Machine Learning as Intelligent tool for Churn Prediction in Telecommunication Industry

Megha Gupta
M.Tech, CSE

Anju Bhandari Gandhi, PhD
Associate Professor, CSE

S. C. Gupta, PhD
Chairperson

ABSTRACT
In most of the sectors dominated by the cost of changing the landscape, wear of the activities is painted or blurred by the clients (clients who wish to modify or change their suppliers for various reasons). This phenomenon is widespread in the telecommunications industry and all aspects related to its lead to the belief that it increases sharply. Since the market is very competitive and the number of prepaid customers, "it increases, it is vital that companies to actively confront with the distraction of customers, the identification of behaviour that could ultimately be created." In this article we have proposed A hybrid model machine learning to predict distortion in mobile telecommunications networks The experiments were carried out using a variety of tools for machine learning, together with a set of real data from the open data provider to evaluate the manufacturer's productivity. The results show that the new hybrid model is more accurate than the individual methods. Auto learning is a subset of artificial intelligence, commonly used statistical techniques to give computers the ability to "learn" the data.

Keywords
Churn prediction, logistic regression, and telecommunications, voted perceptron.

1. INTRODUCTION
The Indian telecommunications industry is one of the fastest proliferating sectors in the world. Mobile prepaid customers are vibrant in changing their mobile operators within very short span. Churn rate increases pungently in parallel to the growth of prepaid mobile subscribers. Customer churn happens to be the most challenging issue for mobile industry irrespective of its rapid growth. This in turn entangled with disloyalty and as the industry saturates, it becomes imperative for the mobile operators to redesign service plans with new offerings to enhance customer loyalty. Customer retention, therefore, is becoming critical to sustain customer base. In this regard it is essential to investigate the basis for switching of the prepaid mobile users in India. Machine learning is a subset of artificial intelligence that often uses statistical techniques to give computers the ability to "learn" with data. Machine learning is sometimes conflated with data mining, where the latter subfield focuses more on exploratory data analysis and is known as unsupervised learning.

For the implementation of machine learning we are able to use many algorithms and they are the following:

Support Vector Machines (SVMs) are controlled learning models with associated learning algorithms that analyse the data used for classification and regression analysis. Given a set of training examples that are labeled as belonging to one or another of two categories, learning algorithm SVM constructs a model that sets new examples in one or another category, making it a non-probable binary linear classifier, When data is not labelled, supervised learning is not possible, and unattended approach to learning, attempts to find a natural grouping of data in groups, and then assign new data for those groups formed is required.

The Bayesian Network is a (probabilistic) model of the probabilistic graphic model that represents a set of variables and their conditional target-acyclic graphic (DAG) dependencies. For example, a Bayesian network may represent probable links between diseases and symptoms. Given the symptoms, the network can be used to calculate the probabilities of different diseases.

Artificial Neural Networks (ANN) is based on a collection of linked units or nodes called artificial neurons that dissolve neurons in the biological brain. Any connection, like synapses in the biological brain, can transmit a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal for additional artificial neurons associated with it. In ANN embodiments, the link between artificial neurons is a real number, and the output of each artificial neuron is calculated by a non-linear function of the sum of its inputs. The connections between artificial neurons are called "edges". Artificial neurons and edges usually have a weight that adapts as learning progresses. Weight increases or decreases signal strength in relation. Artificial neurons may have a threshold so that the signal is sent only if the extra signal passes that threshold. Usually artificial neurons are added in layers. Different layers can perform different types of transformations at their inputs. The signals pass from the first layer (the input layer) to the last layer (the output layer), probably after several passes of the layers. The Recurrent Neural Network (RNN) is a type of artificial neural network in which the links between the nodes form a sequence oriented graph. This allows you to display dynamic time sequence behaviour. Unlike the front neural networks, RNNs can use their internal state (memory) to process the input sequences. This makes them applicable for tasks such as non-segmented, related handwriting recognition or voice recognition.

The term "repeating neural network" is used to indiscriminately designate two broad categories of networks of similar general structure where it is impulse-limited and the other is an endless impulse. Both types of networks show temporary dynamic behaviour. A limited impulse repetitive network is directed acyclic graphics that can be developed and replaced with neural network feedback strictly while repetitive net infinite impulse is a cyclically directed graph that cannot unwind.

2. OBJECTIVES
The objectives of work proposal are as follows:

i.   To review the relevant studies about churn analysis on telecommunications industry presented in the last five years, particularly in the last two years, and introducing these up-to-date studies in the literature.

ii. To determine the data mining methods frequently using churn implementations.

iii. To study and implement methods that can be used in further studies.

3. LITERATURE REVIEW

Literature review is limited to the telecommunications domain, though the churn prediction process will be similar in most settings. After data is collected, it must be labelled with the outcome variable, in this case churn. Then models are trained to predict the impact of independent descriptive variables on the dependent churn variable. Finally, the models are evaluated and put into use for future predictions. Fig 1 provides a summary of churn prediction literature from the telecommunications domain. Telecommunications can be further divided into prepaid and post-paid services. Prepaid services, which are offered without a subscription or contract, tend to include the least amount of customer information and even what is available may not always be accurate due to sharing SIM cards. These accounts may often be shorter in duration when compared to post-paid subscription services and completely anonymous to the service provider. Therefore, attributes based on past and current account behaviour are used to make accurate predictions about future behaviour, including eventual churn. Customer churn prediction is a supervised classification problem where customers are either churners or non-churners.

(**Sherif M Badr, 2017**) this study offers a sample that can offer which client will be organized and who will not. For this purpose, the author uses different technologies for data collection. By using different methods, I make sure to accurately diagnose the customer and identify the reasons that cause user to maintain. With regard to every technology and poor backdrop and popular populace, great help has been made to provide and present user spending, and has given them a high level of status. As related to the clustering technique, it has been proven that Diesel's clustering is suitable for both data and is compatible with data, as well as providing non-raw material to cluster-2 dryers. As long as Airriers and FA-Grad rules are concerned, they have been enforced to define behavioural behaviour of every client and should help in highlighting or blocking certain services of this business organization.

(**Abhinash Mishra,2017**) In this research report, the writer utilizes one-third class-based class and competes with the current leading class, a comprehensive proposal for the customer's former customer analysis in the Telecommunications industry. Experimental results show that routine performance is a great classifier to combat competition with other performance modes, such as reliability, sensitivity, attribute, and error rate. Each user is expecting a good service or reward from service providers. By opening the right customers, be more embarrassed by the faster services because it is very difficult to convey the original buyer. The initial profit type can prevent the company from preventing harm by following customer behaviour. In future, there are two important ways to solve the chronic problem by teaching the ability or ability to learn.

(**Duyen DO, 2017**) this document presented a decision to classify an unbalanced dataset identifying users who are at risk of switching off the service in the telecommunications sector. It mainly focuses on the elements of the engineering and modelling phases, which play a major role in this process. The purpose of this study is to provide a methodology for solving the problem of a popular real-life problem - the

extremely unbalanced data for solving problems for customers. This also applies to other issues in which the data are not balanced. In addition, the paper provided a comparison of these problems to improve it in the future. However, the study has a limit. Engineers' characteristics are based on the history and rules of the most used customers. It is difficult to plan a conservation strategy.

(**Sandra Mitrovic, 2017**) this newspaper is four feet away. First of all, we created the original graphical RFM plug-in that appears in both negotiation and structural information. Second, we order two additional relax random stand generation and grid generation look to the rest of the original node2vec additional halfway, this is a big, great schedule for large ideas. Third, experience and performance testing showcase an exhibition where F-enhanced RFM networks and construction from these RFMs representational training are detained. Finally, this study reflects the NDF's typically retained learning-scale with CD-R graphics for RFM frames to use. This is an initial study that provides a series of interesting questions. Our future for us wanted to make the series of art RFM make the FIFA format ensures that, from efficiency, performance, and with FIB EMA to him, that the invention of the VCR FM was enabled by the Network The work of the various different work versions. We know that the advantage of a bad result is to use a more efficient RFJ optimization network. In addition, it is interesting to see what will come parameters changes, that is, the walk and the measurement of the length of the walk number, the additional increase in the AUC lift scores will be the same. Finally, it assumes that in the order form the selection effect is homophile occupied by entities and their and structural equivalence phenomena is still an open question.

(**Dilip Singh Sisodia, Somdutta Vishwakarma, Abinash Pujahari, 2017**) in this article, find out that machine-translation algorithms work for employee targets and ultimately can exit their executive information and environment-based organization. From experimental results, Random Forest definitely achieves all other classifiers, which are sequential. This study helps organizations identify the causes, causing the staff to leave the organization and take appropriate action to reduce them. This study requires further research to reduce forecasts of error.

(**Ammar A Ahmed, 2017**) this article first provides information about customer revenue and the benefits of telecommunications forecasting to improve the purpose of this document. We reviewed and analysed in detail some of the most important existing pressure prediction methods and collected them in a table for the measured data set to detail. The most important aspect of our review is that the focus is not only on accurately predicting threads, but also the reasons why the transition as well as the methods of handicaps. All this method aims to use the client method, some of which use the direct method based on machine learning and meta-heuristic algorithms, while others improve with indirect methods of data pre-processing and feature selection methods. From this analysis it can be concluded that the most accurate pressure forecasting is achieved using hybrid methods and not individual algorithms. The purpose of the forecasting study has a larger scope that influences us to design our own hybrid model in the future.

(**Paul Bertens, Anna Guitart and ´Africa Peri´a˜nez, 2017**) The results show that a method based on conditional survival complexes can simulate peace both in playtime and on the level, accurately predicting at what level the player will go and how long he will play. This indicates that the model is

resistant to different data splits and is applied to different types of response variables. Although the Cox regression rates were comparatively good, they require a lot of manual effort and suffer from scaling problems that make them unsuitable for the production environment. On the other hand, the survival kits offered are easily adaptable to other varieties and are used in parallel, which can be used not only for multiple cores but also for multiple machines. It gives game developers the ability to get complete Survival Probability curves for each player, and to predict in real-time not only when the player leaves the game, but also at what level they will do it and how many hours they will play before giving up.

(**Rong Zhang, Weiping Li, Tong Mo, 2017**) A good model can be used by a common insurance provider and benefits many companies, such as NEW CHINA LIFE INSURANCE COMPANY LTD. The most advanced methods are either 1) generative models such as logistic regression, complex engineering or 2) deep models that simultaneously dominate the functions or classification models. Partial patterns can be better stored, while deep patterns can be better generalized, but cannot be generalized sufficiently with insufficient data. Therefore, we offer the Deep & Shallow Combination Model (DSM) to take advantage of both the embossing and generalization strengths in a model and to train together on flat and deep patterns. The joint training is performed by omitting the gradients of both the shallow and deep parts at the same time during the back propagation, with minimal stochastic gradient landing optimization.

(**Ionu Brându oiu, Gavril Toderean, Horia Beleiu, 2016**) in this article, we have proposed an improved data mining method for predicting prepaid mobile. When evaluating the results from a technical point of view, we observe that to predict both the chain and the non-woven, models overall accuracy is 99.10% BN, 99.55% MLP and 99.70% SVM. In practice, the models have a very good performance of 99% to 100% and predict the brackets in the mobile phone company. Employee decision-making may involve different marketing techniques to obtain a kitchen appliance based on predictors that are more relevant for evaluating the performance of a model. These putty prediction models can also be used in other customer reaction patterns, such as: inter-sales, up-sell or customer purchases. The predictive efficiency of this data acquisition methodology can be further improved by effectively integrating the machine translation algorithms as described in the group's training structure. In addition, text extraction and social media acquisition techniques could be used along with the proposed data mining method to further reduce the exchange rate.

AUTHOR	METHODS	DATA SET
Sherif M Badr	classification, back propagation and multi-layer perceptron	Real data form IBM Watson analytics
Abhinash Mishra	Decision Tree, Naive Bayes, Support Vector Machines	From a web link http://www.ics.uci.edu~mlearn/MLRepository.html
Duyen DO	AdaBoost, Extra Trees, KNN, Neural Network and XGBoost.	The dataset with 98% non-churners and 2% churners, precision and recall of the model are 45.71% and 42.06%, respectively.
Sandra Mitrovic	RFM model, Baseline methods	Social network data
Dilip Singh Sisodia	KNN, LSVM, Naïve Bayes, Decision Tree, Random Forest	Employee details from Kaggle website
Ammar A Ahmed	ANN, Churn Prediction, KNN	Mobile industry
Paul Bertens	Social Games, churn prediction, survival analysis	Uses previous churn of mobile social games
Rong Zhang	Deep and Shallow Model, Joint Training, Machine Learning; Deep Learning; Insurance Churn Prediction Service	NEW CHINA LIFE INSURANCE COMPANY LTD
Ionu Brându oiu	Bayesian networks, call detail records, churn prediction, support vector machines, neural networks, principal component analysis, telecommunications	Asian mobile company prepaid customers data

Fig: 1 Churn Prediction Literature Comparison Chart

Machine learning as intelligent tool

Machine Learning (ML) is a category of algorithms that software applications can use to more accurately predict results but are not clearly programmed. The basic training of machine training is to develop algorithms that can receive input data and use statistical analysis to predict the output while updating the results as new data becomes available. Machine translation algorithms are often classified as supervised and unsupervised. Supervised algorithms require a data scientist or data analyst with machine translation capabilities to provide both input and output, and provide feedback on predictive accuracy during algorithm training. Unsupervised algorithms do not have to be trained with the desired output data. Instead, they use a reciprocal approach called deep learning to review data and draw conclusions. Continuous learning algorithms, also known as neural networks, use sophisticated processing tasks as supervised learning systems, including image recognition, speech-to-text, and natural speech generation. The complete approach how processing done is shown in figure 2.
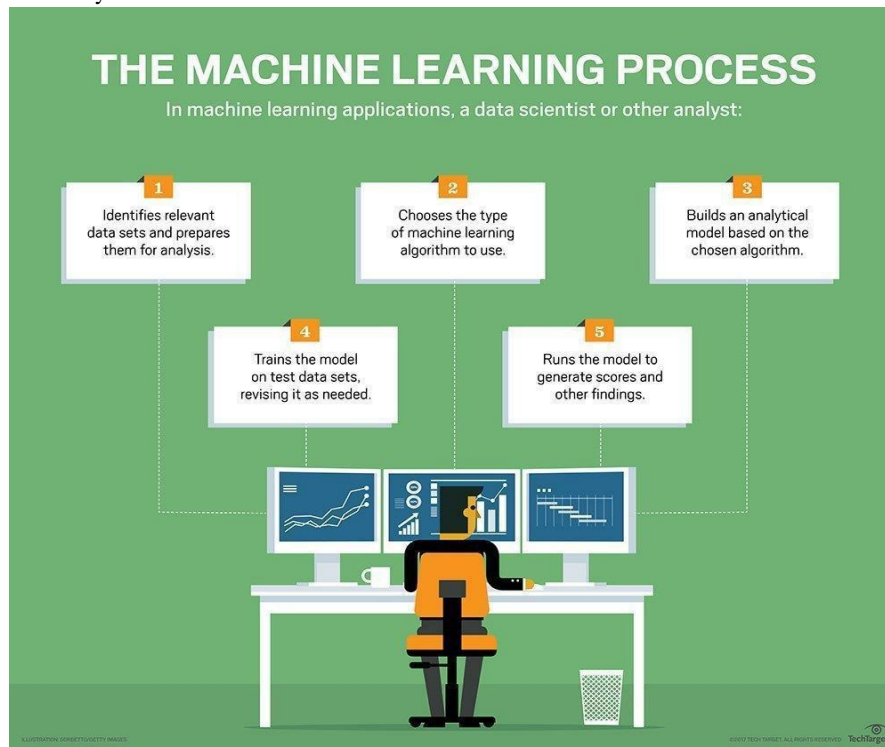

Fig 2: Process of Machine Learning

4. WAYS TO REDUCE CUSTOMER CHURN

The complete process and way to reduce Customer Churn is explained as below in sub points 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8 and as shown in Fig 3.
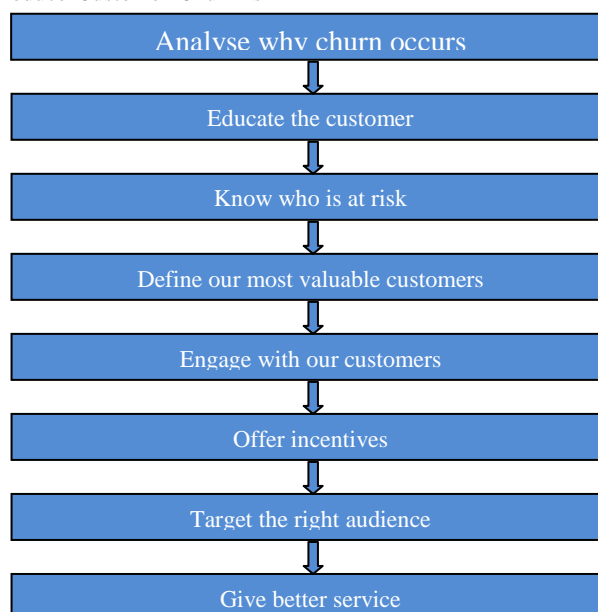

Fig: 3 Ways to Reduce the Customer Churn

4.1 Analyse why churn arises

Yes, this can be an obvious aspect, but let's emphasize again that we just have to understand why customers have decided to leave. The easiest way is to talk with the client. And, speaking, i mean really talking: the best choice is for our customers on the phone. In this way, we can show that we can really take care of it and find out immediately what has happened. Do not laziness by sending a customer survey, simply call them and find out why they're leaving. This gives us an instant feedback on whether our product solves client issues or causes implementation problems.

4.2 Educate the Customer

Of course, this trick, while avoiding false fraud, flows from top to bottom. We need to have enough good education or support materials to help increase customer loyalty and reduce bribery. We offer free training, seminars, video tutorials and product demonstrations to make our customers feel comfortable and informed. In other words, we not only need to provide them with tools that work, but they also provide training on how to use these tools with the most profit. In this way, we show the full potential of our products and services and ensure that our customers are on board.

4.3 Know who is at Risk

The best way to avoid emigration is to prevent it from the start, right? There is always a group of customers who are more likely to go away than others, so it's in our interest to find out what strikes this dangerous edge.

In this way, we can reach them in time to stay. It's easy to find those who are approaching the "risk group". Find out which customers do not communicate for some time. Did you ask for something like brochure, quotation or just more information and we forgot to continue it? If we all know it, we will be more active in avoiding the crisis. After analyzing the causes of emigration, we also learned about specific activities or the lack of our clients. This knowledge can help us find out if someone acting like this will soon leave our company.

4.4 Define our most Valuable Customers

As desperate as it sounds, we distinguish the most valuable customers from others and go a step further to make sure they get at least what they sign up for. Why? Well, let's be honest, these are the customers we usually want to keep. Particular attention should be paid to valuable customers as they bring the greatest turnover. The story of how we interact with customers can show how they are involved at every stage, whether they have product issues, and whether these issues are resolved. So what we can do is leave our customers in profitability groups, the desire to go and their chance to stay

positive about our offer. In this way, we can better predict customer turnover.

4.5 To Engage with our Customers

The second thing that helps us avoid delays is to involve our customers with our product. Provide our customers with the reasons to continue returning with the benefits of everyday products. In other words, we need to make ourselves (our products, services, offer, etc.) Part of our daily workflow for our customers. Commitment can be achieved by providing sufficient and clear content about the key functional benefits of our product and offering regular news such as announcements of offers, special offers or sales that are ahead of us. As the following IBM studies show, the increase in direct presence is the largest method of customer retention. Companies use site and social media channels to connect with their customers.

4.6 Offer Incentives

Another one is to offer incentives such as discounts and special offers for those who may have a deficit. But! Make sure we've properly judged whether offering incentives is good for you. This means that we need to make sure that the cost of our client engagement campaign is not higher than the customer profits we want to save. All in all, we should not spend a lot of money on customers who are unlikely to receive significant revenues.

4.7 Target Audience

Regardless of the complexity of all our detention tricks, they can get into the drain if they attract an incorrect audience. Here, I think that if our first interaction with the client is "free" and "cheap" then we risk to attract people who do not want our values. Such "freebies" collectors are unlikely to leave. It is better to focus on those who value the long-term value of products, and believes that investing in good quality is an asset. We should deal with them better.

4.8 Better Service

We waited for this type before, or not? Yes, this is the most obvious way to keep our customers on our side. In fact, poor customer service is a leading customer. According to Oracle's Client Experience Impact Report, the two main reasons why customers leave the company are incompetent and ruthless employees and unnecessarily slow service. According to the Forum Corporation survey, the worst service is 70%.

5. ANALYSIS AND RESULT

The complete analysis and result is shown in fig 4 and fig 5where we represent the prepaid market growth of data in US.
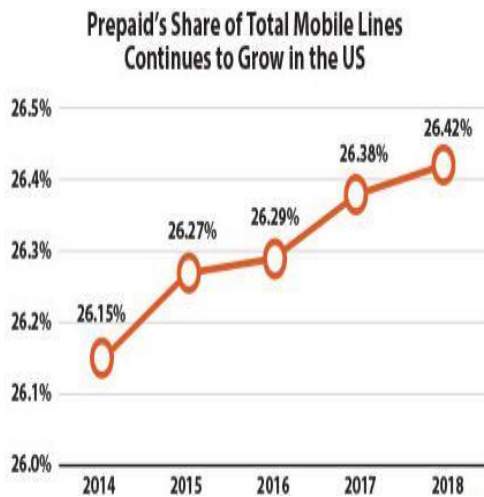
Fig 4: Prepaid Market Growth in percentage in US

Fig 5: Prepaid Market Growth in Data in US

According to a recent Euro Monitor International report, over the last five years, the use of prepaid mobile services has changed significantly in the United States. Although prepaid services in most Western European countries declined, and with low growth in countries such as Australia and the United Kingdom, the United States recorded a record high growth of 11%. Mobile prepaid sales amounted to 19% of the US market, so US operators need to continue, and significant growth opportunities for Yankee Group over the next 3-4 years, prepaid growth following the compound annual growth rate (CAGR) is over 12% projected for years. It is 3 percentage points faster than the entire telecommunications sector. It is more than 26% of the $ 27.5 billion total Yankee Group is expected to receive from the 2016 industry. Three major players are responsible for the transition to prepaid mobile devices:

- The economic downturn that US consumers have been comparing? Shop and move from long-term contracts.

- Improved positive attitude to prepaid mobile services due to improved smartphone and data transfer speeds

- Increased Internet usage for younger participants (12-20) who tend to look for cheaper plans

Mobile service providers must understand the reasons for the increase in prepaid payments to adapt their

6. CONCLUSIONS AND RECOMMENDATIONS

Mobile operators have to develop the policies by devising the strategies in the following areas

6.1 Network integration with banks

By integrating with banks, mobile operators reduce operating costs and add added value to customers, such as fees, activation, billing, etc.

6.2 Campaign management

Through design through strong awareness programs through road shows, innovative mass-market advertising with different systems, data and voice services, and campaign management, can be achieved by maintaining the desired customer base service offerings accordingly. You also need to understand what problems these managers are causing. Mobile Prepaid
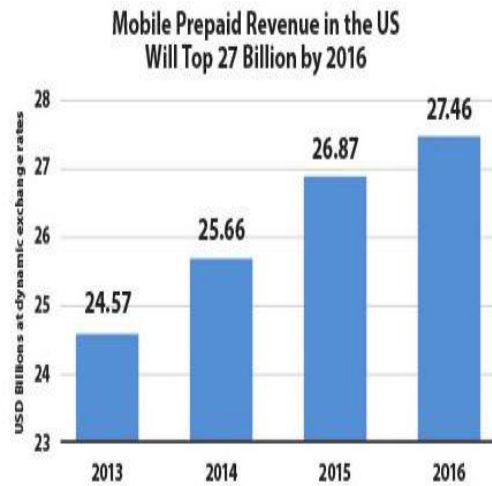
Growth Challenges In order to maintain the predicted high growth rate of prepaid plans, operators must develop proposals to address the following challenges:

- Low average revenue per user (ARPU). As previously mentioned, prepaid mobile customers are worth purchasers. Operators should check plans and prices to encourage the use and improvement of the ARPU.

- Lack of loyalty to high turnover - thanks to the easy availability of phone numbers and the growing contract without a contract number, customers have little reason to remain loyal to one operator. Price-sensitive customers can continually look for better deals and discounts from other service providers. Operators looking for demographics and the needs of specific customer segments can offer better service and pricing plans that boost loyalty.

- Premium pricing devices. Advantages of introducing prepaid rates are a major barrier to introducing prepaid services. It is therefore very necessary to look for new ways to reduce the cost of smartphones and offer attractive financing opportunities.

6.3 Mobile Analytics

Based on mobile analysis, a high-tech network with high signal strength should be developed to reduce the jar.

6.4 Participation management

Results oriented. Participatory customer initiatives need to be developed to develop various marketing elements.

6.5 Social Media Marketing

Acquaintance with the reach of social media marketing service providers should build awareness and route their advertising opportunities.

6.6 Affiliate Marketing

Direct advertising, online contests, multi-level marketing, with Internet services such as targeted real-time promotions without calling, SMS, MMS, closed user groups, announcing annual, monthly contracts with a special allowance, monthly free minutes, make a mobile wallet services community

Centre, Buzz marketing e-procedures, banks, subsidized data cards, e-evaluation mobile wallets, etc. This chip rate can be

reduced as most customers are Internet users.

6.7 Mobile Banking Opportunities

Would have a solution that goes beyond making and receiving payments, and allows the end-user to manage on-demand savings, and possibly use a wider range of products, secure store values and credit and insurance. Arup, M & A, regulation, licensing and technology coverage. Therefore, Indian mobile operators need to update information about this interface.

6.8 Concentrated Strategy

Development Hybrid Models Compatible with Smartphones, Tablet PCs, Socket Plugs, and enhance e-process, Sum of Differential Data, Speed, Service and Cost High-Tech Carriers, Cross-Border Creation Package Deals Offer subsidized mobile phones and fast new prepaid service configuration.

6.9 Different pricing strategies

Over voice and data services are developed, for example, for monthly quota services, combined with reduction of interest rates on the automatic call-back system of prepaid lines, SMS, MMS, roaming, data plans, etc.

6.10 Youth Marketing Strategies

It is necessary to attract this segment with compatible services for smartphones, tablets, sockets, campaign and event marketing, is offered concept, marketing has been developed and special privileges such as SMS, MMS, free ringtones, ringtones and the bags offer data services.

6.11 Region-Oriented Strategies

Designed based on any city or region to identify customer needs, such as network retail stores, expanding the ability of mobile banking, relating to the concept of Cooperative Marketing - Cross & Up sales.

7. REFERENCES

[1] Praveen et al., Churn Prediction in Telecom Industry Using R, International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Volume-3, Issue-5, May 2015.

[2] Chuanqi Wang, Ruiqi Li, Peng Wang, ZonghaiChen,"Partition costsensitive CART based on customer value for Telecom customer Churn Prediction, Control Conference (CCC), September 2017

[3] U.S. Commercial Service Vietnam, "Vietnam Market for Telecommunications Equipment and Services," The U.S. Commercial Service, 2014.

[4] Saravanan, M., Manoj, P., Smitha, G. B., & Lakshmi, V. (2015). Aatish-A New Profile-Based Recommendation Services for Mobile Telecom Network Subscribers. In Network Intelligence Conference (ENIC), 2015 Second European (pp. 160-164).

[5] R. Punnoose and P. Ajit, "Prediction of employee turnover in organizations using machine learning algorithms," Int. J. Adv. Res. Artif. Intell., vol. 5, no. 9, 2016.

[6] Ionuţ Brânduşoiu, Gavril Toderean, and Horia Beleiu "Methods for churn prediction in the pre-paid mobile telecommunications industry." In 2016 International Conference on Communications (COMM), pp. 97-100. IEEE, 2016.

[7] A. Perianez, A. Saas, A. Guitart, and C. Magne. Churn prediction in mobile social games: Towards a complete assessment using survival ensembles. In 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pages 564–573, Oct 2016.

[8] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," arXiv preprint arXiv: 1701.07875, 2017.

[9] R. Johnson and T. Zhang, "Effective use of word order for text categorization with convolutional neural networks," arXiv preprint arXiv: 1412.1058, 2014.

[10] T. Sumathi. "Churn Prediction on Huge Sparse Telecom Data Using Meta-heuristic." International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), vol.5, no.7, pp.574-577, 2016

[11] Reinartz, Werner J., and Vita Kumar. "The impact of customer relationship characteristics on profitable lifetime duration." Journal of marketing 67, no. 1 (2003): 77-99.

[12] Sohn, So Young, and Yoonseong Kim. "Searching customer patterns of mobile service using clustering and quantitative association rule." Expert systems with Applications 34, no. 2 (2008): 1070-1077.

[13] Thomassey, Sébastien, and Antonio Fiordaliso. "A hybrid sales forecasting system based on clustering and decision trees." Decision Support Systems 42, no. 1 (2014): 408-421.

[14] Wei, Chih-Ping, and I-Tang Chiu. "Turning telecommunications call details to churn prediction: a data mining approach." Expert systems with applications 23, no. 2 (2002): 103-112.