

FACULTY OF ARTS AND SOCIAL SCIENCES



INDIVIDUAL COURSEWORK COVERSHEET

Coursework Details			
Module Name and Code (please check front of module handbook)	MANM354 MACHINE LEARNING AND VISUALISATIONS		
Coursework Title	Customer Retention by Reducing Customer Churning		
Date of Submission	11 th June, 2018	Word Count	2997

Student Details			
Student URN (7-digit number on Uni card)	6476333	Student Name	VEDANT RUPARELIA
Programme	MACHINE LEARNING AND VISUALISATIONS		

Student Declaration <i>To be agreed by Student</i>
<p>Please refer to the University of Surrey Regulations for the Conduct of Examinations and Other Forms of Assessments and your departmental Student Programme Handbook for more information on Academic Misconduct and Plagiarism.</p> <p>Declaration: <i>I confirm that the submitted work is my own work and that I have clearly identified and fully acknowledged all material that is entitled to be attributed to others (whether published or unpublished) using the referencing system set out in the programme handbook. I agree that the University may submit my work to means of checking this, such as the plagiarism detection service Turnitin® UK. I confirm that I understand that assessed work that has been shown to have been plagiarised will be penalised.</i></p> <p><i>'In completing this work, I have been assisted with its presentation by [state name and contact details of assistant] and, if requested, I agreed to submit the draft material that was completed solely by me prior to its presentational improvement.'</i> (This paragraph is to be deleted where it is not relevant).</p> <p>By completing and submitting this form, I confirm that:</p> <ul style="list-style-type: none"> • I have read and fully understand the University's Regulations and guidance on Academic Misconduct and Plagiarism • This submission is my own work and all coursework is the product of my own research and study. • All quotes and sources have been fully and properly attributed and referenced. • This work has not been previously submitted, in full or in part, for assessment at this or any other institution • No effort has been made to subvert plagiarism detection processes of the University • This submission may be transferred to and stored in the Turnitin Plagiarism Detection database for plagiarism detection now and in the future • I understand that all required work must be received within the published deadline • I understand that work received after the published deadline will be penalised in line with University Regulations • I understand that any request for mitigating circumstances must be made formally via the Student Liaison Officer, using the appropriate form and including evidence; the application and associated evidence must be received by the stipulated date • I have not lent my work to any other student. • I have not colluded with any other student in the production of this work.

- The word count I have declared is precise.

Table of Contents

Introduction	4
Dataset	4
Technical Perspective.....	5
Some Prerequisites	5
Pre-processing.....	5
Deep Learning Model using Keras	8
Logistic Regression Modelling.....	10
Business Perspective	11
Business and Churn.....	11
Project Life Cycle.....	11
Visualisations	13
Algorithms.....	15
Scoring Methods and Interpretations.....	15
Customer Lifetime Value (CLV)	18
Conclusions and Recommendations	22
Appendix 1	25

Customer Churn Analysis

Introduction

Customer churn arises when there is a termination in association between a customer and some company/business for whatsoever reason and thus costs the company a great amount. Customers are the base that empowers any business. Even if a single customer leaves the system, company suffers with losses. Further, as understood from previous researchers to acquire new customer in comparison to keeping the existing one is 5-6 times more expensive. As a result, businesses have started taking steps to reduce customer churn. One industry that focusses on churn rates with great importance is the telecommunications sector, in which will be discussed in this report. In regards with some previous studies in this field, it can be interpreted that Machine Learning can help and proves to be a boon in such situations. We use Deep Learning with Keras for this analysis in place of the traditional machine learning algorithms.

Dataset

The dataset that we will consider is the large telecommunications carrier company which wants to improve the profitability by reducing the churning of the existing customers because acquiring new ones in regards with existing costs them more and reduces the profit. Predicting churn is important issue but due to what factors it occurs is the most critical aspect for business that provides services that are contract based.

Technical Perspective

Some Prerequisites

- We use some R libraries for carrying out this study which are installed using the `install.packages()` function. The use of each is explained further in the report. These libraries can then be loaded by using the `library()` function.
- As we are going to use Keras package, we will need to first install (if not already installed) this package by using the `install_keras()` function.
- Once this is done we import the dataset using the `read.csv()` function.

Pre-processing

Real world data is messy and hence needs sufficient pre-processing to reach a stage where modelling algorithms can be applied. Preprocessing is carried out in following steps:

1) Data Pruning

Pruning is basically the process of deleting unnecessary rows or columns which are not required as part of our analysis. "CustomerID" is deleted as we do not require to model the unique identifier. 11 missing values in "Total charges" are dealt with by using the `drop.na()` function from `tidyr` as we would still be left with 99.8% of the records. A `Prune_Function()` is created in R for solving this purpose.

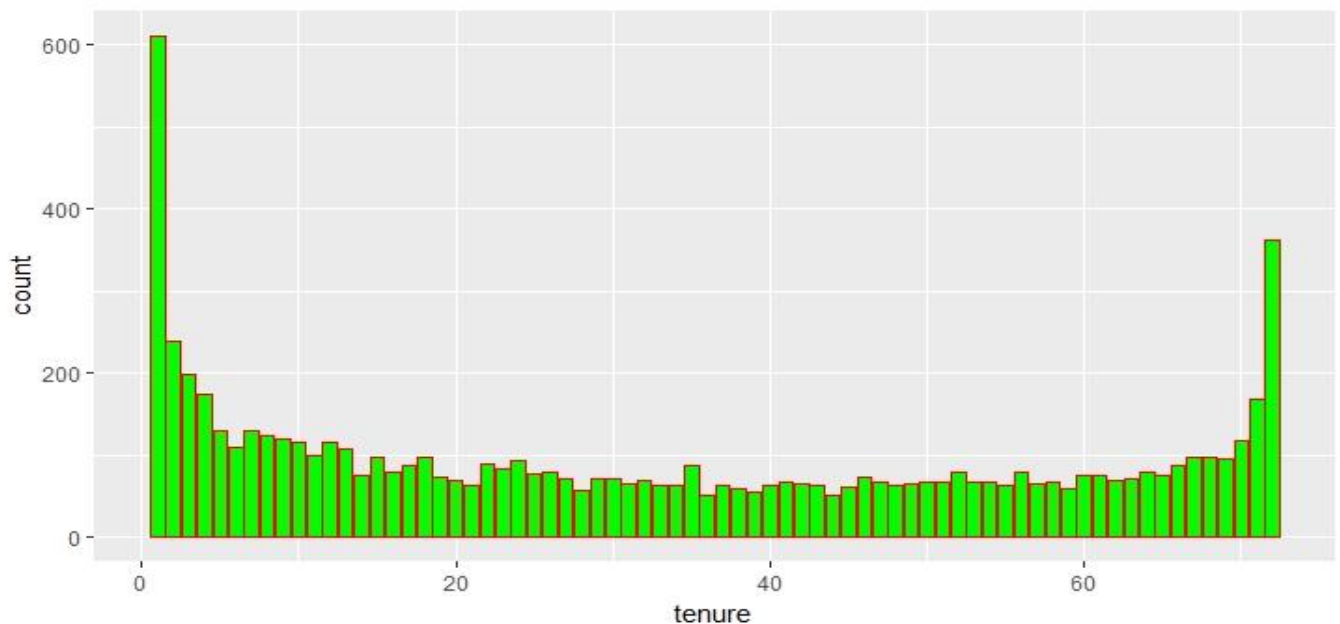
2) Data Splitting

Splitting the data into train and test is done using the `initial_split()` function from `rsample` library. This package provides useful sampling of data and returns a `rsplit` object. Ratio of 70:30 is used as it provides better accuracy in comparison others.

3) Data Transformation/ Exploratory Factor Analysis

Artificial Neural Networks (ANN) work best when they are scaled, encoded using one hot and centered. Additionally, some other transformations can also be used to find associations between the fields which would help in modelling.

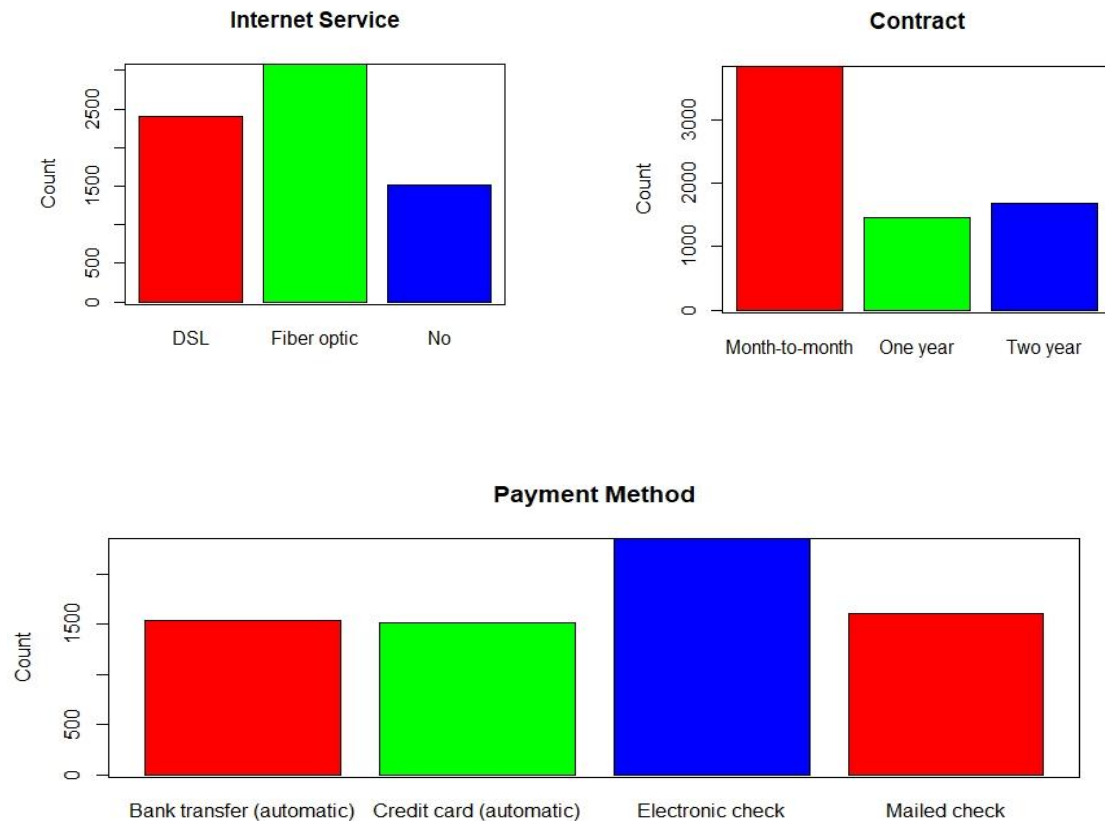
- “Tenure” is divided into 6 groups of 1 year each by creating a TenureGroup() function as it is a numeric feature and works well when discretised into groups.



- The “Total charges” field data is not normal and hence to deal with it we need to discretise the data for normal distribution. There are different ways but we use Log transformation here which is done using the log() function. The Log and Total Charges are tested for correlation for surety. This is done using the correlate(), focus() and fashion() functions from tidyr basically used for formatting. Transformation will increase the accuracy of ANN considering the magnitude between churn and log value.

	rowname	Churn
1	TotalCharges	-.20
2	LogTotalCharges	-.24

- One Hot Encoding technique is used to convert the categorical fields to sparse data which is necessary while modelling ANNs. These train with speed and accuracy when one hot encoded, scaled, normalised and centered. We have 3 multiple category field which needs to be encoded:



- Finally, we include all of these into a `recipe()` function developed by Max Kuhn. It takes an “`recipe_object`” argument like all models take and the transformations are added using `step()` functions like `step_discretize()`, `step_log()`, `step_dummy()`, `step_center()`, `step_scale()`. The last step is to use `prep()` function for passing the dataset. After this we apply the recipe to the `bake()` function for transforming the data which can be understood for modelling. Finally, the target vectors are created.

Data Recipe

Inputs:

```

role #variables
outcome 1
predictor 19

```

Training data contained 5243 data points and no missing data.

Operations:

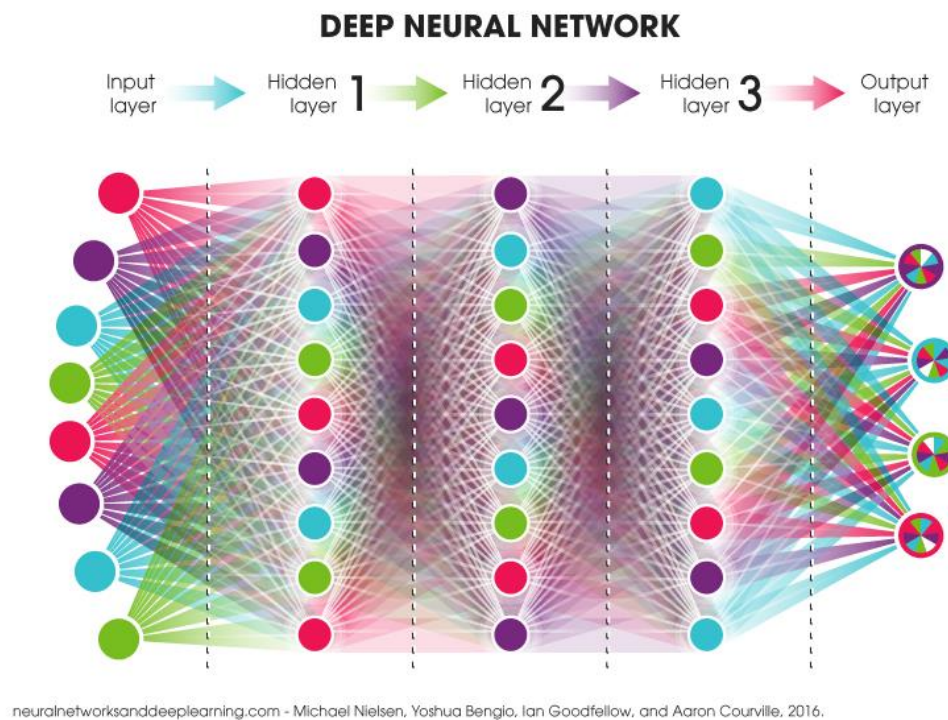
```

Dummy variables from tenure [trained]
Log transformation on TotalCharges [trained]
Dummy variables from gender, Partner, Dependents, tenure, PhoneService, MultipleLines, ... [trained]
Centering for SeniorCitizen, MonthlyCharges, TotalCharges, gender_Male, ... [trained]
Scaling for SeniorCitizen, MonthlyCharges, TotalCharges, gender_Male, ... [trained]

```

Deep Learning Model using Keras

The basis is to train a multi-layer neural network, a deep neural network known as multi-layer perceptron. The logic behind this is that it gives more accurate and precise predictions. There are several packages available but keras package is used because it has straightforward implementation, it is fast and easy to implement.



Steps for modelling neural networks:

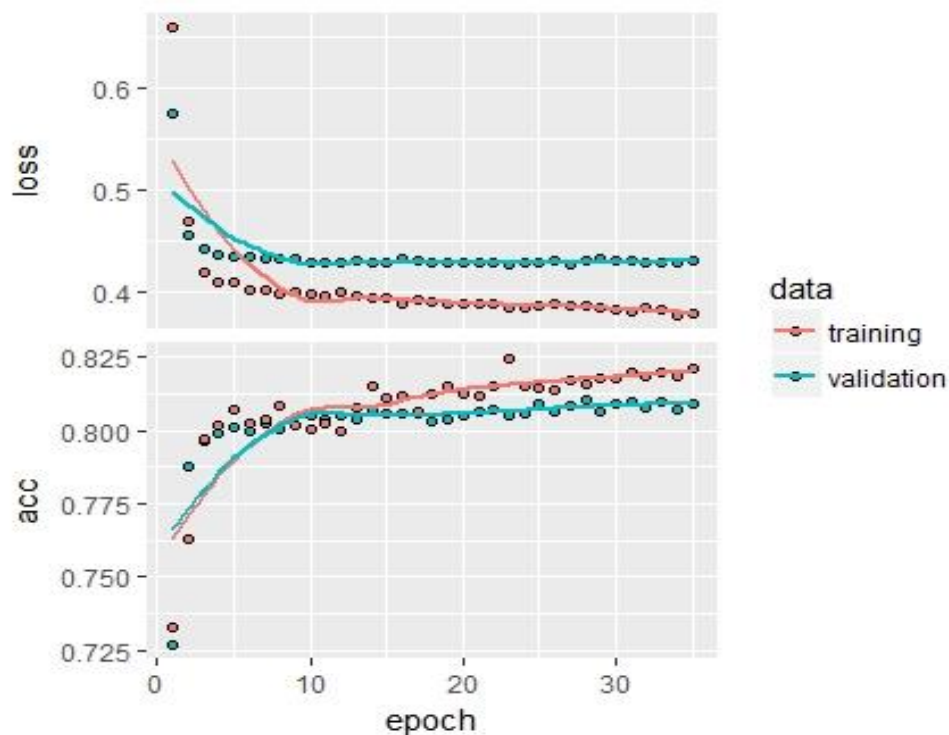
- 1) By using `keras_model_sequential()` function, we initialise the keras model that we are going to build. It contains a layer stack.
- 2) We then apply layers to this model starting with the input layer where we supply our data. Next, we create 2 hidden layers using `layer_dense()` function. These provide weights to non-linear activation. Parameters are selected as `kernel_initializer="uniform"` and `activation = "relu"`. These are basic optimisation parameters. Dropout layers are created next using `layer_dropout()` function to overcome over-fitting by 10% using the rate parameter. Lastly, we parameterise the output layer with basic binary parameters as `kernel_initializer = "uniform"` and the `activation = "sigmoid"`.

- 3) Last step involves compiling the model using compile() function and again using basic and popular binary parameters as optimizer = “adam”, loss = “binary_crossentropy” and metrics = c(“accuracy”) of checking the accuracy.

The final model looks like this:

Model		
Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 16)	576
dropout_1 (Dropout)	(None, 16)	0
dense_2 (Dense)	(None, 16)	272
dropout_2 (Dropout)	(None, 16)	0
dense_3 (Dense)	(None, 1)	17
Total params: 865		
Trainable params: 865		
Non-trainable params: 0		

fit() function is used to run neural network to check the accuracy between training and validation which should be minimal. Parameters like batch_size =50 which usually should be high which shows number of samples per gradient, epoch = 35 which should also be high as training cycles are controlled by this and validation_split = 0.30 to overcome overfitting by using 30% validation data. We can plot history using the plot() function:



We visualise the accuracy of validation and loss levelling off. The curve has started to flatten which is an indication that we should stop training with the following results:

```
Trained on 3,425 samples, validated on 1,468 samples (batch_size=50, epochs=35)
Final epoch (plot to see history):
val_loss: 0.4392
val_acc: 0.8065
loss: 0.3894
acc: 0.8216
```

Logistic Regression Modelling

The LR model was developed which provided a similar accuracy to ANN. As the target variable is categorical, logistic regression is suitable.

```
glm(formula = Churn ~ ., family = binomial(link = "logit"), data = train_data_reg)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.9584	-0.6910	-0.2949	0.6973	3.0468

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.6280550	0.9689229	0.648	0.516856
genderMale	-0.0191201	0.0772742	-0.247	0.804574
SeniorCitizenYes	0.2099271	0.0989888	2.121	0.033946 *
PartnerYes	-0.0950171	0.0919388	-1.033	0.301378
DependentsYes	-0.0618555	0.1066691	-0.580	0.561994
PhoneServiceYes	0.2003618	0.7778355	0.258	0.796724
MultipleLinesYes	0.3839899	0.2119708	1.812	0.070060 .
InternetServiceFiber optic	1.6681775	0.9592626	1.739	0.082031 .
InternetServiceNo	-1.7460257	0.9705067	-1.799	0.072005 .
OnlineSecurityYes	-0.1481216	0.2124403	-0.697	0.485653
OnlineBackupYes	0.0411673	0.2108909	0.195	0.845231
DeviceProtectionYes	0.0720854	0.2093465	0.344	0.730594
TechSupportYes	-0.2310916	0.2167215	-1.066	0.286285
StreamingTVYes	0.5655001	0.3933895	1.438	0.150574
StreamingMoviesYes	0.5544487	0.3913909	1.417	0.156597
ContractOne year	-0.7647554	0.1282191	-5.964	2.45e-09 ***
ContractTwo year	-1.6460301	0.2173680	-7.573	3.66e-14 ***
PaperlessBillingYes	0.3399827	0.0886443	3.835	0.000125 ***
PaymentMethodCredit card (automatic)	-0.0588562	0.1352118	-0.435	0.663353
PaymentMethodElectronic check	0.3101323	0.1128098	2.749	0.005975 **
PaymentMethodMailed check	0.0006885	0.1374428	0.005	0.996003
MonthlyCharges	-0.0325022	0.0381119	-0.853	0.393765
tenure_group1-2 Years	-0.8339022	0.1141952	-7.302	2.83e-13 ***
tenure_group2-4 Years	-1.2557166	0.1193873	-10.518	< 2e-16 ***
tenure_group4-5 Years	-1.4207402	0.1673031	-8.492	< 2e-16 ***
tenure_group5+ Years	-1.6338884	0.2041315	-8.004	1.20e-15 ***

Tenure groups, Paperless Billing, Contract, were the variables that were found to be statistically significant. These can be used as selected features for other models as well.

(Appendix A for more info)

For calculating the estimates `Estimations_Function()` is the own function created in R.

Business Perspective

Business and Churn

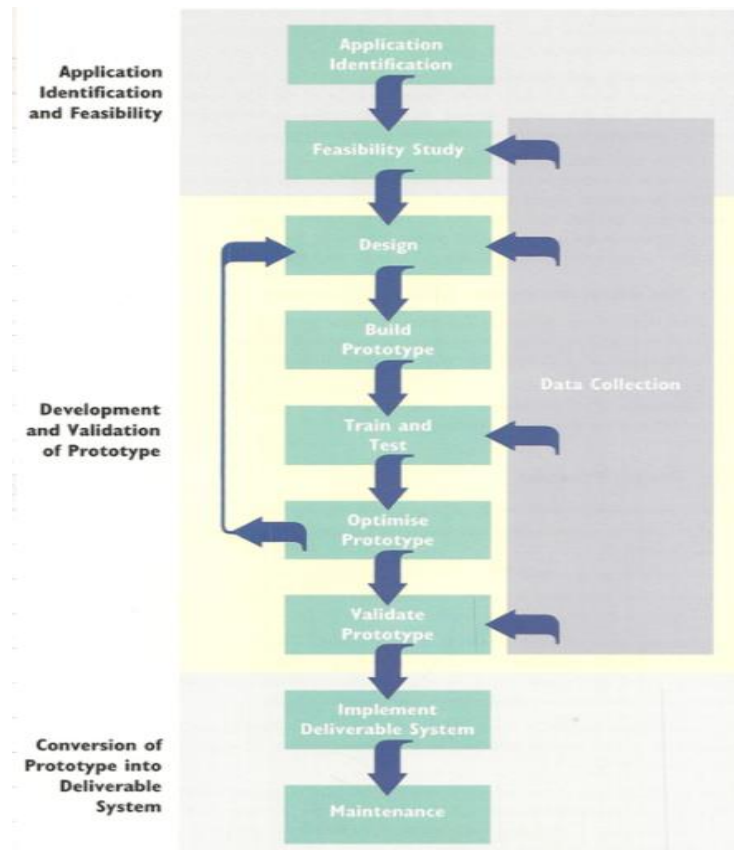
A business that handles the churn rate in an organised way is a sign of a balanced and mature business. Telecommunication industry is the best example of this where customers change network providers quite frequently due to the competitiveness of the industry. Most of the cases of churn happen here because significant knowledge about churn drivers have been generated.

There are not as such strategies which are in use that can help these companies to reduce churn. Machine learning has great importance in this field and past research has proved that by using modelling on the large datasets of these industries, analysis and trends can be determined on predicting and reducing churn.

Another aspect of this is vast number of customers means huge turnover and if customers start to churn this would impact the profit of the company and in turn the performance. Hence, it is of utmost importance to reduce customer churning in this sector, because gathering new customers rather than retention of existing results a more expensive cost approach.

Project Life Cycle

As this study has a greater organisational impact, we have used the project life cycle to implement the entire process and analysis.



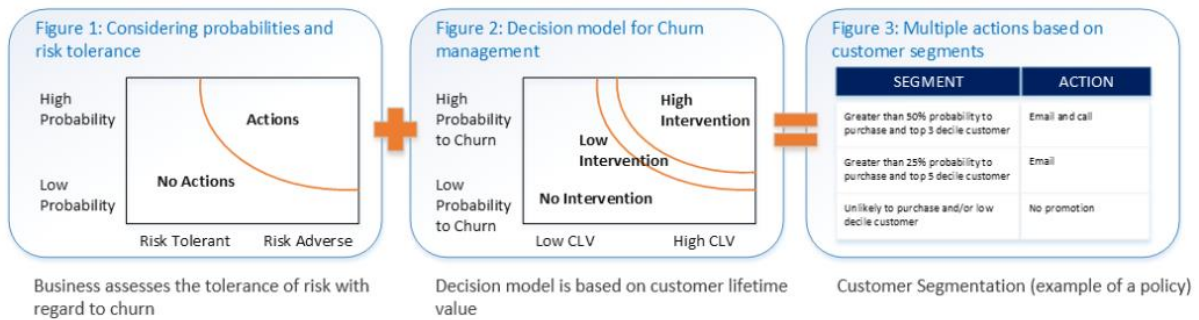
This covers the entire process of the project starting from gathering ideas and implementing it to validating it and bringing it to reality to make it work in real-time.

A basic methodology that can be implemented to solve the problem of churn which can be depicted and explained as below.

Stage 1: Modelling risk to account for factors affecting risk and profit

Stage 2: A Decision Model that helps to take into account the affects on customer lifetime value (CLV) and the probablity of customers leaving the system.

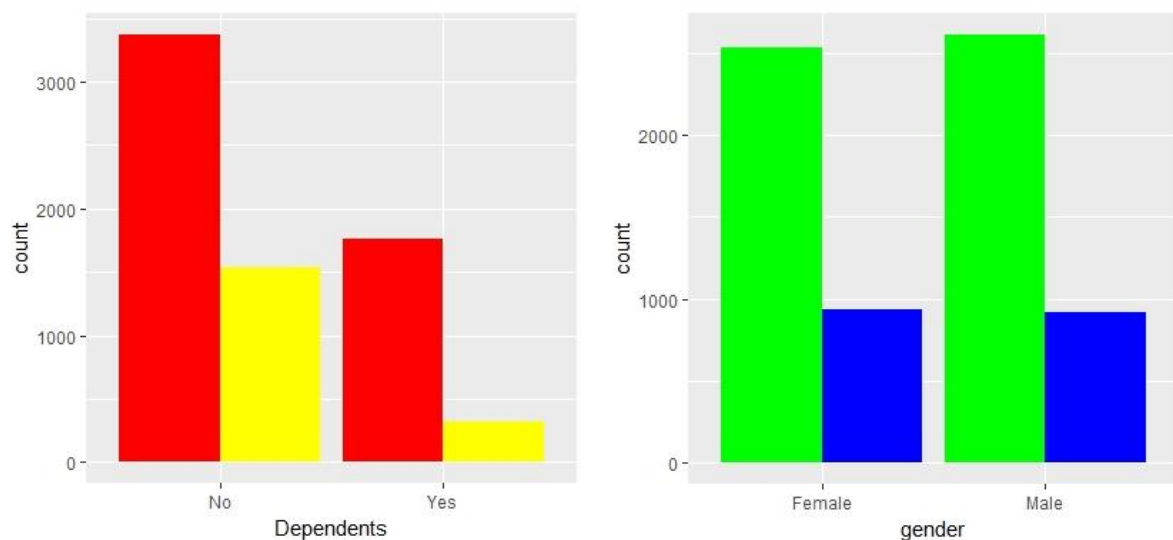
Stage 3: A futher qualitative analysis that allows for actions that helps deliver best performance with regards to customers and company.

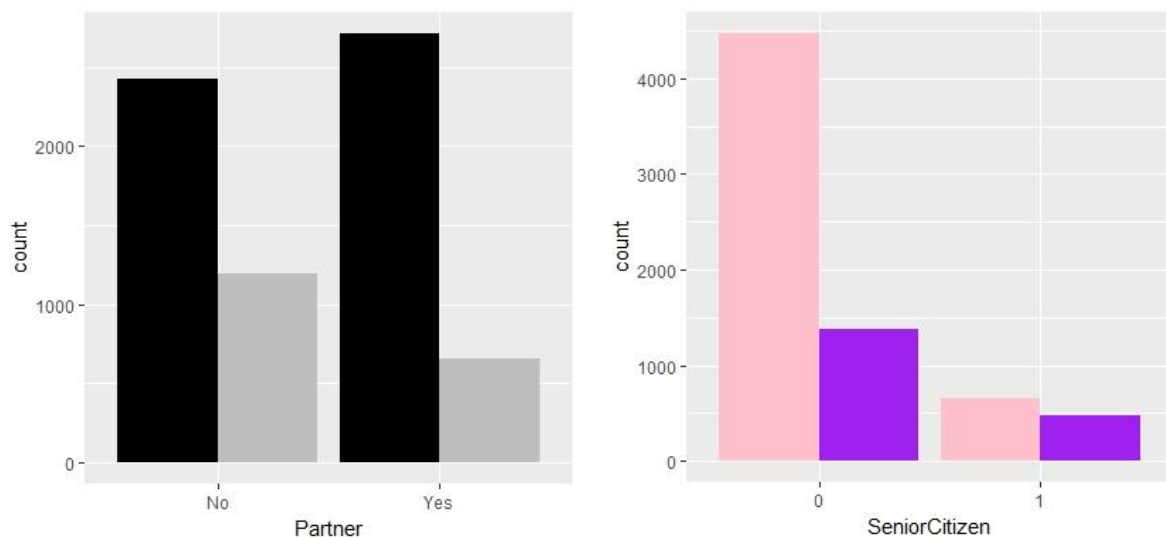


A systematic interaction between the models is important if we were to deliver a comprehensive approach of reduction in churn. The complete cycle that the model follows is a generalised structure, applicable to many business problems. Machine Learning and Artificial Intelligence implementation is ideal for these problems, as they collect exhaustive noisy data about the customers and considering the growth of technologies like Deep Learning, Story Telling and many more, businesses will flourish employing multiple layers of smart solutions that enhance customer satisfaction and performance of the industry.

Visualisations

We will first start with some visualisations which can be used as a part of management decisions. Some basic visualisations at the initial stage always help in deciding strategies while modelling ahead.





Conclusions from these visualisations indicate that

- 1) Customers with no dependents are likely to churn more than the ones that have more dependents.

It can also be found out that the cost of no dependents churning is affecting the company the highest and this needs to be critically figured out that why are the customers with no dependents churning more and what can be done to retain them. Cost and number of customers that churn can be shown below.

CHURNING WITH REGARDS TO NO DEPENDENTS

TOTAL NUMBER OF CUSTOMERS	1532
TOTAL COST	2227347

- 2) There is not as such discrimination in gender of the people that churn.
- 3) Customers with no partners churn more than with partners.

It is depicted that customers with no partners churn more and this affects the business as well but comparatively its less than no dependents and some strategy needs to be

implemented while modelling so that we can reduce churning and enhance retention. The table shows the statistics

CHURNING WITH REGARDS TO NO DEPENDENTS

TOTAL NUMBER OF CUSTOMERS	1196
TOTAL COST	1295126

It would be interesting to check that how many customers are there with no dependents and no partners that churn, considering their correlation with literature it should be high.

- 4) The percentage of Senior Citizens churning is quite high, 471 of them churned.

Algorithms used in this prototype are:

- 1) Deep Learning with Keras
- 2) Logistic Regression

Scoring Methods and Interpretations

	PRECISION	RECALL	ACCURACY	PREDICTION CONFUSION MATRIX		
DEEP LEARNING	0.658	0.528	80%	Truth		
				Prediction	no	yes
				no	1354	275
				yes	160	308
LOGISTIC REGRESSION	-	-	80%	FALSE	TRUE	
				0	944	82
				1	198	173

These models are evaluated based on the Confusion Matrix, Recall, Precision and Accuracy. These are the performance indicators which in general test the performance of the models and the levels of accuracy the models predict. When implementing machine learning, the aim is not to develop a perfect model but to evaluate the probability of correct prediction generation.

Precision

It indicates that how often the model has predicted as “yes” and it is yes.

$$precision = \frac{true\ positives}{true\ positives + false\ positives}$$

Recall

It indicates that how often the predicted “yes” is correct.

$$recall = \frac{true\ positives}{true\ positives + false\ negatives}$$

Precision and Recall parameters are critical for business cases. The main aim of the business is to balance the costs of customers that want to leave and targeting them for retention, instead of targeting the whole set which will include loyal customers.

In our case, precision and recall are quite good which should be above 0.5 ideally and are actually above.

Accuracy

Finally, the accuracy gives the proportion of correct cases determined by the model.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

The accuracy of all 3 models is approximately around 80% which means that the models are correctly predicted without any biases and no perfect model is found amongst all 3.

Confusion matrix

A confusion matrix determines results in a tabular form by combining the actual vs predicted outputs of a machine learning model. It basically consists of 4 quadrants which can be evaluated as following in our case:

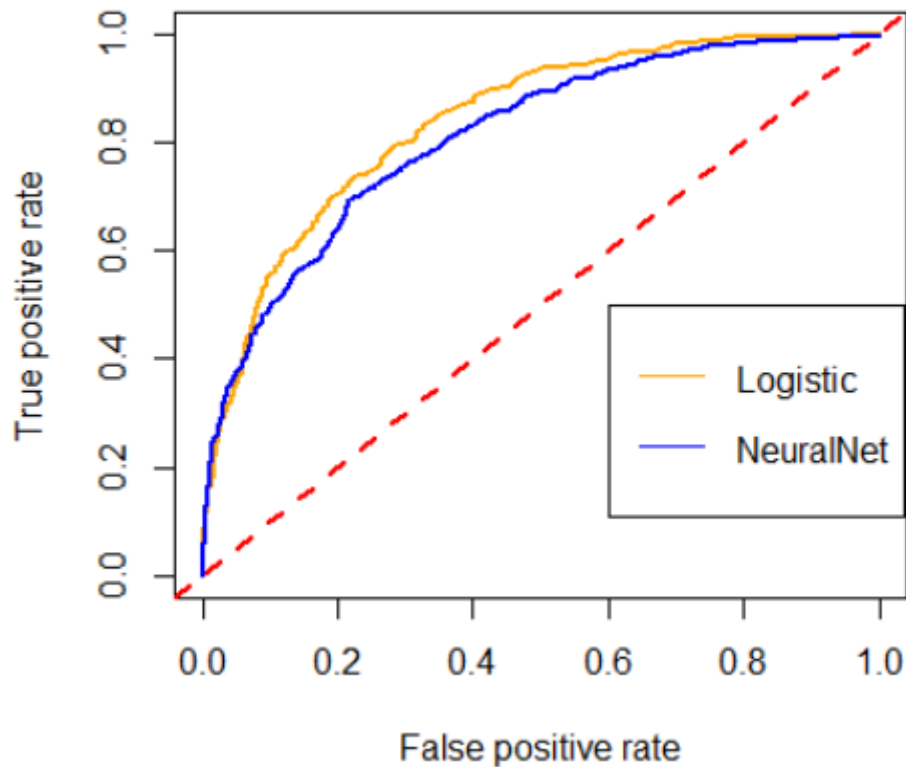
TN – we entice all these subscribers by spending 10% of their 12-month revenue and assume we then retain all of them. So, we gain 90% of their 12-month revenue.

FP – we miss enticing these subscribers and so they leave. Cost of this customer loss is equal to their 12-month revenue to the business, plus we need to replace them, which is \$750 each expenditure.

FN - we wrongly entice these subscribers by spending 10% of their 12-month revenue which is a loss to the business.

TP – describes the number that model correctly predicts the churn customers.

Another way of identifying the models estimate is by using the ROC curve. The area under the curve provides the estimation of the model and values being more than 0.5 shows that the models performed well.



Customer Lifetime Value (CLV)

It is the future profits that can be generated by a customer by considering the discounted values. The most important values for the business are the cost and revenue estimates. Forecasting revenues is difficult and hence companies tend to be more critical of this. It is also the current value of the future cash flows attributed by the company to the customer during the entire association of the customer with the business. Machine Learning Models are required to predict these features which would help the business know where they stand.

The formula that can be implemented for calculating the CLV is:

$$CLV = \sum_{t=1}^n (r)^t \frac{P_t}{(1+d)^t}$$

where,

- tt is a time period, e.g. the first year($tt=1$), the second year($tt=2$)
- nn is the total number of periods the customer will stay before he/she finally churns
- rr is the retention rate/possibility
- $pipi$ is the profit the customer will contribute in the Period t
- dd is the discount rate

The CLV is a problem until we optimise it using threshold values and then select the data according to the best selected threshold based on performance. The thresholds used are:

THRESHOLD	TP	TN	FP	FN	ACCURACY
0.1	350	514	512	21	62%
0.2	324	671	355	47	71%
0.3	273	806	229	98	78%
0.4	231	868	158	140	79%
0.5	187	928	98	184	80%
0.6	139	971	55	232	79%
0.7	86	1001	25	285	78%
0.8	29	1002	4	342	75%
0.9	0	1026	0	371	73%
1.0	0	1026	0	371	73%

We select the threshold with 0.3 as it provides the best predictions in terms of the Uplift, Profit and ROI.

These parameters are calculated as follows:

It is the TotalCharges per subscriber record that we will need to use to make the calculation depending upon where the classification falls in the quadrants TP, FP, TN or FN. The

threshold is set from the model which gives FPR 35% and TPR is 85% and we use 20% of the dataset as test data to generate.

The mean of TotalCharges field is selected so that it's a more genuine approach in working through the calculations considering the field does not follow normal distribution and to avoid any biases.

UPLIFT		FPR=27%
		TPR=79%
NET REVENUE FROM SUBSCRIBERS WE RETAIN BY SPENDING THE 10%.	(TP*90%*TotalCharges)	\$560,933
WRONGLY ENTICED SUBSCRIBERS	(-FP*10%*TotalCharges)	-\$50226
LOST REVENUE DUE TO MISSED SUBSCRIBERS THAT WE DID NOT ENTICE AND SO LOST	(-FN*TotalCharges)	-\$223734
COST TO REPLACE THE LOST SUBSCRIBERS	(-FN*\$750)	-\$73500
	Sum for Total uplift from using the model on the test dataset	\$213473

We can consider presenting this as a comparison to “if no model existed”. The cost to the business would have been all those who churned in the dataset, $P=TP+FN$. If we assume that the same subscriber revenue lost due to churn is exactly replaced when these are substituted with new subscribers by spending \$750 to acquire them, then we simply get:

$$P = TP + FN$$

$$P = 371$$

$nomodel = (P * \$750)$, where P is the number of subscribers that churned

$$= 371 * 750 = \$278250$$

In this case *nomodel* is \$278250.

The additional 'profit' to the business when the model is in place is: *nomodel*-uplift, so \$64777

Return on Investment (ROI)

Alternatively, the business might examine being the gain from investment divided by the investment. Using the above, the investment is:

		FPR=27%, TPR=79%
SUBSCRIBERS WE RETAIN BY SPENDING THE 10%.	(TP*10%*TotalCharges)	\$62326
WRONGLY ENTICED SUBSCRIBERS	(-FP*10%*TotalCharges)	-\$50226
COST TO REPLACE THE LOST SUBSCRIBERS	(-FN*\$750)	-\$73500
	Invested based on the model churn predictors	\$186052

The model gained the revenue from (subscribers we retain- subscribers we lost)

$$= (\$560,933 - \$223734) = \$337199.$$

The amount invested by the business=\$186052.

$$ROI = \frac{\text{investment}}{\text{gain from investment}}$$

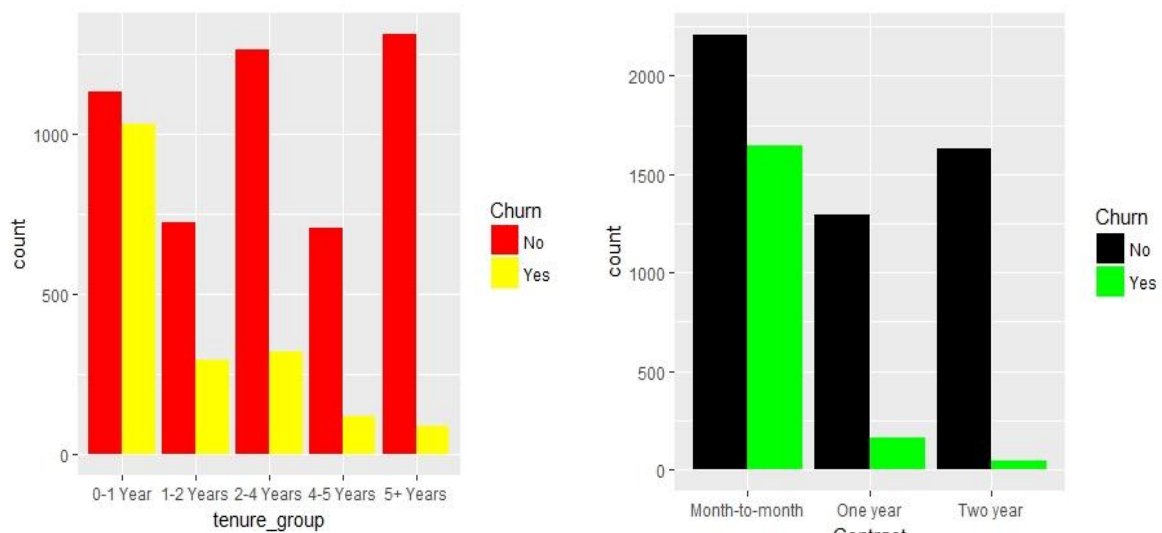
$$= 181 \%$$

All these figures state that our model fits well.

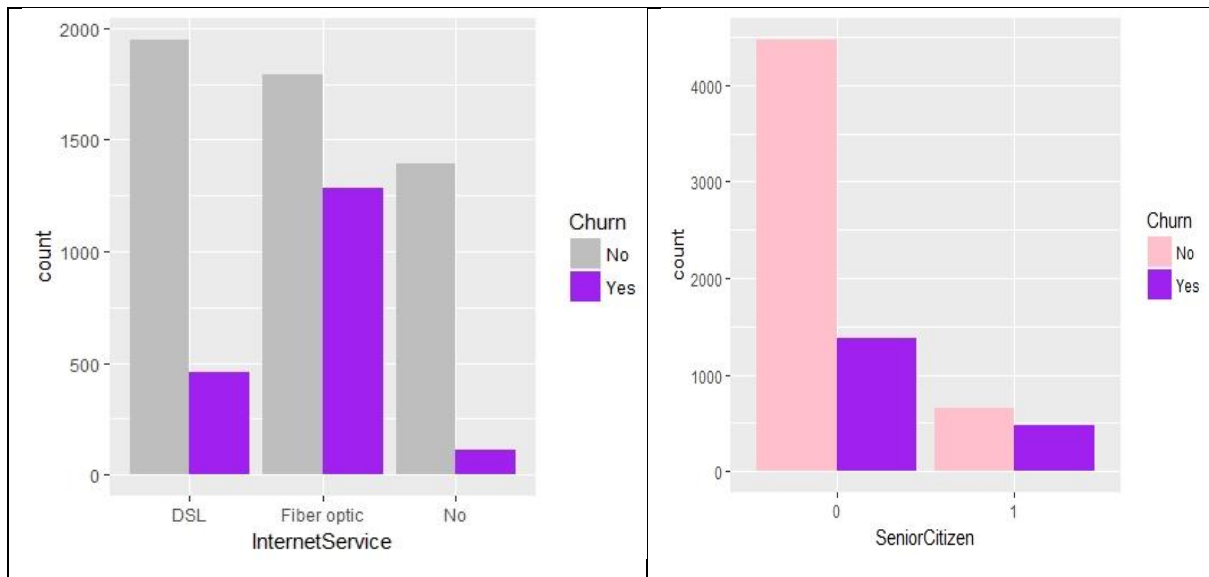
Conclusions and Recommendations

Feature Investigation based on highly significant variables suggested the following conclusions based on which recommendations can be provided:

- 1) Customers with Tenure group 0-1 year are more likely to churn. The trend shows that higher the tenure group lower the churn rate and hence it is recommended to attract customers for higher tenure.

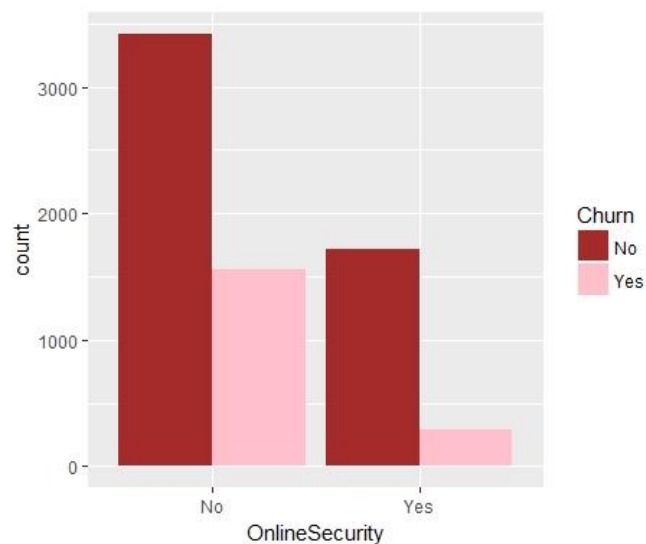


- 2) Customers with one and two-year contracts are less likely to leave and hence it is recommended to attract customers for a larger contract period.
- 3) Customers that use fibre optic cable service are more likely to leave and hence it is recommended to attract customers with other types.

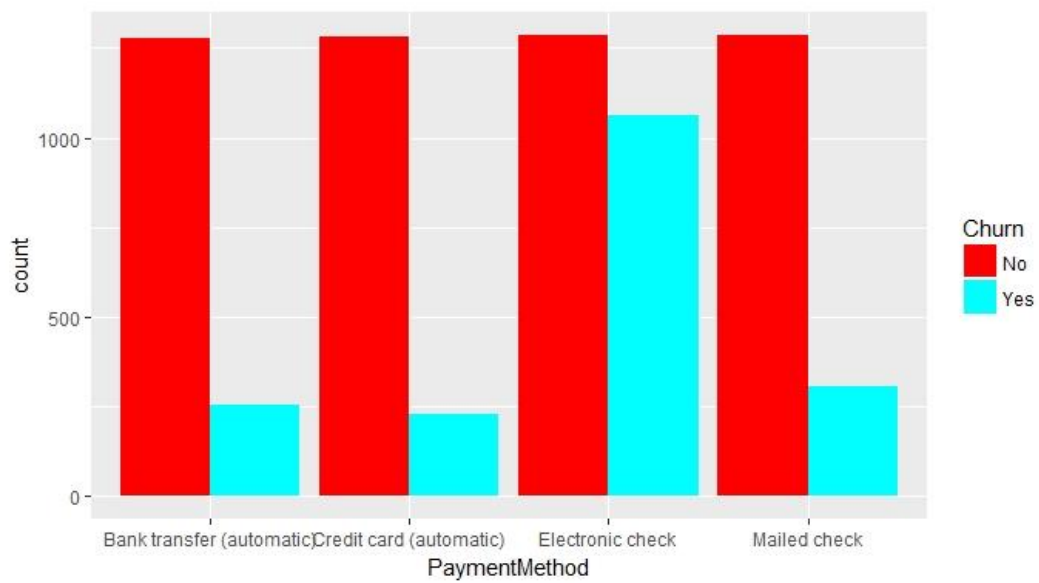


4) It is suggested to attract non-senior citizens because these are more likely to churn.

5) Customers without online security are more likely to churn and hence it should be taken care of providing online security to avoid churning.



6) Customers who have opted for Electronic Check as their payment method are more likely to churn and hence this method should be avoided in future.



Customer churning is a very expensive issue. The good news is that the models which we built are sufficiently accurate and can be applied to this business area to reduce customer churning. Moreover, the revenue generated is genuine and the analysis suggest that the above features needs to be investigated and taken care of.

Appendix 1

Preprocessing in Logistic Regression (LR)

LR only accepts binary values and hence SeniorCitizen field is converted to binary values. Missing Values are deleted using `complete.cases()` function. Tenure is converted to 6 groups for better analysing that which groups churns more because the data is widely distributed. `Tenure_group()` function is created for the same. `customerID`, `tenure` and `MonthlyCharges` fields are converted to null that is they are removed as they are unique identifiers.

Data Partition is done in the ratio 80:20 using the `data_partition()` function.

Modelling in LR

`Glm()` function is used to model LR. We provide `family = binomial` as we are predicting a categorical target variable. `Summary()` is used to plot model summary.

`Anova()` function is used to analyse the deviance when adding variables one by one.

The results of Anova are:

Analysis of Deviance Table

Model: binomial, link: logit

Response: Churn

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			5592	6473.8		
gender	1	0.36	5591	6473.4	0.547765	
SeniorCitizen	1	116.89	5590	6356.6	< 2.2e-16	***
Partner	1	135.21	5589	6221.3	< 2.2e-16	***
Dependents	1	45.59	5588	6175.8	1.461e-11	***
PhoneService	1	0.72	5587	6175.0	0.395009	
MultipleLines	1	6.78	5586	6168.3	0.009201	**
InternetService	2	543.05	5584	5625.2	< 2.2e-16	***
OnlineSecurity	1	182.39	5583	5442.8	< 2.2e-16	***
OnlineBackup	1	88.94	5582	5353.9	< 2.2e-16	***
DeviceProtection	1	51.19	5581	5302.7	8.396e-13	***
TechSupport	1	91.72	5580	5211.0	< 2.2e-16	***
StreamingTV	1	0.03	5579	5210.9	0.861604	
StreamingMovies	1	0.95	5578	5210.0	0.329213	
Contract	2	305.99	5576	4904.0	< 2.2e-16	***
PaperlessBilling	1	15.29	5575	4888.7	9.212e-05	***
PaymentMethod	3	46.03	5572	4842.7	5.578e-10	***
TotalCharges	1	107.08	5571	4735.6	< 2.2e-16	***
tenure_group	4	85.82	5567	4649.8	< 2.2e-16	***

Signif. codes:	0	'***'	0.001	'**'	0.01	'*' 0.05 '.' 0.1 ' ' 1

Odds ratio provides the odds of an event happening and is very important with regards to LR. We use

exp(cbind(OR=coef(LReg_Model), confint(LReg_Model)))

to plot the odds ratio using MASS library.

	OR	2.5 %	97.5 %
(Intercept)	0.6964252	0.4905521	0.9864409
genderMale	0.9782227	0.8480006	1.1284753
SeniorCitizenYes	1.2368570	1.0297357	1.4851142
PartnerYes	1.0135476	0.8539289	1.2033147
DependentsYes	0.8096408	0.6628720	0.9875051
PhoneServiceYes	0.6765038	0.5087161	0.9006867
MultipleLinesYes	1.3125656	1.0998312	1.5673146
InternetServiceFiber optic	2.5964974	2.0956215	3.2235249
InternetServiceNo	0.3953830	0.2900072	0.5359759
OnlineSecurityYes	0.7302524	0.6046312	0.8808818
OnlineBackupYes	0.8580586	0.7232199	1.0181490
DeviceProtectionYes	1.0046409	0.8434204	1.1970791
TechSupportYes	0.7290901	0.6026963	0.8809924
StreamingTVYes	1.2296351	1.0267063	1.4733920
StreamingMoviesYes	1.3708107	1.1451272	1.6422234
ContractOne year	0.5247067	0.4142904	0.6616534
ContractTwo year	0.1945396	0.1278709	0.2888182
PaperlessBillingYes	1.3966349	1.1853094	1.6467021
PaymentMethodCredit card (automatic)	1.0037573	0.7797562	1.2918977
PaymentMethodElectronic check	1.4992932	1.2170747	1.8505448
PaymentMethodMailed check	1.1084551	0.8591902	1.4316435
TotalCharges	0.9999073	0.9997835	1.0000328
tenure_group1-2 Years	0.4101738	0.3259060	0.5151097
tenure_group2-4 Years	0.3124103	0.2247521	0.4315592
tenure_group4-5 Years	0.3065506	0.1784963	0.5200672
tenure_group5+ Years	0.2702459	0.1326064	0.5400655

We then use predict function to predicted the actual vs predicted outputs using predict() function and plotting the confusion matrix. LR proves to be a good prediction with an accuracy of 80% approximately.