results derived from Precision and Recall is also further supported by the F1-Measure of 63%. The Precision value is indeed lower that we expected. The lower value is understandable since our dataset contains a huge disparity between the "Yes" and "No" classes.

Training		
	Reference	
Prediction	No	Yes
No	5673	0
Yes	980	1112

Testing		
	Reference	
Prediction	No	Yes
No	962	10
Yes	211	186

Table 9: ANN – confidence matrix

However, we also need to consider the implications of all the performance measures together to assess our results. From *Table 9*, we can see that our model is able to correctly classify 84% of the observations in the testing dataset. Out of the total observations, only 15.4% were misclassified as "Yes". We can discern that, regardless of the low Precision value, our model still performs well in classifying our observations. It has highly accurate classifications of observations and does well in limiting the number of inaccurate classifications. It is also well-balanced as it provides similar values for both the datasets.

CONCLUSION AND RECOMMENDATIONS:

Based on our analysis, we have concluded that a few of the customer demographics and behavior are more important than others. Gender, Education Level and Income Level of customers seem to have very little effect on the outcome. However, a customer's Employment Status is very telling of their response to our campaign and the value they bring to the company. Customers who are employed or retired bring in the most value to the company but are less likely to respond to our campaign. Customers who are unemployed bring less value to the company but are more likely to respond to our campaign.

Another important determinant of response is the Renew Offer Type. From our cluster analysis, we found that the response rate for Offer 2 was higher than the response rate for offer 1. The segments that were offered more on Offer Type 1 have a lower response rate than the segments who were offered as many offers of Offer Type 2. While we do not know what aspects of the offer affect a response, we recommend that the company investigate into the offers being sent to customers. Instead of a blanket approach on all offers, it would be better to curate the offers based on customer demographics and customer behavior.

We would recommend that the company survey a small sample of each customer segment identified in our cluster analysis to be able to offer them the right deal. If our company can get more responses from clusters that have a higher average Customer Lifetime Value, it would be most beneficial. By curating our offers with regards to the customer segment, we are able to ensure that more of the customers become receptive of our services.

Other variables of importance are Marital Status and Location code. We have found, based on customer lifetime value and total claim amount, married customers are more profitable than single customers. We have also found that the customers living in Suburban areas have a higher response and are more profitable to the company.

From our decision trees model, our initial decision tree (after over-sampling) has 53 nodes. A bigger tree is more difficult to interpret .We note that the root node indicates a balanced model as the minority class and majority class have a 50/50 distribution. The variables of importance from the initial decision tree are Renew Type Offer 1, Renew Type Offer 2,Location Coded- urban, Claim amount of less than $667 ,and Retired Employment Status. After tuning our model using

5 fold cross-validation, the variables of importance that are observed are Customer Lifetime Value, Total Claim Amount, Income, Monthly Premium Auto, Months Since Last Claim, Employment Status Retired, Renew Offer Type Offer 2 and Education Level.

It is important that the company analyzes the Customer Lifetime Value for the different customer segments and compile a list of the higher customer value from the Total Claim Amount, Income, Monthly Premium Auto, Months Since Last Claim, Employment Status Retired, Renew Offer Type Offer 2 and Education Level, and ensure that there is targeted marketing campaign to those customers that have a higher value. These customers could also have certain promotions and incentives to keep them responding to campaigns and interested in purchasing insurance and avoid them churning out of the business.

We employed four tools to assist in our decision making (Cluster Analysis, Decision Trees, Naive Bayes and Artificial Neural Networks). Among those four tools, ANN has provided us with the best outcome. The model is able to correctly classify 84% of the observations in the testing dataset. It is highly accurate in classifying the observations and performs well in minimizing the number of inaccurate classifications as well. The model returns an inaccurate classification of 15% of the observations but considering our correct classifications, the model is still very relevant. By correctly classifying a large portion of the dataset, we will be able to minimize the portion that is incorrectly classified, and thus minimize the cost incurred due to inaccurate correction. If we combine the insights derived from Cluster Analysis, we can specifically target a customer base and make predictions for that group. If we are able to improve the number of positive responses received, our model can deliver better performance and increase its predictive power, and in turn, minimize our incurred costs even more.