



## International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 10 • Number 31 • 2017

### Opinion Mining and Sentiment Analysis of Mobile Internet Usage

P.S. Rajeswari<sup>a</sup> K. Sadasivan<sup>b</sup> and Nithya preetha<sup>c</sup>

<sup>a</sup>Assistant Professor, Faculty of management, SRM University, Kattankulathur  
E-mail: rajeswari.PS@ktr.srmuniv.ac.in, psrajee\_2003@yahoo.co.in

<sup>b</sup>Associate Professor, Faculty of management, SRMUniversity, Kattankulathur

<sup>c</sup>Assistant Professor, Faculty of management, SRMUniversity, Kattankulathur

**Abstract:** India is the second largest in the online market proliferated with the subscriptions over 460 million internet users. By 2021, there will be about 635.8 million internet users in India. Successively the mobile internet usage has tremendously attained implausible growth and the internet penetration rate allures with respect prepaid segment when compared with post-paid segment. It is predominant to sense the factors influencing the mobile internet usage. This would facilitate the profitability rate since this market prevails with hectic competition and volatile as the customer churn is cardinal especially due to the internet offer wars. Henceforth it is primary to vigil and mine their opinions and do sentiment analysis of customer's preferences on reasons behind the mobile internet usage. Thus this study was mainly to examine the customer opinions and sentiments, to enhance the customer loyalty. The research study was descriptive and survey method was employed using structured questionnaires as sampling instrument. Sample size was 1102 and TamilNadu, one of the states of India was selected as the sampling framework. Neural networks were used to perform the sentiment analysis. From sentiment analysis, one could understand the factors influencing mobile internet usage Based on this study operational strategies were devised to improve customer loyalty.

**Keywords:** opinion mining, sentiment analysis, artificial intelligence, customer preferences.

#### 1. INTRODUCTION

##### 1.1. Back ground

The Indian Telecommunications Industry is one of the fastest proliferating sector in the World and India is projected to bench the second largest telecom market globally by 2017. Indicators are clearly revealing increased competition inducing the customers to hop for low cost options. This in turn entangled with disloyalty and as the industry saturates, it become imperative for the mobile operators to shift their focus from rapid acquisition strategies to strategies which helps to maintain and enhance margins from existing customer base.

##### 1.2. Impact of the problem

Though many service industries are affected by the churn phenomenon, the problem is extremely acute in the telecom industry, with customers joining and quitting in short periods. According to research firm Gartner, India's churn rate is anywhere between 3.5 percent to 6 percent per month, one of the highest in the Asia-Pacific region. Considering that the cost of acquiring a new customer is as high as Rs 3,000, the losses are immense.

### **1.3. Opinion Mining and sentiment analysis**

Current research is a process with focus on Indian cellular market by which data from customer behavior is aggregated and analysed to gain customer mind map, enabling each business to help make better and quicker business decisions. This information is used by the cellular businesses for direct marketing and customer relationship management. Customer analytics plays a very important role in predicting customer behavior and shaping future customer interactions.

Opinion Mining and sentiment analysis can provide the exact opinions and sentiments of customers with regard to the different bases of customer segmentation groupings; profitability analysis (which group of customers lead to the most profit over time); personalization (the ability to market to individual customers based on the data collected about them); event monitoring, what-if scenarios (how likely is a customer or customer category that bought one product to buy a similar one); and predictive modeling (for example, comparing various product development plans in terms of likely future success given the customer knowledge base). Data collection and analysis are viewed as a continuing and iterative process and ideally over time business decisions are refined based on feedback from earlier analysis and to make consequent decisions.

The efforts of product development, marketing, client services and other departments are fuelled by customer actions, or at least assumptions about the customer. Recent statistics depicts very high churn in this industry, is mainly rooted by the Youth segment. Hence this study of customer analytics is mainly focusing on Youth segment to lead not only to better and more productive customer relations in terms of sales and service but also to improvement in supply chain management (lower inventory and speedier delivery) and thus lower costs and more competitive pricing.

It facilitates to assess the Customer profitability index and Customer lifecycle; Customer analytics enables an operator to gain a better understanding of the variables that influence customer churn. It enables the Telco to understand which customer is likely to leave and why, which in turn can help the company take the necessary measures to counter it.

Extract of customer analytics provides the telecom company with a sliced and diced view of the customer base, thereby empowering it to treat each customer differently as per needs. The customer attributes typically considered in a churn analysis can be broadly categorised into customer demographics, contractual data, technical quality data, billing and usage data and events-type data. But the most commonly used historic variables include the time a customer spends on air, the number of calls he makes and the revenue generated from that customer.

The predictive information becomes crucial as it gives the service provider a window to proactively fix the glitches in service and contain churn, thereby improving bottom lines. The solution also helps identify cross-sell and up-sell opportunities, which can have a further positive impact on the operator's bottom line.

### **1.4. Subject Area**

Opinion mining and sentiment analysis was performed on customer profile, level of satisfaction, customer loyalty and their buying behavior with regard to the mobile internet usage.

## **2. RESEARCH METHODOLOGY**

### **2.1. Research problem**

Even though Indian mobile telecommunications marked tremendous growth, it is facing very high churn and disloyalty in the market. It is highly critical to analyze the factors influencing customer loyalty and their level of satisfaction on the cellular providers.

## **2.2. Purpose**

The purpose of this research is to perform Opinion Mining and sentiment analysis by identifying the operational factors that are influencing customer buying behavior, level of satisfaction and loyalty with regard to Indian Mobile telecommunications.

## **2.3. Objective**

Recent trend line shows the growth and prospects of youth marketing and especially in the field of Telecommunications, they are marking tremendous development. At the same time indicators are highlighting very high churn rate in this segment. Hence this Research focuses on Youth segment and Research objectives are listed below.

1. To perform opinion mining and sentiment analysis with respect to Indian mobile telecommunications.
2. To develop the conceptual model on customer churn to examine the behavioral constructs about the mobile service provider.

## **2.4. Research Design**

“Exploratory” type of research design is adopted for this study.

## **2.5. Data collection**

The research work is in need of firsthand information. Primary and Secondary data are collected for this survey.

## **2.6. Primary Data**

Survey method was adopted for collecting the primary data. As mentioned above in the research objective, youth segment falling under the age group of 12 to 34 are selected as the respondents for the research. Questionnaire was structured using the objective set for this study.

## **2.7. Secondary Data**

The secondary data had been collected from the previous Research findings, scholarly reports, telecommunication reports, respective marketing departments and through the different sources of literature such as journals, articles etc.

## **2.8. Sampling plan**

Simple Random sampling method was adopted.

## **2.9. Sample size**

The sample consists of 1102 respondents in the age group of 12 to 34.

# **3. RELIABILITY AND VALIDITY**

## **3.1. Reliability**

Reliability is the consistency of a measurement. Questionnaire had been tested with 30 respondents to check the reliability. With SPSS package, reliability was tested. Cronbach alpha value was 0.782 with no exclusions implying the acceptability of the questionnaire for the research work.

### 3.2. Validity

Internal and external validity were checked with the respective sources and every hypothesis is represented by a question in the questionnaire so that they can be tested and measured.

### 3.3. Pilot study

Pilot study was conducted with 30 respondents.

## 4. DATA ANALYSIS

### 4.1. Neural Networks

#### 4.1.1. Artificial Neural Network Model

Neural networks are nonlinear data driven, self-adaptive approach and as powerful tool for modeling. It imitates human brain and can process complex data. Multilayer perceptron on (MLP) and radial basis function (RBF) networks are supervised in the sense that the model predicted results can be compared against known values of the target variables. The parameters of the networks should be found by applying optimization methods. The optimization is done with respect to the approximation error measure. SPSS neural networks option allows fitting MLP and RBF networks and saving the resulting models for scoring. Multilayer perceptron networks are used when parameters are non linear and radial basis function networks are used if the parameters are linear. The results show that the MLP networks are superior in memory usage and classification time. The procedure adopted in this study was Multilayer Perceptron network since the parameters used for modeling were non linear.

#### 4.1.2. Multilayer Perceptron Network Model

The multilayer perceptron trained with back propagation algorithm is able to perform a general non linear output –input mapping from the dimension of input space to the dimension of output space. It produces a predictive model for one or more dependent (target) variables based on the values of the predictor variables. The dependent variables can be nominal, ordinal or scale. Multilayer perceptron procedure can find more complex relationships. If MLP network allows a second hidden layer, each unit of the second hidden layer is a function of the units in the first hidden layer, and each response is a function of the units in the second hidden layer. In this study two hidden layers for modeling using multilayer perceptron approach was used.

#### 4.1.3. Model Output Using Testing Sample

The factors from the output of exploratory factor analysis and MANOVA are fed as the factors for developing the base model. The factors of churn were taken as dependent factors, the factors of customer satisfaction were considered as Independent factors and city was taken as a covariate. Several trials were made by taking all possible covariates subjected for this study, and cities were seemed to be the significant co-variant. The estimated results of the base model are given in the following tables

**Table 1**  
Case processing summary

<i>Sample</i>	<i>N</i>	<i>Percent</i>
Training	773	70.3%
Testing	327	29.7%
Valid	1100	
Excluded	2	
Total	1102	

The case processing summary given in the above table shows that 773 samples were assigned to the training sample, 327 to the testing sample and 2 of the cases were excluded from the analysis.

**Table 2**  
**Network Information**

<i>Layers</i>	<i>Particulars</i>	<i>Factors</i>
Input Layers	Independent variables	Number of Mobile phones usage(X1) Design of the Mobile Phone(X2) Video(X3) Type of the datacard(X4) Place of purchase(X5) Awareness(X6) Personal Use(X7) Alternative choice due to non-availability(X8) Tariff(X9) Changing the brand(X10) Net Speed(X11) Signal strength(X12) Good Impression on the corporate Image(X13) Recharge Vouchers(X14) Tariff Rates(X15) Internet Services(X16) Customer Care(X17) Time taken for Complaint Resolution(X18) Welfare to the society(X19) Pulsing/Timing(X20) Quality of Coverage(X21) Easiness in Subscription(X22) Recharge Process(X23) Functional Product(X24) Quick customer complaint redressed(X25) Application Process(X26) Reach of Customer services(X27) Readiness of Customer care(X28) Availability and Easiness of services(X29) Front End services(X30)

<i>Layers</i>	<i>Particulars</i>	<i>Factors</i>
		Customer Services(X31)
		Call centres(X32)
		Personalization(X33)
	Covariate	District
	Number of Units	186
	Rescaling Method for Covariates	Standardized
Hidden layers	Number of Hidden Layers	1
	Number of Units in Hidden Layer 1	12
	Activation Function	Hyperbolic tangent
Output Layers	Dependent Variables	Social Media(Y1)
		Internet Facility(Y2)
		Moral/Ethical Values(Y3)
		Regulatory certainty of operations(Y4)
	Number of Units	4
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

The network information in the above table displays information about the neural network and is useful for ensuring that the specifications are correct. The number of units in the input layer is the number of covariates plus the total number of factor levels. One hidden layer was requested and the procedure has chosen 12 units in the hidden layer. A separate output unit was created for the scale- dependent variable. They were rescaled by the standardized method, which required the use of the hyperbolic tangent activation function for the output layer. Sum-of-squares error was reported because the dependent variables were scaled.

**Table 3**  
**Model summary**

<i>Training Sum of Squares Error</i>	<i>30.14</i>
Average Overall Relative Error	.020
Stopping Rule Used	1 consecutive step with no decrease in error
Training Time	0:00:04.53
Testing Sum of Squares Error	10.17
Average Overall Relative Error	.034

The model summary in above table displays information about the results of training and applying the final network to the holdout sample. Sum-of-squares error was displayed because the output layer has scale dependent variables. This was the error function that the network tries to minimize during training. Note that the sums of squares and all following error values were computed for the rescaled values of the dependent variables. The relative error for each scale-dependent variable was the ratio of the sum-of squares error for the dependent variable to the sum-of-squares error for the “null” model, in which the mean value of the dependent variable was used as the predicted value for each case. The relative errors were fairly constant across the training, testing, and hold out samples, which gives confidence that the model was not over-trained and that the error in future cases, scored by the network would be closed to the error reported in this table. The estimation algorithm stopped because the error did not decrease after a step in the algorithm.

**Table 4**  
**Independent variable importance**

<i>Factors</i>	<i>Importance</i>	<i>Normalized Importance</i>
Number of Mobile phones usage(X1)	0.075	100.0%
Design of the Mobile Phone(X2)	0.050	66.2%
Video(X3)	0.018	23.5%
Type of the data card(X4)	0.025	32.5%
Place of purchase(X5)	0.041	54.7%
Awareness(X6)	0.045	59.8%
Personal Use(X7)	0.040	53.1%
Alternative choice due to non-availability(X8)	0.027	36.3%
Tariff(X9)	0.028	36.6%
Changing the brand(X10)	0.033	43.7%
Net Speed(X11)	0.035	47.1%
Signal strength(X12)	0.049	64.6%
Good Impression on the corporate Image(X13)	0.021	28.1%
Recharge Vouchers(X14)	0.029	39.0%
Tariff Rates(X15)	0.021	27.2%
Internet Services(X16)	0.039	52.1%
Customer Care(X17)	0.021	28.3%
Time taken for Complaint Resolution(X18)	0.020	27.1%
Welfare to the society(X19)	0.026	34.4%
Pulsing/Timing(X20)	0.028	37.7%
Quality of Coverage(X21)	0.058	77.0%
Easiness in Subscription(X22)	0.027	36.4%
Recharge Process(X23)	0.026	34.8%
Functional Product(X24)	0.021	27.6%
Quick customer complaint redressed(X25)	0.036	47.5%
Application Process(X26)	0.023	30.0%
Reach of Customer services(X27)	0.019	25.5%
Readiness of Customer care(X28)	0.027	35.4%
Availability and Easiness of services(X29)	0.025	32.6%
Front End services(X30)	0.036	47.5%
Customer Services(X31)	0.025	33.7%
Call centres(X32)	0.004	5.9%
Personalization(X33)	0.075	100.0%
City	0.050	66.2%

The importance of independent variables were listed as per the simulated pattern of brain functioning of the customers using the synaptic weights of each neurons. Hence the customers were giving importance to the mobile phone usage and purchase of new phones with advanced technology and designs. They were giving priority to internet services and personalization –e services. In connection to this they expect wider coverage of network with high quality and signal strengths. Place of purchase also plays significant factor for the customer. They seek good image of the corporate and awareness over various schemes and services. As they are main churning factors for the customer, the mobile operators should focus on compatible mobile phone based services.

They can revamp their conventional mode of networking and avail the technology transfer to implement robust technology in network coverage and signal strength. Since majority of the customers were using data services, they could increase the options in schemes.

Mobile Operators could also allocate budget in raising their level of awareness about their data schemes through rigor market campaigning. Even though mobile operators were highly competitive in making price wars they failed to fill the need gaps of the customers. From the above analysis customers were highly confined to Internet usage. Bharti Airtel, Reliance and other players were very conscious in reducing price tariffs especially in voice calls and not very much on data packages. Hence they had to classify and devise the need based data schemes.

#### 4.1.4. Suggestions

Opinion mining and sentiment analysis thus indicating the path way for the Indian Mobile for increasing the customer loyalty and to turn around the churn rate, For that Indian mobile should:

1. Keep its good performance on quality of phone call, coverage, quality of SMS and pay attention on these three areas because they are in the first level, which means they have great importance on customer loyalty
2. Improve its performance on customer service, advertisements about corporate image, inquiring phone fee system and corporate social responsibility and pay more attention on customer service, because it is in the first level.
3. Improve its performance on rating price of given quality significantly and pay great attention on this factor.

## 5. CONCLUSION

The telecom industry, especially the mobile industry of India is undergoing a transformation and the number portability is bringing about imperatives worthy enough to carry out high-end research. This study is one such attempt to enhance the exposure on Opinion mining and sentiment analysis and it is also expected to facilitate the marketers to design the essential operational parameters for scheming the retention strategies and to enhance Customer Experience management.

## REFERENCES

- [1] ACSFI: Fornell, C. Johnson, M. D., 1996. The American customer satisfaction index: Nature, purpose, and findings, *Journal of Marketing*, 60. 7-19
- [2] Ahn, J. H., Han, S. P., Lee Y. S., 2006. Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry, *Telecommunications Policy*, 30, 552-569.
- [3] Amine, A., 1998. Consumers' true brand loyalty: the central role of commitment, *Journal of Strategic Marketing*. 6. 305-319



- [4] Anderson, E. W. Sullivan, M. W. 1993. The antecedent and consequences of customer satisfaction for firms. *Marketing Science*, 12. 125-144.
- [5] Andreasen, A. Drumwright, M. 2001. In: Andreasen, A. (Ed.), *Alliances and Ethics in Social Marketing in Ethical Issues in Social Marketing*. Washington, DC: Georgetown University Press,
- [6] Andres K. 2007, Affecting customer loyalty: do different factors have various influences in different loyalty levels?, *University of Tartu - Faculty of Economics & Business Administration Working Paper Series*. 58. 3-30.
- [7] Aydin, S. Ozer, G. Arasil, O. 2005, Customer loyalty and the effect of switching costs as a moderator variable: A case in the Turkish mobile phone market. *Marketing Intelligence & Planning*. 23. 89-103.
- [8] Bandyopadhyay, S. Martell, M. 2007. Does attitudinal loyalty influence behavioral loyalty. *Journal of Retailing & Consumer Services*. 14. 35-45.
- [9] Barich, H. Kotler, P. 1991, A framework for marketing image management, *Sloan Management Review*. 32.94-104.
- [10] Bitner, M. J. Hubbert, A. R. 1994. *Encounter Satisfaction versus Overall Satisfaction versus Quality, Service Quality New Directions in Theory and Practices*. London: Sage Publications. 72-95.
- [11] Blodgett, J.G. Hill, D.J. Tax, S.S. 1997, The effects of distributive, procedural, and interactional justice on post-complaint behavior, *Journal of Retailing*. 31. 185-210.
- [12] Bodet, G. 2008. Customer satisfaction and loyalty in service: Two concepts, four constructs, several relationships. *Journal of Retailing & Consumer Services*. 15. 156-163.
- [13] Brady, M.K. Cronin, J.J. Brand, R.R. 2002. Performance-only measurement of service quality: a replication and extension. *Journal of Business Research*. 55. 17-31.
- [14] Brown, T.J. Dacin, P.A. 1997. The company and the product: corporate associations and consumer product responses. *Journal of Marketing*, 6. 68–84.
- [15] Brunner T. A. Stöcklin, M. Opwis, K. 2008. Satisfaction, image and loyalty: new versus experienced customers. *European journal of marketing*. 42. 1095-1105.
- [16] Burnham, T. A., Frels, J. K., Mahajan, V., 2003. Consumer switching costs: a typology antecedents, and consequences, *Journal of the Academy of Marketing Science*, 31, 109-126.
- [17] Butcher, K. Sparks, B. O’Callaghan, F. 2001, Evaluative and relational influences on service loyalty. *International Journal of Service Industry Management*. 12. 310-328.
- [18] Coussement K and Van den Poel D (2008), Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques, *Expert Systems with Application*, Vol. 34, no. 1, pp. 313-327.
- [19] Donio, J., P. Massari, and G. Passiante, (2006). Customer Satisfaction and Loyalty in a Digital Environment: An Empirical Test,” *Journal of Consumer Marketing*, 23(7), 445-457.
- [20] Drucker, P. F. (1973), *Management: Tasks, Responsibilities, Practices*, Harper and Row, New York. Etzel, Walker, and Stanton (2004), *Marketing*, 13th Ed., Tata McGraw Hill, New Delhi.
- [21] Fildes, R. (2002), Telecommunications demand - A review, *International Journal of Forecasting*, Vol. 18, Pp. 489-522.
- [22] Gefen, D. (2002), Customer loyalty in e-commerce, *Journal of Association for Information Systems*, (3), 27–51.
- [23] Geppert, C. (2002), *Customer Churn Management: Retaining High-Margin Customers with Customer Relationship Management Techniques*, KPMG & Associates.
- [24] Grönroos, C. (2000), *Service Management and Marketing: A Customer Relationship Management Approach*, Wiley, New York, NY.
- [25] Hutt, Michael D. and Thomas W. Speh (2004). *Business Marketing Management: A Strategic View of Industrial and Organizational Markets*, 8th ed., Mason, OH: South-Western.
- [26] Johnson, W. C. and A. Sirikit (2002). *Service Quality in the Thai Telecommunication Industry: A Tool for Achieving a Sustainable Competitive Advantage*, *Mgt. Decision*, Vol.40, No. 7.

- [27] Kobsa, A. (2002). Personalized hypermedia and international privacy, communications of the ACM, Vol.45 pp64-7.
- [28] Kotler, P. (2007). Marketing Management. Analysis, Planning, Implementation, and Control, 9th ed., Prentice-Hall International, Englewood Cliffs, NJ.
- [29] Kotler et al. (2002). Understanding the relationships among brands, customers and resellers, Journal of the Academy of Marketing Science, Vol. 28, No. 1, pp. 17-23.
- [30] Parasuraman, A., V.A. Zeithaml and L.L. Berry (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for future research , Journal of Marketing, Vol.58, pp.111-124.
- [31] Perreault, W.D. Jr., McCarthy,E.J. (2002).Basic Marketing, McGraw-Hill/Irwin, Boston Reinartz, W.J., V. Kumar (2003). The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration, Journal of Marketing, Vol. 67 No.1, pp.77-99.
- [32] Rigby, D.K., Reichheld, F.F., and Schefter P. (2002). Avoid the four perils of CRM, Harvard Business Review 80 (2): 101–109.
- [33] Thomassey S and Fiordaliso A (2006). A hybrid sales forecasting system based on clustering and decision trees,Decision Support Systems, vol. 42, no. 1, pp. 408-421.
- [34] Wei, C. and Chiu, I. (2002). Turning telecommunications call details to churn prediction: A data mining approach, Expert Systems with Applications, 23(2) pp.103-112.
- [35] Zeithaml, V.A. and M. J. Bitner (2000). Services Marketing: Integrating Customer Focus across the Firm,2nd ed., McGraw Hill, New York, NY.