**RESEARCH ARTICLE** 



# Is price a significant predictor of the churn behavior during the global pandemic? A predictive modeling on the telecom industry

Vishal Shukla<sup>1</sup> · Sanjeev Prashar<sup>2</sup> · Bhartrihari Pandiya<sup>3</sup>

Received: 25 June 2021 / Accepted: 4 November 2021 / Published online: 22 November 2021 © The Author(s), under exclusive licence to Springer Nature Limited 2021

#### Abstract

The recent pandemic has affected the world in many aspects including, communication. Telecom and internet-based communication have witnessed a drastic upsurge due to lockdown and consequent work-from-home situation, being termed as the new normal. Due to the low switching costs and stiff competition, telecom service providers are struggling hard to attract new customers and prevent the existing ones from switching to rival telecom service providers. Hence this study is undertaken with an aim to discover the key factors in the order of their relative worth that could be focused upon by the telecom companies to prevent their customers from churning. For the study, linear discriminant analysis was applied on the collected dataset to predict the customers' churn behavior. It was found that tariff rates for domestic calls and the number of calls made to the customer service were the significant predictors of the customer churn behavior.

Keywords Churn behavior  $\cdot$  Telecom service  $\cdot$  Customer loyalty  $\cdot$  Customer service  $\cdot$  Predictive modeling  $\cdot$  Discriminant analysis

# Introduction

The global pandemic caused due to the novel coronavirus led to a nationwide lockdown in India starting mid-March 2020. People were confined to the boundaries of their homes and offline activities were significantly replaced by online activities. This became a new normal leading to an increase in internet and telephony usage. The COVID-19 pandemic has changed the pattern of telecom consumption habits for not only internet usage but also for telephonic conversation. Due to the nationwide lockdown and norms related to social distancing, and subsequent confinement to homes, an upsurge in usage of telecom services was witnessed. Even when this pandemic ends, it is expected that this habit of

 Vishal Shukla vishal.azh@gmail.com
 Sanjeev Prashar sprashar@iimraipur.ac.in
 Bhartrihari Pandiya

bhartrihari.nits@gmail.com

<sup>1</sup> Presidency University, Bangalore 560064, India

<sup>2</sup> Indian Institute of Management, Raipur 493661, India

<sup>3</sup> Presidency College, Bangalore 560024, India

staying and working remotely will remain in the coming times as people and various organizations globally have gradually adjusted to this new normal (De et al. 2020). This creates an increased demand for the internet and telephony services, both in numbers and quality of services, leading to stringent competition in the telecom industry.

Being locked down and immobile in their homes, increased the internet usage and calling as people had more leisure time too. This allowed them to engage more with browsing online content or social media activity. There is a drastic increase in the usage of video conferencing services, like Zoom, content delivery services like Akamai (Branscombe 2020). This has led to a race between telecom operators to provide cheap data and calling plans so that they can take a competitive edge and increase their revenues.

Around 5.74 million subscribers requested mobile number portability (MNP) in March 2020, but this number increased by 7.12% to 6.18 million in November 2020 (TRAI Report 2020). This unabated increase in requests for number portability within eight months provides firm evidence of dissatisfied customers, which is a matter of serious concern for telecom service providers.

Though revenues might be short-termed, to have longterm sustainability, telecom services providers need to work upon customer loyalty as the market is gradually saturating (Hwang and Kim 2018). The customers are exhibiting churn behavior and are switching from one telecom operator to another based on certain unidentified reasons. To sustain their profitability, telecom operators must make a serious effort to identify the factors responsible for customer churning.

Therefore, this research has been endeavored to identify the variables of churning intention in the telecom market and their relative worth. These variables of churning, once identified, can specify the reasons for churning. In addition, this research also attempts to find solutions that can lead to a decrease in the probability of customer churn from their telecom service provider. The service providers can draft the requisite strategy, using these variables, to reduce the churning of their customers.

To fulfill these objectives, the research questions framed were:

RQ1. What are the variables that lead to churning behavior in the telecom sector?

RQ2. What is the relative worth of each of these variables that lead to churning behavior in the telecom sector?

RQ3. Which of the variables of churning behavior lead to a decrease in the probability of customer churn?

The paper has been organized is in the following format: it opens with the introduction of the theme of the research, i.e., how the pandemic has added to the churning in the telecom industry. The subsequent part details the overview of the research work available in the extant literature and identification of the literature gap, followed by the outline of the material and methods employed in this research work. The description of the data is followed by the information about the discriminant analysis, problem formulation, and data pre-processing. Thereafter, exploratory data analysis has been described. The next section details about the development of the prediction model and its testing divided into eight steps, which illustrate the step-by-step analysis, their respective outputs, and their interpretations. Finally, the findings of the research work and concluding remarks have been presented, followed by the managerial implications and suggestions for telecom service providers. The paper closes with a discussion on limitations along with the future scope of the study.

# Literature review

The literature review has been undertaken to comprehend the various aspects of churning behavior. The varied methods and techniques used by previous researchers in predicting the churn have been discussed in the context of the present research work. Besides this, the changing consumption behavior of customers during the current pandemic situation was also studied. Churning refers to the customer loss because of switching to a different service provider from their present ones during a specific phase of time (Mock 2011). Companies aim to maximize customer retention as it is a major concern across all industries including telecommunication (Rajamohamed and Manokaran 2018; Choudhari and Potey 2018). The timing of the prediction of churning behavior is equally critical. Earlier the prediction, the better it is for the organizations. As per Alboukaey et al. (2020), the monthly churn prediction is partly inefficient; hence daily churn prediction models are required. Thus, churn management includes the identification of valuable customers who might exhibit churning behavior and being proactive to prevent the churn.

It is an accepted concept that retaining the present customers is more economic than acquiring new ones (Rust et al. 1995; Heskett et al. 1994). Van den Poel and Lariviere (2004) observed that the customer retention rate results in a noteworthy effect on businesses. The customers switch from one option to other for gaining the advantage of the best deals possible (Capizzi and Ferguson 2005; Bellizzi and Bristol 2004) or under the influence of their friends' churning behavior (Ferreira et al. 2019).

Geetha and Kumari (2012) did a thorough study of the usage pattern in case of non-revenue earning customers causing churn in revenue. Customers with changes in usage revenue along with more use of value-added services are expected to show churn behavior.

The details of some other research work related to churning behavior in the Indian telecom sector have been delineated in Table 1.

Various researches have shown that customer buying intention has been affected during this pandemic period (Islam et al. 2021; Wang and Na 2020; Laato et al. 2020; Kabadayi et al. 2020). Laato et al. (2020) studied the purchasing behavior of the customers in unusual times of pandemic with the stimulus–organism–response approach. A strong connection was posited between self-intention to isolate and unusual purchase intent. It proves customer behavior is directly linked to self-isolation time. Many organizations shifted their digital infrastructure to facilitate work-fromhome (WFH) for the time being (Khetarpal 2020; Akala 2020). The meetings and interactions are held online due to the global pandemic situation.

Conclusively, it is noted that despite a large number of studies that have been undertaken to decipher churning behavior as well as customer behavior during the pandemic, there exists a research gap. Evidently, there are limited researches in this area of customer churn during the pandemic period in the telecom sector, and no research has been attempted, which categorizes the variables of churning in order of their relative worth. The usage of a linear discriminant analysis leads to a predictive examination where the prognosis of indicators of churning can be done. The

Authors	Dataset	Tools/method used	Results
Rai et al. (2020)	Customers of telecommunication companies	CART algorithm	Proposed a classification model call tree for telecom customer churn prediction
Ranjan and Sood (2020)	Twitter handle of five major Indian telecom brands	Monthly sentiment score	Proposed model using month wise sentiment score for telecom churn prediction
Bhattacharjee and Dash (2020)	Online content by users in communi- ties of social media	Netnography approach	Factors of churning—issues of data speed, ineffective relationship build- ing, service area coverage issues and billing issues
Borah et al. (2020)	Primary data from a telecom company	Resource exchange theory and a multi-method approach	Chances of customer churn is high in case of process failures when compared to outcome failures
Mukhopadhyay et al. (2021)	Indian telecom companies	Machine learning	Model of churn prediction and forecast- ing of churning behavior

Table 1 Literature review of churning behavior in the Indian telecom sector

significance of the proposed study of the churning variables is that the telecom service providers can focus on the more significant ones on a priority basis and pay lesser heed to the less critical ones.

The present study focuses on the switching behavior during the global pandemic and that is the novelty in approach as the telecommunication industry is in a spot of concern in these changing circumstances. Many studies pertaining to predicting customer churn behavior have been conducted using various classification techniques, but none of these have used the discriminant analysis approach that makes this study unique in terms of methodology.

# **Materials and methods**

## **Data description**

The dataset used in this study was acquired from the three private sector telecom companies operating in the Delhi-NCR region of India. The dataset consisted of 2500 customer details and contained information about the number of months they were subscribed to the services of the company (subscription time), the number of voicemails they received (num\_voicemail), the total amount they spent on the domestic calls made (dom\_call\_charges), the total amount they spent on international calls made (intl\_call\_ charge), the total amount they spent on availing internet services (total internet charge), the total number of calls they made to the customer service of their telecom service provider (num\_ccs\_calls), network quality (network\_quality) measured in terms of the availability of the network ranging from 1 to 5 where '5' and '1' represent the best and the worst respectively, internet speed (int\_speed) categorized as 1 = 4G, 2 = 3G, and 3 = 2G. Also, average call drops per day (call drops ave day) were measured.

All these variables were the independent variable X and we had a dependent variable Y named churn consisting of two categories—0 and 1, where 0 indicates the customers who remained with the service provider or those who didn't churn, while 1 represents that the existing customer churned and switched to another service provider. The entire data was for a period of nine months ranging from March 2020 when the first phase of lockdown was enforced in India to November 2020, when maximum unlock was announced.

### Methodology adopted: discriminant analysis

According to Hair et al. (2014), discriminant analysis is an appropriate technique for statistical treatment for situations where the outcome is categorical in nature and the predictor variables are metric or continuous. This method uses the information offered by the independent variable X to classify the dependent variable Y into two groups. In this study, Y is classified as a categorical variable or group consisting of customers who do no churn (denoted by 0) and customers who churn (denoted by 1).

This statistical technique entails obtaining a variate that is the linear combination of independent variables, which optimally discriminates between objects in the prior defined groups. The equation for discriminant analysis is almost similar to that of the multiple regression and is expressed as follows:

$$Z_{jk} = a + W_1 X_{1k} + W_2 X_{2k} + \dots + W_n X_{nk}$$

where  $Z_{jk}$  is discriminant Z score of discriminant function *j* for object *k*, *a* is intercept,  $W_i$  is discriminant weight for independent variable *i*,  $X_{ik}$  is independent variable *i* for object *k*.

Though there are many other classification techniques like logit regression, ANN, cluster analysis, etc., but Hair

Table 2         Target variables	'data.frame' 2500 observations of 10 variables		
	\$ subscription_time	int 128 137 84 147 141 168 95 93 73 130	
	<pre>\$ num_voicemail</pre>	int 25 0 0 0 37 0 0 0 0 0	
	<pre>\$ num_ccs_calls</pre>	int 1 0 2 0 0 1 3 3 1 0	
	\$ dom_call_charges	num 56.1 48.7 59.8 36.2 58.6	
	\$ intl_call_charge	num 2.7 3.29 1.78 1.92 3.02 3.02 3.32 2.19 3.51 2.57	
	<pre>\$ total_internet_charge</pre>	num 16.78 10.3 5.26 8.76 18.87	
	<pre>\$ network_quality</pre>	int 3 5 2 4 3 4 2 1 3 2	
	\$ int_speed	int 1 1 2 2 1 1 1 3 2 1	
	<pre>\$ call_drops_ave_day</pre>	int 2 2 1 2 1 0 3 2 0 1	
	\$ churn	Factor w/2 levels "0", "1": 1 1 1 1 1 1 1 1 1	

1

0.1432

Table 3	Baseline churn rate	0	
		0.8568	

et al. (2014) noted that discriminant analysis is the most suitable for testing the postulate 'the group means of a set of independent variables for two or more groups are equal.' Hence, we have chosen this technique for our study.

#### Problem formulation and data pre-processing

At the initial step, the data file was processed using R programming to convert the target variables into categorical variables. To run a discriminant analysis, the target or the outcome variable Y must be categorical. The output is shown in Table 2.

After the conversion of the target variables to categorical variables, we determined the baseline churn rate for the dataset. The output is given in Table 3. From the above table, it is obvious that the dataset is imbalanced due to the baseline churn rate being only 14.32%, meaning there are more zeros than ones. This will considerably impact our interpretation of the model performance measures.

In the next step, we explored our dataset to have a preliminary insight on how our outcome variable *Y*, labeled as churn, is impacted by the dependent variables taken for the study.

## **Exploratory data analysis**

The boxplots (Fig. 1) reflect that the subscription time (No. of months the customer stayed with the service provider) has no impact on customer churn intention. The second boxplot (Fig. 2) indicates that customers who churn, for them the number of voicemail messages is lower as compared to those who don't. Tariffs for domestic calls (Fig. 3) and international calls (Fig. 4), and internet usage charge (Fig. 5), are generally higher for subscribers who churn as against those who don't. This

















charges vs. churn



indicates that the customers who churn, realize at some point in time, that they are paying more money for the services and are not happy with the value for money that they are getting. The availability of the network quality (Fig. 6) is low for those who churn in contrast to those who don't. Internet speed (Fig. 7) and average call drops in a day (Fig. 8) do not seem to be significant predictors of churn intention.

With respect to the number of calls made to the customer service by the customers, we found that the churn rate is relatively high (Fig. 9). This is a clear indication that the customers who churned, had attempted to contact the customer service many times, but failed to get a satisfactory resolution to their problems.

## Development and testing of the prediction model

Based on the assumptions derived from the exploratory data analysis, we further explored our variables using discriminant analysis. Our data had variables exhibiting various dimensions, so we scaled these to avoid the influence of the range of each variable on the discriminant coefficients. Thereafter, we divided the dataset into two parts-training and testing sets. The training and testing dataset ratio was kept as 70:30. Seventy percent of the dataset (1750) observations were kept for the training purpose, and the remaining 750 observations were kept for the testing purpose. The baseline churn rates for training data and test data are given in Tables 4 and 5.

The output given in the above tables indicate that the distribution of partition data is correct.







**Fig. 8** Boxplot of average call drops in a day vs. churn





Table 4Baseline churn rate fortraining data	0	1
	0.8565714	0.1434286
Table 5         Baseline churn rate for           test data	0	1
	0.8573333	0.1426667

We further developed our prediction model using the training dataset and then tested the performance of the developed model using the test dataset. This was done in eight steps.

# Step 1: testing the significance of discriminant function using MANOVA

In this first step, we used MANOVA for testing the significance of the discriminant function. For this purpose, we formulated a null and an alternate hypothesis as given below:

 $H_0 \ \mu_1 = \mu_2 = ... = \mu_k$  (The independent variables taken in the study are not significant discriminators of the churning and non-churning groups).

H<sub>1</sub>  $\mu_1 \neq \mu_2 \neq ... \neq \mu_k$  (At least one independent variable is a significant discriminator of the churning and non-churning groups).

The output of the MANOVA test is shown in Table 6.

From this, we can see that Wilks  $\lambda$  for MANOVA is 0.8883, which is very close to 1. This indicates that there is a relatively low extent of discrimination in the model. However, we also see that the *p*-value is highly significant, and thus null hypothesis stands rejected. From these results, we can infer that the discriminant model is vastly significant.

#### Step 2: developing the Fisher discriminant function

Here, we identified a set of attributes, which discriminate the churning out customers from those who won't. For this, we developed the Fisher discriminant function (FDF) using the R packages DiscriMiner and MASS. Since our outcome variable *Y* has only two values, so we will be having only a single discriminant function represented by DF1. The coefficients of the discriminant function are present in Table 7.

Then, the independent X variables were sorted based on their coefficients in descending order. This provides us an understanding of how each X variable is having an influence on differentiating the Y variable (Table 8).

From the above results, we can infer that the domestic call charges had the maximum impact, followed by the number of calls made to the customer service.

# Step 3: Differentiating the Wilks $\lambda$ and the individual independent variables

The output given in Table 7 indicates that the *p*-value is insignificant for the subscription time, network quality, internet speed, and average call drops in a day, but is highly significant for the other *X* variables. From this, it can be interpreted that all the *X* variables, except the subscription time, network quality, internet speed, and average call drops in a day, are excellent predictors in terms of differentiating our customer groups.

# Step 4: estimating the correlation between the independent *X* variables and the discriminant function

The correlation coefficients of each of the variables help to ascertain the comparative importance of each of the X variables. From Table 7, it is obvious that the domestic call charge is the most noteworthy variable in its ability to discriminate the churning and non-churning groups.

# Step 5: records classification based on discriminant analysis of X variables and predicting Y variables for the test set

From the above results, it is observed that the discriminant model that we developed is significant. This information was used to classify records belonging to either of the group of churning customers or non-churning customers depending on the X variables. The lda() function in R was used to categorize records based on the value of X variables and predict the class and probability for the test set. The output is presented in Table 9.

The difference in the group means from Table 9 gives us a better idea of the factors that have the highest and

#### Table 6 MANOVA output

e-16***

Significance codes '\*\*\*' 0.001, '\*\*' 0.01, '\*' 0.05

# Table 7 Fisher discriminant function output

Descriptive discriminant analysis

\$Power (discriminant power)				
	Correlation ratio	Wilks $\lambda$	F statistic	<i>p</i> values
subscription_time	0.00003103	0.99996897	0.05424851	0.81585616
num_voicemail	0.00623413	0.99376587	10.96562001	0.00094709
num_ccs_calls	0.04461582	0.95538418	81.63045915	0.00000000
dom_call_charges	0.04702368	0.95297632	86.25333266	0.00000000
intl_call_charge	0.00670115	0.99329885	11.79263079	0.00060854
total_internet_charge	0.00715553	0.99284447	12.59801512	0.00039637
network_quality	0.00012328	0.99987672	0.21552171	0.64253135
int_speed	0.00001860	0.99998140	0.03251450	0.85692401
call_drops_ave_day	0.00047384	0.99952616	0.82866666	0.36278387
\$Values (table of eigenvalues)	)			
	Value	Propor	tion	Accumulated

DF1	0.111	100.000	100.000
\$discrivar (discriminant var	iables)		

	DF1
Constant	-0.014875
subscription_time	-0.009916
num_voicemail	-0.240598
num_ccs_calls	0.659023
dom_call_charges	0.686517
intl_call_charge	0.288374
total_internet_charge	0.300188
network_quality	-0.040032
int_speed	0.049298
call_drops_ave_day	0.005643
\$discor (correlations)	

	DF1
subscription_time	0.01667
num_voicemail	-0.23624
num_ccs_calls	0.63200
dom_call_charges	0.64883
intl_call_charge	0.24493
total_internet_charge	0.25310
\$Scores (discriminant scores)	

	$z_1$
1	0.5183
3	0.7038
6	- 1.5439
7	1.0632
9	0.4767
10	- 1.5397

Tab	le 8	Discriminant	variable	coefficients	(descending	order)
-----	------	--------------	----------	--------------	-------------	--------

Constant	-0.014875
dom_call_charges	0.686517
num_ccs_calls	0.659023
total_internet_charge	0.300188
intl_call_charge	0.288374
int_speed	0.049298
call_drops_ave_day	0.005643
subscription_time	-0.009916
network_quality	-0.040032
num_voicemail	-0.240598

significant contribution to discriminate between the two groups. The above output reflects that the group means difference is highest in the case of domestic call charges and lowest for average call drops in a day. The linear discriminant coefficients presented in Table 9 also reflects a similar pattern and indicates the independent variable that has maximum contribution to group partition.

# Step 6: visualization of the groups

Further to have better insights into how our probable churning customers look like as compared to their non-churning counterparts, we visualized the data in the form of graphs as shown in Fig. 10.

Figure 10 gives a graphical visualization of the groups that we created through the discriminant analysis. They closely resemble Wilks  $\lambda$  value of 0.8883, which was arrived at Step 1 (MANOVA output). These graphical visualizations clearly indicate that though our discriminant model is significant, both the groups are not entirely exclusive from each other. Rather, they have some overlap between them.

#### Step 7: making predictions on the test set

Now having built the model and performed its training, we used our test dataset to predict the churning or non-churning behavior of the customers based on the independent variables of the testing dataset. To make predictions on the testing dataset, we applied the discriminant model that was constructed with the training dataset. The aim of performing this step is to measure the improvement in the performance

#### Table 9 Records classification based on discriminant analysis

		-			
Prior prob	abilities of groups				
0					1
0.8565714	Ļ				0.1434286
Group mea	ans				
	subscription_time	num_voicemail	num_ccs_calls	dom_call_charges	intl_call_charge
0	0.0008472916	0.04401104	-0.07230058	-0.08059323	-0.00963103
1	0.0167712294	-0.18362145	0.52675205	0.53510858	0.22189972
Group mea	ans				
	total_internet_charge	netwo	rk_quality	int_speed	call_drops_ave_day
0	-0.04759535	-0.00	5860001	0.002880849	0.02034905
1	0.18805228	-0.03	37341866	0.015230175	-0.04208807
Coefficient	ts of linear discriminants				
					LD1
subscriptic	on_time				-0.009915827
num_voice	email				-0.240598021
num_ccs_o	calls				0.659022835
dom_call_	_charges				0.686516594
intl_call_c	charge				0.288373562
total_inter	net_charge				0.300188294
network_q	Juality				-0.040031938
int_speed					0.049298274
call_drops	_ave_day				0.005642648



Fig. 10 Visualization of the groups created through discriminant analysis

of our discriminant model on a new dataset. The output is presented in Table 10.

## Step 8: evaluation of the model performance measures

Table 10 indicates the model performance on the test data. The error rate of the model comes out to be 0.143. Hence, the accuracy of the model can be calculated as 1 - error rate, which comes out to be 1 - 0.143 = 0.857 = 85.7%. This is fairly good. However, during the initial data exploration, we found that our data is not balanced, therefore accuracy cannot be the standalone predictor of the fitness of the model. So, to ascertain that whether this model is a robust model, we considered looking into some other measures of model performance.

Sensitivity could be another indicator of the model performance. Also known as the true positive rate, this is the percentage of samples that are truly positive that give a positive result using the test in question (Steward 2019). Sensitivity is the ability to predict accurately about the customers who are likely to churn. In our model, the sensitivity comes out to be 11.2%, which is extremely low. For a better predicting of churn, it is necessary that the model should pick the positives as positives. Another measure is specificity, which is the percentage of samples that test negative using the test in question that are genuinely negative (Steward 2019). This is also termed as the true negative rate. In our case, the specificity came out to be 98.1% and this reflects a lack of balance between the two performance measures. To get an optimum balance for sensitivity and specificity, we decided to vary the threshold of the model to some other values from the default value of 50%. The respective outputs are given in Table 11.

Further, we compared the values of sensitivity and specificity at various threshold values. The respective outputs are given in Table 12.

Now, we have the accuracy, sensitivity, and specificity at various threshold levels as indicated in Table 12, but we will take into account the optimal cut-off or threshold value. For this, we plot these three values for different thresholds and the point of intersection will be the optimal threshold value. The output from the plot is given in Fig. 11.

In Fig. 11, we can see that the three values for threshold values of 15%, 16%, and 17% intersect with each other. But, if we compare the accuracy, sensitivity, and specificity of these three threshold values, we find that the accuracy, sensitivity, and specificity values at the threshold value of 15% are more balanced as compared to the other values. Thus, it is apparent that the optimum threshold for the model is 15%.

 Table 11
 Confusion matrix at different threshold levels

	pred.1	pred.0
At threshold of 30%		
actual.1	34	73
actual.0	42	601
At threshold of 20%		
actual.1	59	48
actual.0	99	544
At threshold of 17%		
actual.1	70	37
actual.0	137	506
At threshold of 16%		
actual.1	72	35
actual.0	150	493
At threshold of 15%		
actual.1	73	34
actual.0	165	478
At threshold of 10%		
actual.1	86	21
actual.0	272	371

Table 10 Model performance on test data

				I	pred.1				pred.0
actual.1 12				12				95	
actual.0		12							631
	ER	Sens	Spec	Precision	Recall	TPR	FPR	F	Youden
	0.143	0.112	0.981	0.5	0.112	0.112	0.0187	0.183	0.0935

 
 Table 12
 Accuracy, sensitivity, and specificity at different threshold levels

ER	Sens	Spec
At 30%		
0.1533333	0.317757	0.9346812
At 20%		
0.196	0.5514019	0.8460342
At 17%		
0.232	0.6542056	0.7869362
At 16%		
0.2466667	0.6728972	0.7667185
At 15%		
0.2653333	0.682243	0.7433904
At 10%		
0.3906667	0.8037383	0.5769829



Fig. 11 ROC plot of sensitivity and specificity for various threshold values

Upon drawing predictions at this threshold, the accuracy and specificity comes to be 74% and sensitivity is very close to these values at 68%.

# **Findings and conclusions**

Based on the discriminant coefficients and the correlation ratios provided by the model, we can predict and conclude that an increase in domestic call charges, number of calls made to the customer service, internet usage charges, and international call charges are the strong predictors of the customer churn intention and increase the probability of a customer churn. Thus, the relative worth of the variables that lead to churning behavior in the telecom sector is the value contribution of research. The higher the worth of the factor, the greater is its role in the churn intention of the customers. Domestic call charges are the best indicator of the churning behavior because a slight increase in the price of domestic calls disrupts existing tariffs, emanating confusion in the mind of the customer that may lead to switching from the telecom service provider. Our finding confirms the findings of Chadha and Bhandari (2014), according to which tariff is one of the major antecedents of customers switching toward MNP. The price sensitivity has increased in the pandemic due to a reduction in the income levels of the people. The number of calls made to the customer service is a strong indication of frustration or uncertainty in customers' minds. Apart from some queries to the customer service, most of the calls are made to register the complaints and grievances to vent out their problems and ask for solutions. The aim should be to process the calls and settle the query as soon as possible. The internet charges and the speed are issues of concern for customers because the usage has increased manifolds since the digital upsurge. A speedy internet at affordable rates is expected by customers. Our this finding is in congruence with the findings of Dey et al. (2020), who conducted a study in the UK telecom market and concluded that speed is an important factor that influences customer satisfaction and churning intention. The scale of international calls is increasing day by day due to globalization so the charges also need to be reasonable and competitive so that customers can make international calls at affordable prices. The model also gave a meaningful insight that an increase in the number of voicemail messages, network quality, and duration of the subscription may decrease the probability of customer churn. In these situations, they are more likely to continue with the existing telecom service provider. Thus, this study adds more dimensions to the churning factors identified by the study conducted by Geetha and Kumari (2011) such as the proportion of local and STD calls made to other networks along with higher usage of value-added services.

#### Implications and suggestions

The Indian telecom sector is undergoing intense competition and switching cost in this sector is low. Thus, it becomes imperative to understand customer expectations and their viewpoint regarding their switching and retention behavior. The facility of MNP has empowered customers to a great level and in such situations, it becomes pertinent for the telecom service providers to get into the nuances of customer's expectations with respect to internet and telephony services (Chadha and Bhandari 2014). The meaningful insights given by our discriminant model can help the telecom service providers and the telecom regulators formulate strategies to reduce the customer churn and cases of MNP, which is huge as indicated in a 2020–2021 Report by the Telecom Regulatory Authority of India.

The key lessons that these organizations can take from this study are to find an amicable solution to resolve the customer issues within the first or the second call. Dignity and respect are the essential components of customer loyalty (Bahri-Ammari and Bilgihan 2019), and these are reflected in how quickly a customer's problem is addressed. If a customer has to make multiple calls to customer service pertaining to an issue, then it may result in churn intention. Also, there should be a robust, smooth, and organized escalation procedure for attending to those issues which are not resolved within two calls. A good network availability for calling and smooth internet surfing experience matters a lot to customers and a poor network leads to churning. Thus, telecom service providers should focus on providing the best network quality to the customers.

Finally, to address the most important factor, i.e., the price, the telecom service providers should offer more lucrative and affordable plans. In the present market situation, customers are informed and can compare several value offerings of similar products. To maximize the customer base, companies can cross-sell products like more data bundles, tariff discount vouchers, etc., which is the need of the hour in the present scenario, where online classes and work-from-home have become the new normal. Ultimately, to retain the customers, the telecom service providers should show some empathy toward their customers and should provide them better internet and telephony services at affordable prices rather than cashing the situation. This would help develop better customer relations and loyalty that would ultimately reduce the churn intention of the customers. Though these findings are derived from the study conducted on the telecom industry but can be generalized to different service sectors/industries.

#### Limitations and future scope of the study

This study has used only one classification technique to predict the churn behavior, and that too only in the context of telecom services. Similar studies could be further extended to other services such as healthcare, entertainment, education, etc. Also, many other machine learning models like logistic regression, decision tree, random forest, clustering, and Neural Network models could be applied and the comparison of the accuracy of these models could be studied. A collection of data from other countries could provide meaningful insight for comparing the churn intention of the telecom customers across different countries. The churn intention may also be affected by cultural and societal norms and these variables were not a part of this study. A future study can be done by taking into account the cultural and societal differences among different countries and their effect on customer satisfaction and loyalty.

# References

Akala, Adedayo. 2020. More big employers are talking about permanent work-from-home positions. https://www.cnbc.com/2020/ 05/01/major-companies-talking-about-permanent-work-fromhome-positions.html. Accessed 9 Feb 2021.

- Alboukaey, Nadia, Ammar Joukhadar, and Nada Ghneim. 2020. Dynamic behavior based churn prediction in mobile telecom. *Expert Systems with Applications* 162: 113779. https://doi.org/ 10.1016/j.eswa.2020.113779.
- Bahri-Ammari, Nedra, and Anil Bilgihan. 2019. Customer retention to mobile telecommunication service providers: The roles of perceived justice and customer loyalty program. *International Journal of Mobile Communications* 17 (1): 82–107. https://doi. org/10.1504/IJMC.2019.096518.
- Bellizzi, Joseph A., and Terry Bristol. 2004. An assessment of supermarket loyalty cards in one major US market. *Journal of Consumer Marketing* 21 (2): 144–154. https://doi.org/10.1108/ 07363760410525704.
- Borah, Sourav Bikash, Srinivas Prakhya, and Amalesh Sharma. 2020. Leveraging service recovery strategies to reduce customer churn in an emerging market. *Journal of the Academy of Marketing Science* 48, no. 5: 848-868. https://doi.org/10.1007/ s11747-019-00634-0
- Branscombe, Mary. 2020. The Network Impact of the Global COVID-19 Pandemic. *The New Stack*. https://thenewstack. io/the-network-impact-of-the-global-covid-19-pandemic/. Accessed 12 Feb 2021.
- Capizzi, Michael T., and Rick Ferguson. 2005. Loyalty trends for the twenty-first century. *Journal of Consumer Marketing* 22 (2): 72–80. https://doi.org/10.1108/07363760510589235.
- Chadha, Suresh K., and Namita Bhandari. 2014. Determinants of customer switching towards mobile number portability. *Paradigm* 18 (2): 199–219. https://doi.org/10.1177/0971890714558708.
- Choudhari, Atul Sunil, and Manish Potey. 2018. Predictive to prescriptive analysis for customer churn in telecom industry using hybrid data mining techniques. In 2018 Fourth international conference on computing communication control and automation (ICCUBEA), 1–6. IEEE.
- De, Rahul, Neena Pandey, and Abhipsa Pal. 2020. Impact of digital surge during COVID-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management* 55: 102171. https://doi.org/10.1016/j.ijinfomgt.2020.102171.
- Dey, Bidit Lal, Wafi Al-Karaghouli, Stanimir Minov, Mujahid Mohiuddin Babu, Angela Ayios, Syed Sardar Mahammad, and Ben Binsardi. 2020. The role of speed on customer satisfaction and switching intention: A study of the UK mobile telecom market. *Information Systems Management* 37 (1): 2–15. https://doi.org/ 10.1080/10580530.2020.1696526.
- Ferreira, Pedro, Rahul Telang, and Miguel Godinho De. Matos. 2019. Effect of friends' churn on consumer behavior in mobile networks. *Journal of Management Information Systems* 36 (2): 355–390. https://doi.org/10.1080/07421222.2019.1598683.
- Geetha, M., and Jensolin Abitha Kumari. 2012. Analysis of churn behavior of consumers in Indian telecom sector. *Journal of Indian Business Research* 4 (1): 24–35. https://doi.org/10.1108/17554 191211206780.
- Hair, Joseph F., William C. Black, Barry J. Babin, and Rolph E. Anderson. 2014. *Multivariate data analysis: Pearson new international edition*. Essex: Pearson Education Limited.
- Heskett, James L., Thomas O. Jones, Gary W. Loveman, W. Earl Sasser, and Leonard A. Schlesinger. 1994. Putting the serviceprofit chain to work. *Harvard Business Review* 72 (2): 164–174. https://doi.org/10.1016/j.neucom.2016.12.009.
- Hwang, Shin Young, and Seongcheol Kim. 2018. Does mIM experience affect satisfaction with and loyalty toward O2O services? *Computers in Human Behavior* 82: 70–80. https://doi.org/10. 1016/j.chb.2017.12.044.
- Islam, Tahir, Abdul Hameed Pitafi, Vikas Arya, Ying Wang, Naeem Akhtar, Shujaat Mubarik, and Liang Xiaobei. 2021. Panic buying

in the COVID-19 pandemic: A multi-country examination. *Journal of Retailing and Consumer Services* 59: 102357. https://doi.org/10.1016/j.jretconser.2020.102357.

- Kabadayi, Sertan, Genevieve E. O'Connor, and Sven Tuzovic. 2020. The impact of coronavirus on service ecosystems as service megadisruptions. *Journal of Services Marketing* 34 (6): 809–817. https://doi.org/10.1108/JSM-03-2020-0090.
- Khetarpal, Sonal. 2020. Post-COVID, 75% of 4.5 lakh TCS employees to permanently work from home by '25; from 20%. Business Today, 30. https://www.businesstoday.in/current/corporate/ post-coronavirus-75-percent-of-3-5-lakh-tcs-employees-perma nently-work-from-home-up-from-20-percent/story/401981.html. Accessed 21 Feb 2021.
- Laato, Samuli, A.K.M. Najmul Islam, Ali Farooq, and Amandeep Dhir. 2020. Unusual purchasing behavior during the early stages of the COVID-19 pandemic: The stimulus–organism–response approach. Journal of Retailing and Consumer Services 57: 102224. https://doi.org/10.1016/j.jretconser.2020.102224.
- Mukhopadhyay, Debajyoti, Aarati Malusare, Anagha Nandanwar, and Shriya Sakshi. 2021. An Approach to Mitigate the Risk of Customer Churn Using Machine Learning Algorithms. In *Machine Learning for Predictive Analysis*, 133-142. Springer, Singapore. https://doi.org/10.1007/978-981-15-7106-0\_13
- Mock, Donald. 2011. Comparing carrier churn. http://www.fool.com/ investing/general/2011/06/15/wireless-smackdown-churn.aspx. Accessed 17 March 2021.
- Rai, Sandeep, Nikita Khandelwal, and Rajesh Boghey. 2020. Analysis of customer churn prediction in telecom sector using cart algorithm. In *First International Conference on Sustainable*

Technologies for Computational Intelligence, 457-466. Springer, Singapore. https://doi.org/10.1007/978-981-15-0029-9\_36

- Rajamohamed, R., and J. Manokaran. 2018. Improved credit card churn prediction based on rough clustering and supervised learning techniques. *Cluster Computing* 21 (1): 65–77. https://doi.org/10.1007/ s10586-017-0933-1.
- Rust, Roland T., Anthony J. Zahorik, and Timothy L. Keiningham. 1995. Return on quality (ROQ): Making service quality financially accountable. *Journal of Marketing* 59 (2): 58–70. https:// doi.org/10.1177/002224299505900205.
- Steward, Karen. 2019. Sensitivity vs Specificity. *Technology Networks*, https://www.technologynetworks.com/analysis/articles/sensitivity-vs-specificity-318222. Accessed 19 Feb 2021
- TRAI. 2020. Press Release No. 493/2020. https://www.trai.gov.in/sites/ default/files/PR\_No.49of2020\_1.pdf. Accessed 21 Feb 2021
- Van den Poel, Dirk, and Bart Lariviere. 2004. Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research* 157 (1): 196–217.
- Wang, H. Holly., and H.A.O. Na. 2020. Panic buying? Food hoarding during the pandemic period with city lockdown. *Journal of Integrative Agriculture* 19 (12): 2916–2925. https://doi.org/10.1016/ S2095-3119(20)63448-7.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.