

Examining Churn and Loyalty Using Support Vector Machine

Ali Dehghan

Wesleyan College, 4760 Forsyth Road, Macon, GA 31210, USA

E-mail: adehghan@wesleyancollege.edu

Theodore B. Trafalis

The University of Oklahoma, 202 West Boyd, Norman OK 73019, USA

E-mail: ttrafalis@ou.edu

Received: October 16, 2012

Accepted: December 4, 2012

Online Published: December 5, 2012

doi:10.5430/bmr.v1n4p153

URL: <http://dx.doi.org/10.5430/bmr.v1n4p153>

Abstract

The present study attempts to contribute to the knowledge of how customer loyalty could be assessed, using support vector machine. Additionally, the relationship between customer loyalty rate and churn rate is studied. After literature review and elaborating the concepts of customer loyalty, churn and support vector machine, a comprehensive data analysis on a telecommunication company's customers' dataset, utilizing support vector machine, has been developed. The proposed dataset consists of 1300 customers of a telecommunication company. The results imply that the proposed data analysis tool could be used for recognizing and determining the strategies leading to higher customer loyalty and lower churn.

Keywords: Support Vector Machine, Customer Loyalty, Churn

1. Introduction

Telecommunication industry is among the most rapid growing sectors among high technology industries. Various telecommunication companies offer a wide range of products, plans and services. All these service providers compete tightly to firstly retain their existing customers and secondly, attain new customers. Variety of telecommunication services and products with competitive prices make customer loyalty and churn in the middle companies' attentions. The more customer loyalty, the less churn rate. Thus, preventing customer churn is a significant task for telecommunications and finances companies (Tsai & Lu, 2009). The purpose of this paper is twofold: first, definitions of customer loyalty, customer satisfaction and churn are reviewed and their relationships are practically explained. Second, the applications of Support Vector Machine in assessing the customer loyalty are studied.

2. Literature Review

Satisfaction occurs when someone feels he has achieved his goals (Sheldon and Elliot, 1999). Customer satisfaction is a critical success factor for traditional or online business system (Ho and Wu 1999), and is a famous and established term in different sciences. Expectations and experienced service performance are two factors that impact customer satisfaction (Shahin, 2006; Dehghan and Shahin, 2011). Perceived performance is impacted by a) the customer's perception of service quality, b) marketing mix, and brand name and c) Image of the company (Andreassen and Lindestad, 1998a). Several researchers believe that satisfaction attitude is formed by customers when the performance they received from the products surpasses their pre-purchase expectations (Oliver, 1980). Some of the well-known definitions of customer satisfaction are:

- 1) Kotler (2000) states "Satisfaction is a person's feelings of pleasure or disappointment resulting from compared a product's perceived performance (or outcome) in relation to his or her expectations" (p. 36).
- 2) Anderson, Fornell and Lehman (1990) determine customer satisfaction as "a post choice evaluative judgment of a specific purchase occasion" (p. 54).
- 3) Athanasopoulos (2000) believes "Customer satisfaction is based, conceptually, on the amalgamation of service quality attributes with such attributes as price and convenience" (p. 192).

When product performance is less than expected performance, negative disconfirmation and respectively dissatisfaction occur (Chen-Yu, Williams and Kincade, 2009). Customer expectation is described as a customer's pretrial beliefs about a product (Mckinney, Yoon, and Zahedi, 2002). Customers' predictions about what may happen during an impending transaction or exchange are expectations (Zeithaml and Berry, 1988). Perceived performance is viewed as a customer's perception of how a product's performance meets their needs, wants, and desires (Cadotte, Woodruff and Jenkins, 1987). Perceived quality is the customer's judgment about an organization's excellence or superiority (Zeithaml, 1988). In addition, disconfirmation can be defined as customers' judgments from comparing their expectations and their perceptions of received performance (McKinney, Yoon, and Zahedi, 2002; Spreng, Mackenzie, and Olshavky, 1996).

Churchill and Surprenant (1982) believe that disconfirmation theory is the primary foundation for satisfaction models. Discrepancy between perceived performance and expectation determines satisfaction (Khalifa and Liu 2003). Customers' beliefs about a product are called expectation (Mckinney, Yoon and Zahedi 2002). Perceived performance is how a customer judges an organization's overall excellence (Zeithaml 1988). Disconfirmation is a consumer's perception obtained from comparing their expectations and their perceptions of received performance (Mckinney et al., 2002).

According to Fang, Chiu and Wang (2011), "Customer satisfaction is particularly important to the success of stores as it is posited as a major driver of post-purchase phenomena, such as repurchase intentions" (p. 480). Oliver (1999) has defined customer loyalty as "a deeply held commitment to rebuy or repatronize a preferred product/service consistently in the future, thereby causing repetitive same-brand or same brand-set purchasing, despite situational influences and marketing efforts having the potential to cause switching behavior" (p. 34). Customer loyalty is considered important since it impacts long-term profitability positively (Ribbink, Van Riel, Liljander and Streukens, 2004) and customer retention is generally very related to the long-term growth of a company (Eid, 2011).

Losing a customer either definitely or partially (re-buying rate reduction) will impact sales volume negatively, meaning more marketing activities are needed to keep them attracted (Zins, 2001). The relationship between a customer and a seller after the first transaction is defined as customer loyalty (Hallowell, 1996; Dehghan, 2006). Kuehn (1962) believes loyalty is the probability of product repurchase. Loyal customers are the best ones, because they are less costly to serve, they usually pay more than other customers and most likely they act as marketers for your company through word-of-mouth (Reinartz and Kumar, 2002). Undoubtedly, loyalty becomes a crucial construct in the burgeoning field of customer relationship management (Ball, Coelho, and Machas, 2004; Soderlund, 2006). Loyal customers are active ambassadors for any businesses. Existing customers are less price-sensitive with less maintenance costs compared with new ones (Lee-Kelly, Davis, and Kangis, 2002). Brands, products, or services can invoke loyalty, creating positive attitudes and behaviors in repeat patronage, additional purchases, and recommendations (Rowley, 2005). This form of loyalty can influence actual and potential customers. Additionally, a base of loyal customers increases the positive feedback of the organization's products and services. Further, they are a more valuable asset, thus reducing the need to seek new customers.

Dehghan (2006) states "Existing customers tend to buy additional services, provide favorable recommendations, have lower maintenance needs and prefer service over pricing unlike new customers' needs" (p. 30). Customer loyalty programs pursue two aims. increase sales revenues by raising the levels of purchase/usage and build a close bond between existing customers and the brand. Achieving these aims will lead to profit increased (Uncles, Dowling, and Hammond, 2003; Dehghan and Shahin, 2011). Dehghan and Shahin (2011) argue, "Those consumers that demonstrate the greatest levels of loyalty toward the product or service activity tend to repurchase more often, and spend more money" (p. 3). According to Reinartz and Kumar (2002), states:

Many advocates of loyalty initiatives argue that loyal customers pay their way because the up-front costs of acquiring them are amortized over a large number of transactions. But, of course, that argument presupposes that the customers are profitable in those transactions. A more plausible argument for the link between loyalty and decreased costs can be built on the idea that loyal customers will be more familiar with a company's transaction processes. Since they need less hand-holding, the company should find it cheaper to deal with them (p. 5).

Customer satisfaction cannot be assessed directly using an objective measure. Hallowell (1996) proposes that customer satisfaction impacts customer loyalty, which in turn influences profitability. Service providers benefit from the connection a loyal customer has, while a satisfied customer has a passive connection (McGarry, 1995). Customer satisfaction can mediate the relationship between perceived quality and customer loyalty (Hsu, 2008).

According to a Bowen and Chen (2001) study, customer satisfaction does not equal customer loyalty. The findings of Tecepi (1999) and Bowen and Chen (2001) verified the nonlinear and asymmetric relationship between customer

satisfaction and customer loyalty. (See Figure 1). There are several ways to assess customer loyalty. Some of the famous academic surveys can be found in Table 1.

3. Relationship between Customer Loyalty and Customer Satisfaction

Customer satisfaction is defined as a result of an effective evaluation, where comparison is made between the expected and perceived performance (Dehghan, 2006). If the perceived performance is less than expected, customers will be dissatisfied. However, when the performance surpasses the expectations, customers will be satisfied (Lin, 2003). Customers may also be loyal because they are satisfied and thus want to continue the relationship (Dehghan and Shahin, 2011). Customer satisfaction can mediate the relationship between perceived quality and customer loyalty (Hsu, 2008). Customer loyalty "prevents customer churn and consolidates retention, thereby constituting an important cause of customer loyalty (Fornell, 1992; Reichheld, 1996).

According to a Bowen and Chen (2001) study, customer satisfaction does not equal customer loyalty. The findings of Tecepi (1999) and Bowen and Chen (2001) verified the nonlinear and asymmetric relationship between customer satisfaction and customer loyalty. (See figure 1)

4. Churn

"Churn management is the term that has been adopted to define customer turnover" (Hadden, Tiwari, Roy and Ruta, 2007). Several reasons are associated with customer churn, for instance, when a customer finds out that a competitor offers some new products which their current supplier is unable to provide. There could be economic reasons like offering the same product with a better price (Kim, Park and Jeong, 2004). Churn rate measures the number of things moving out and in during a specific period of time. Although, this term is used in many contexts, it is vastly used in business showing the number of customers leaves the company or stay in. The churn rate is broadly used in telecommunication industry. Customer movement from one provider to another in telecommunication industry is called customer churn and the operator's process to retain profitable customers counted as churn management (Berson, Smith & Thearling, 2000).

Many researchers argue about the relationship between churn rate and customer loyalty rate. Customers with a high level of satisfaction are likely to remain with their existing providers. Such loyal customers hardly switch to another service provider, in other word, they do not churn.

Hwang, Jung and Suh (2004) concluded with following equation:

$$\text{Customer Loyalty} = 1 - \text{Churn Rate}$$

Since customer loyalty generally refers to a customer's willingness to continue his/her relationship with a certain company or product/service, the decision not to churn from the current mobile operator, dependent on the length of time, could be equated to loyalty to that company. It is an undeniable fact that forecasting the customers' behaviors is one of the service providers' major concerns and an accurate prediction may lead to a set of strategic and tactic retention measures to reduce the customers' movement (Lariviere & Van den Poel, 2004; Hung, Yen & Wang, 2006).

5. Support Vector Machines

Support vector machines are a set of supervised learning methods which are being used for two folds: classification and regression analysis. The support vector machines can separates the data points into two classes. Generally, a support vector machine constructs one or more hyper-planes by mapping the data points into higher dimensional space to make them linearly separated. According to Lee & Stolfo (1998) "Data mining generally refers to the process of (automatically) extracting models from large stores of data. The recent rapid development in data mining has made available a wide variety of algorithms, drawn from the fields of statistics, pattern recognition, machine learning, and database" (p.2).

One of the major applications of Data mining is classification. Classification is the problem of identifying the sub-population to which new observations belong, where the identity of the sub-population is unknown, on the basis of a training set of data containing observations whose sub-population is known (n.d). According to Shaw, Subramaniam, Tan & Welge (2001), "A systematic methodology that uses data mining and knowledge management techniques is proposed to manage the marketing knowledge and support marketing decisions. This methodology can be the basis for enhancing customer relationship management (p.2).

6. Discussion

In this research, a telecommunication company's customers are classified based on several attributers like : State, Account Length, Int'l Plan, V-Mail, Plan, V-Mail Message, Day Minutes, Day Calls, Day Charge, Eve Minutes, Eve

Calls, Eve Charge, Night Minutes, Night Calls, Night Charge, Intl Minutes, Intl Calls, Intl Charge, Customer Service Calls, Churn. The dataset is obtained from “Discovering Knowledge in Data: An Introduction to Data Mining” book. The classified customers help the companies in different ways. For instance, they can allocate adequate marketing budget, they can prioritize the customer preferences meaning which of these attributes have a higher degree of significance. In short, this data analysis will lead to decision making and will help the companies to make the best decision in terms of all aspects of their business.

The original data set contained many categorical dataset or some of them had non-numerical data type. In order to use this data, the researcher had to encode them into integer data type. In the case of states this was not a simple. It requires the state’s attributes to be expanded to fifth. So instead of the state being ‘OK’ it would be either a 1, value meaning the data point fulfills that attribute or a zero, meaning that data point does not fulfill that attribute. This encoding enhanced our data into 68 attributes.

MATLAB version 7.14 was used for the data analysis. Several lines of codes were written to do the data analysis. In order to find the most significant attributes, PCA or principle component analysis was used. As a result, the authors determined that State, Account Length, Int'l Plan, V-Mail, Plan, V-Mail Message, Day Minutes and Day Calls accounted for 98% percent of the variation in the data (figure 2). This means that although, the data spans 68 dimensions, the data looks like an 8-D plane (63% of the datum’s variation can be visualized in figure 2). (See Figure 2 and 3)

7. Prediction Results

To achieve the resulting accuracy, 50 data points were removed from the training data and were put in the testing data. The relevant tests were ran 100 times with different sets of 50 data points in order to find the highest accuracy. The following figure 4 depicts the plot of each of the 100 test and their corresponding accuracies. The optimum median accuracies for class +1 and -1 were 57.16 and 57.18 percent respectively. (See Figure 4)

To optimize the results, the parameters and the C value were tuned in order to improve the accuracy. Since the data was clustered, a clean separating surface is not attainable. The method for choosing parameters, was to choose parameters of a small degree in order to have a parabola which shatters the two clouds. Regarding the trade-off parameter C, it has to remain low in order to account for the outliers that exist in the data set. In figure 4, the accuracies that correspond to these parameter changes can be seen. (See Table 2)

8. Conclusion

The accuracies led us to believe that the data cannot be separated with a high accuracy, however it can be separated. The data was clustered in groups which made the polynomial separation less accurate. Another barrier in our problem was the fact that the data had primarily one class. Approximately 85% of the data was of one class. Scaling the data could lead to a better separation because it would balance the amount of data points in each class. The balance would show the traits of the data more clearly leading to higher separation accuracies.

As was stated, to retain the profitable customers, the service providers should be able to forecast the churners by taking appropriate marketing actions. In this research, support vector machine was used to classify the customer data and make an accurate churn prediction. Using support vector machine may give an interesting opportunity to service providers in other business fields since it has an unbelievable classification capabilities which will eventually lead to more accurate churn prediction.

References

- Aaker, D.A. (1991a). *Brand Equity*. New York, NY: The Free Press.
- Anderson, E. W., and Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction for firms. *Marketing Science*, 12, 125-143. <http://dx.doi.org/10.1287/mksc.12.2.125>
- Ball, D., Coelho, P.S. & Machas, A. (2004). The role of communication and trust in explaining customer loyalty: An extension to the ECSI model. *European Journal of Marketing*, 38(9/10), 1272-1293. <http://dx.doi.org/10.1108/03090560410548979>
- Bowen, J. T., & Chen, S.-L. (n.d.). The relationship between customer loyalty and customer satisfaction.
- Dehghan, A. (2006). *Relationship between service quality and customer satisfaction: In the case of CCG (Customer Centric Group) CO*. Retrieved from <http://www.essays.se/about/Ali+Dehghan/>
- Dehghan, A., & Shahin, A. (2011). Customer Loyalty Assessment-A Case Study in MADDIRAN, the Distributor of LG Electronics in Iran. *Business Management and Strategy*, 2 (1).

- Dick, A.S. & Basu, K. (1994). Customer loyalty: Toward and integrated conceptual framework. *Journal of the Academy of Marketing Science*, 22(2), 99-113. <http://dx.doi.org/10.1177/0092070394222001>
- Eid, M. I. (2011). Determinants of e-commerce customer satisfaction, trust, and loyalty in Saudi Arabia. *Journal of Electronic Commerce Research*, 12(1), 78-93.
- Fang, Y. H., Chiu, C. M., & Wang, E. T. (2011). Emerald Article: Understanding customers' satisfaction and repurchase intentions: An integration of IS success model, trust, and justice. *Internet Research*, 21(4), 479-503 <http://dx.doi.org/10.1108/10662241111158335>
- Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J. & Bryant, B.E. (1996). The American customer satisfaction index: nature, purpose, and findings. *Journal of Marketing*, 60(4), 7-18. <http://dx.doi.org/10.2307/1251898>
- Gummesson, E. (1994). Making relationship marketing operational. *International Journal of Service*, 5 (5), 5-20. <http://dx.doi.org/10.1108/09564239410074349>
- Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2007). Computer Assisted Customer Churn Management: State-Of-The-Art and Future Trends. *Computers & Operations Research*, 34 (10), 2902-2917. <http://dx.doi.org/10.1016/i.cor.2005.11.007>
- Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International Journal of Service Industry Management*, 7 (4), 27-42. <http://dx.doi.org/10.1108/09564239610129931>
- Hoyt, J. E., & Howell, S. L. (2011). Beyond customer satisfaction: reexamining customer loyalty to evaluate continuing education programs. *The Journal of Continuing Higher Education*, 59(1), 21-33. <http://dx.doi.org/10.1080/07377363.2011.544979>
- Jacoby, J. & Chestnut, R. (1978). *Brand Loyalty Measurement and Management.*, New York, NY: John Wiley & Sons.
- Hung, S.-Y., Yen, D. C., & Wang, H.-Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, 31, 515-524. <http://dx.doi.org/10.1016/j.eswa.2005.09.080>
- Kim, M.-K., Park, M. C., & Jeong, H. D. (2004). The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. *Telecommunications Policy*, 28, 145-159. <http://dx.doi.org/10.1016/j.telpol.2003.12.003>
- Kotler, P., and Fox, K. F. (1995). *Strategic marketing for educational institutions*. Prentice-Hall.
- Lariviere, B., Van den Poel, D., & Van den Poel (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. *Expert Systems with applications*, 27(2), 277-285. <http://dx.doi.org/10.1016/j.eswa.2004.02.002>
- Lee, W., Stolfo, S. J., & Mok, K. W. A Data Mining Framework for Building Intrusion Detection Models. *IEEE Symposium*. http://dx.doi.org/10.1007/978-3-642-02617-1_62
- Loveman, G.W. (1998). Employee satisfaction, customer loyalty, and financial performance. *Journal of Service Research*, 1(1), 18-31. <http://dx.doi.org/10.1177/1094670598001000103>
- McKinney, V., Yoon, K., and Zahedi, F. (2002). The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach. *Information Systems Research*, 13 (3), 296-315. <http://dx.doi.org/10.1287/isre.13.3.296.76>
- McGarry, D.E. (1995). The road to customer loyalty. *Canadian Business Review*, 22(1), 35-36.
- Oliver, R.L. (1999). Whence customer loyalty. *Journal of Marketing*, 63(4), 33-44. <http://dx.doi.org/10.2307/1252099>
- Ramanathan, U. & Ramanathan, R. (2011). Guests' perceptions on factors influencing customer loyalty - an analysis for UK hotels. *International Journal of Contemporary Hospitality Management*, 23(1), 7-25. <http://dx.doi.org/10.1108/09596111111101643>
- Reichheld, F.F. (1996). *The Loyalty Effect*. Boston, MA: Harvard Business School Press.
- Rowley, J. & Dawes, J. (1999). Customer loyalty – a relevant concept for libraries? *Library Management*, 20(6), 345-351. <http://dx.doi.org/10.1108/01435129910280474>

- Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need-satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, 76 (3), 482-497. <http://dx.doi.org/10.1037//0022-3514.76.3.482>
- Smith, D.C. & Whan Park, C. (1992). The Effects of Brand Extensions on Market Share and Advertising Efficiency. *Journal of Marketing Research*, 29(3), 296-313. <http://dx.doi.org/10.2307/3172741>
- Soderlund, M. (1998). Customer Satisfaction and its consequences on customer behavior revisited the impact of different levels of satisfaction on word-of-mouth, feedback to the supplier and loyalty. *International Journal of Service Industry Management*, 9 (2), 169-188. <http://dx.doi.org/10.1108/09564239810210532>
- Spreng, R. A., Mackenzie, S. B., and Olshavsky, R. (1996). A Reexamination of the Determinants of Consumer Satisfaction. *The Journal of Marketing*, 60 (3), 15-32. <http://dx.doi.org/10.2307/1251839>
- Tepeci, M. (1999). Increasing brand loyalty in the hospitality industry. *International Journal of Contemporary Hospitality Management*, 11(5), 223-229. <http://dx.doi.org/10.1108/09596119910272757>
- Tsai, C. F., & Lu, Y. H. (2009). Customer churn prediction by hybrid neural networks. *Expert Systems with Applications*, 36(10), 12547-12553. <http://dx.doi.org/10.1016/j.eswa.2009.05.032>
- Uncles, M. D., Dowling, G. R., & Hammond, K. (2003). Customer loyalty and customer loyalty programs. *Journal of Consumer Marketing*, 20 (4), 294-316. <http://dx.doi.org/10.1108/07363760310483676>
- Zeithaml, V. A., Berry, L. L., and Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, 60 (2), 31-46. <http://dx.doi.org/10.2307/1251929>
- Zins, A. H. (2001). Relative attitudes and commitment in customer loyalty models: Some experiences in the commercial airline industry. *International Journal of Service Industry Management*, 12 (3), 269 - 294. <http://dx.doi.org/10.1108/EUM0000000005521>

Table 1. *Questions used for the assessment of customer loyalty (Dehghan and Shahin, 2011, p. 7)*

Author	Questions
Chaudhuri and Holbrook, 2001; Oliver, 1997; Pritchard et al., 1999; Sirdeshmukh et al., 2002; Taylor et al., 2004	- I use heavy equipment from the company I am evaluating because it is the best choice for me. - I consider myself to be a loyal patron of the manufacturer of heavy equipment I am evaluating.
Wong (2004)	- Customer says positive things about retail store XYZ to other people - Customer recommend retail store XYZ to someone who seeks his advice - Customer encourages friends and relatives to shop at retail store XYZ
Colwell et al., (2009)	- I believe I have a strong relationship with my bank and would not leave because of better fees - I believe I have a strong relationship with my bank and would not leave because of better rates - If I need other types of bank service I will choose this bank - I will recommend this bank to others who seek my advice - I will continue to use this bank service
Thuy and Hau (2010)	- I only pay my attention to this bank
Kassim and Abdullah (2010)	- I will recommend the online organization to other people (WoM) - I would recommend the organization's website to others (WoM) - I intend to continue using the online organization (Intent) I prefer the online organization above others (Intent)

Table 2. Accuracy

Parameter	C	Accuracy +1	Accuracy -1	Ave. Accuracy
1,2	1000	57.16	57.18	57.17
1,2	100	57.16	55.66	56.41
1,2	10	57.27	55.34	56.305
1,2	1	57.51	55.98	56.745
0,2	1000	58.55	55.86	57.205
0,2	100	56	55.21	55.605
0,2	10	52.54	68	60.27
0,2	1	57.62	55.68	56.65
0,1	1000	53	66.77	59.885
0,1	100	54.5	66.09	60.295
0,1	10	51.85	69.2	60.525
0,1	1	56.12	65.05	60.585

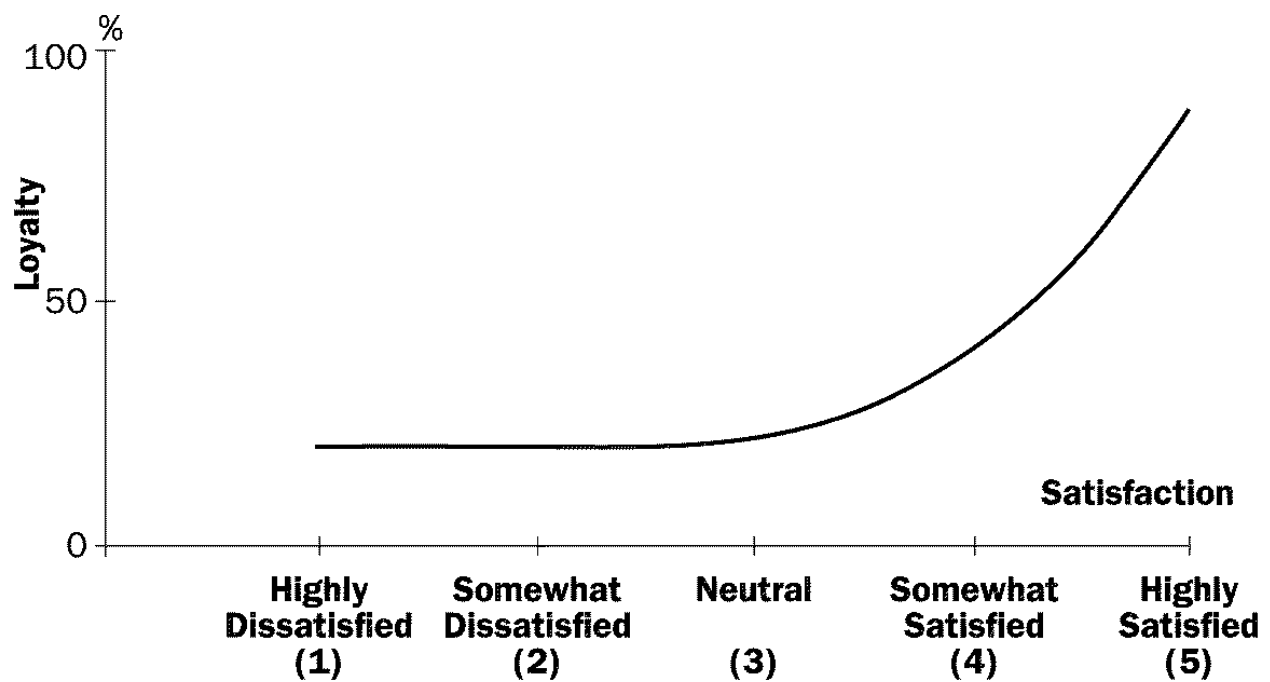


Figure 1. Customer satisfaction / loyalty Relationship

Source: Tepeci, M. (1999). Increasing Brand Loyalty in the Hospitality Industry. *International Journal of Contemporary Hospitality Management*, 11(5), 223-229.

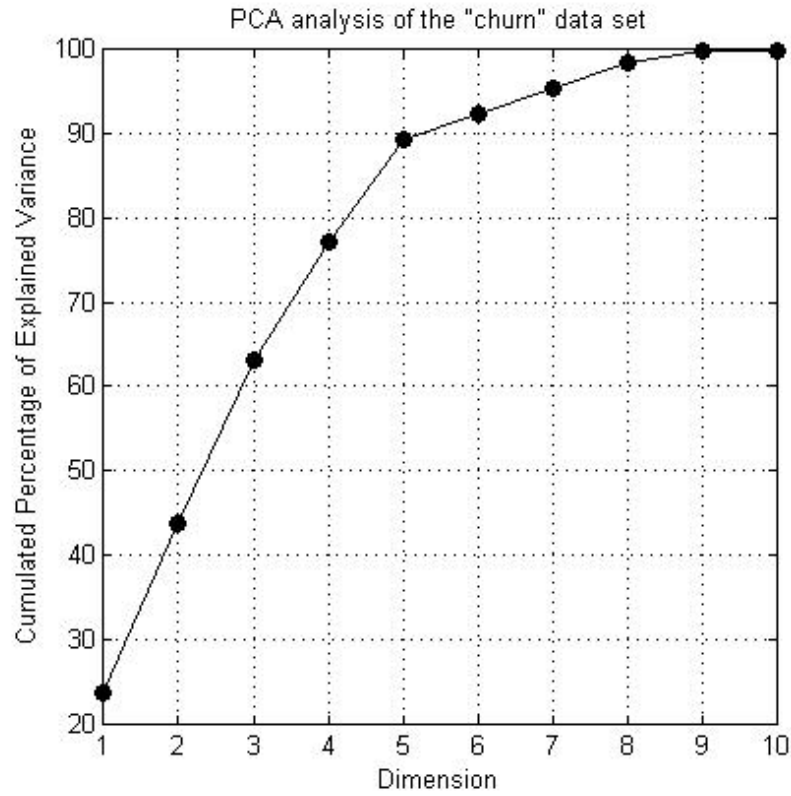


Figure 2. PCA Analysis

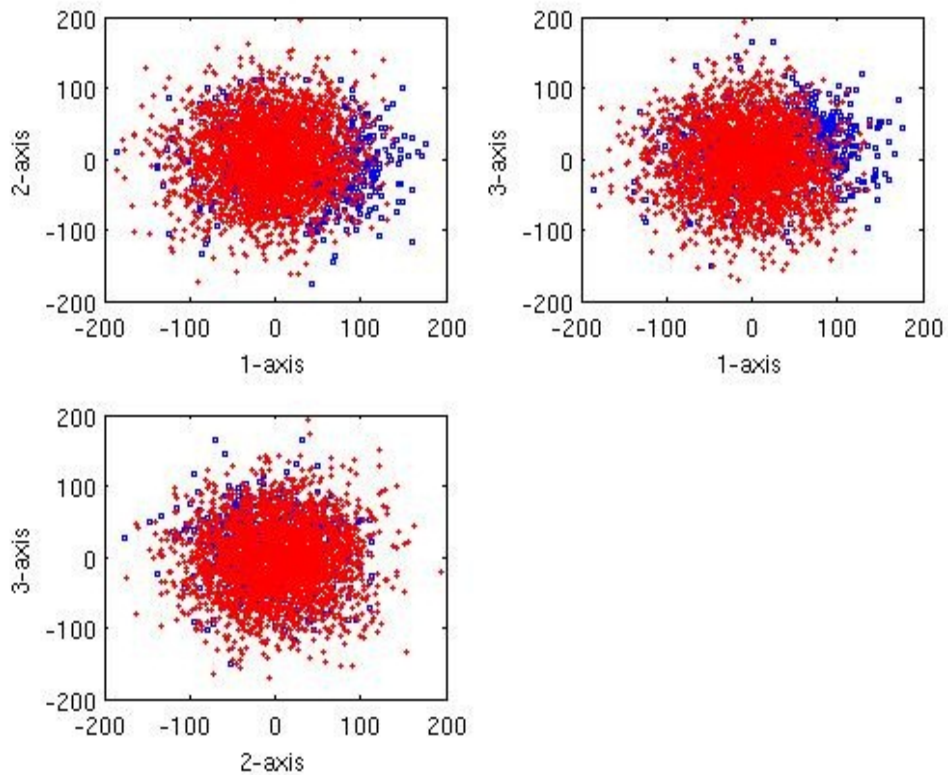


Figure 3. Classification

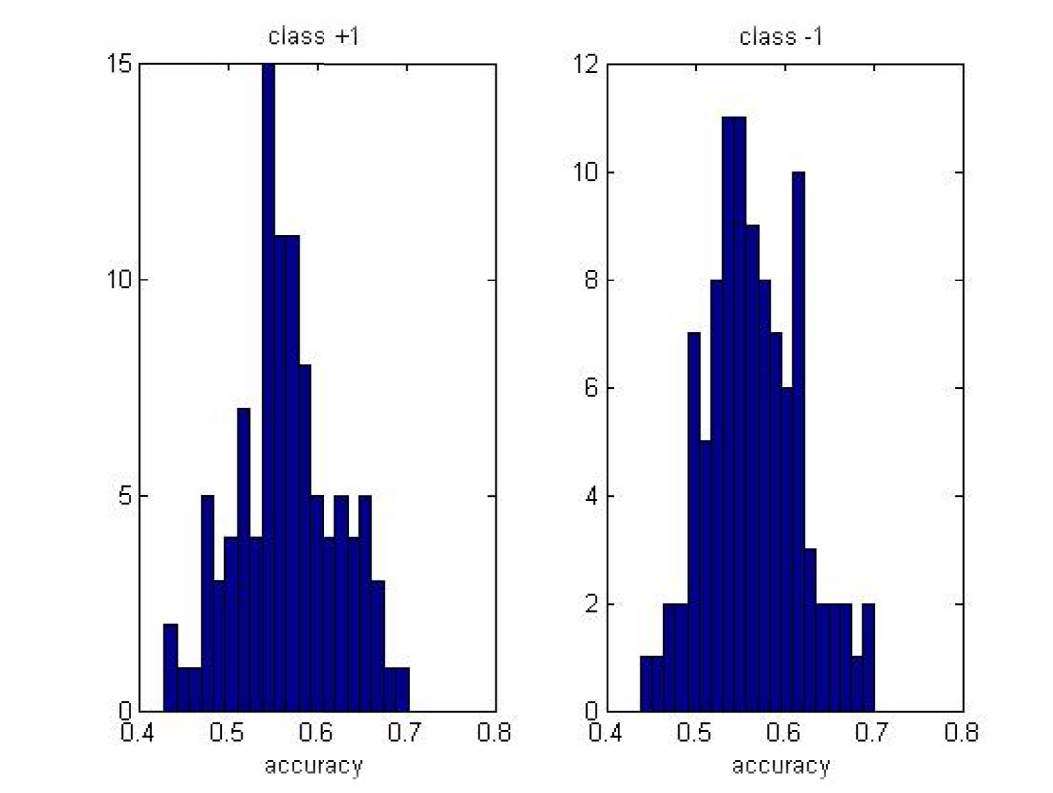


Figure 4. Accuracy