UNIVERSITY OF CAMBRIDGE
Cambridge Service Alliance

The Fallacy of the Net Promoter Score: Customer Loyalty Predictive Model

Mohamed Zaki, Dalia Kandeil, Andy Neely, and Janet R. McColl-Kennedy

This is a working paper.

Why this paper might be of interest to Alliance Partners:

The Net Promoter Score (NPS) is still a popular customer loyalty measurement despite recent studies arguing that customer loyalty is multidimensional. Therefore, firms require new data-driven methods that combine behavioral and attitudinal data sources. This paper provides a framework that holistically assesses and predicts customer loyalty using attitudinal and behavioral data sources. We built a novel customer loyalty predictive model that employs a big data approach to assessing and predicting customer loyalty in a B2B context. We demonstrate the use of varying big data sources, confirming that NPS measurement does not necessarily correspond to actual behavior. Our model utilises customers' verbatim comments to understand why customers are churning.

October 2016

Find out more about the Cambridge Service Alliance:
Linkedin Group: Cambridge Service Alliance
www.cambridgeservicealliance.org

The Fallacy of the Net Promoter Score: Customer Loyalty Predictive Model

Mohamed Zaki, Dalia Kandeil, Andy Neely,
Cambridge Service Alliance, University of Cambridge

Janet R. McColl-Kennedy
The University of Queensland, Brisbane, Australia and
Visiting Professor at the Cambridge Service Alliance, University of Cambridge, UK

Single-question customer metrics have become popular as tools to measure customer loyalty (Wiesel, Verhoef, and de Haan 2012). Indeed, many organizations use the Net Promoter Score (NPS), following the claim that a single question is all that is needed to predict a company's financial performance and growth (Reichheld 2003). Indeed, it is argued that by asking "How likely is it that you would recommend company X to a friend or colleague?", a firm is able to assess overall customer loyalty based on the customer's intention to refer others. Although the NPS measure can be used as a loyalty indicator, it does not offer an explanation of the root cause or causes of a low score. Further, this measure is typically taken at the end of the customer journey, thus potentially masking the underlying issues of concern, which form the basis for identifying improvements. Relying solely on a simple single customer metric is risky and so companies are encouraged to adopt a more nuanced multidimensional approach to better predict customer behavior (Keiningham, et al. 2007; Wiesel, Verhoef, and de Haan 2012). Keiningham et al. (2007) stated that a "combination of VOC metrics universally outperforms the use of only the NPS's recommendation intentions when predicting actual loyalty behavior." Clearly, the widely acclaimed NPS is based on a customer's attitude rather than his or her actual behavior.

Our work contributes to the understanding and management of customer loyalty measurement in the following four ways. First, the study contributes to establishing the unreliability of the NPS or overall satisfaction as a single loyalty measure within B2B complex service organizations. Second, the proposed framework integrates multiple sources of customer data, by including demographic, behavioral and attitudinal customer data when assessing customer loyalty. The combination of multiple data sources when assessing customer loyalty supports the criticism of using a single loyalty metric evident in the literature (Aksoy 2013; Keiningham et al. 2007; Kristensen and Eskildsen 2011; Pollack and Alexandrov 2013). Making a judgment on customer loyalty without consideration of behavioral data is misleading. Third, our predictive analytics model uses big data techniques to predict customer loyalty and identifies customers who are no longer conducting business with the organization. In particular, our study contributes to setting up systematic multi-methods using a big data approach to capture and analyze customers' data from different sources. Fourth, we extended the linguistic text-mining approach introduced by Villarroel Ordenes et al. (2014) to determine the complaint status and emotions of each customer using text-mining textual feedback that divides customers into groups of complainers, neutral or satisfied. The integration of the verbatim comments helps firms to understand *why* customers are leaving or inactive.

NET PROMOTER SCORE CRITICISM

Customer loyalty measurement relies mainly on survey-based metrics and uses the overall satisfaction or repurchase intention scores (Bolton and Drew 1991; Parasuraman 2006). In particular, Reichheld (1993) claimed that the NPS can be used as a standalone metric for measuring customer loyalty and that it is superior to other VOC metrics. Each customer answers the question *"How likely is it that you would recommend company X to a friend or colleague?"* on an 11-point rating scale, ranging from 0 (not at all likely) to 10 (extremely likely). Customers are then grouped based on their chosen ratings into three segments: promoters, passives, and detractors. *Promoters* are defined as customers having the highest referral intention, with a rating of 9 or 10, *passives* with ratings of 7 or 8 and *detractors* ranging from 0 to 6 on the NPS scale (Reichheld 2006). The NPS has been used to predict a company's financial growth (Reichheld 2003).

However, questions have been raised about the veracity of the claims that the NPS is the only score needed. For example, Keiningham et al. (2007) replicated the analyses used in net promoter research and compared the findings of Reichheld (2003) and Satmetrix with the American Customer Satisfaction Index using the same industries employed in Reichheld's study. Their research rejects the claim that the Net Promoter Score (NPS) is the "*single most reliable indicator of a company's ability to grow."* In its macro-level analysis, the study found no real indication that average levels of attitudinal loyalty metrics significantly correlate with the relative change in revenue within the respective industry. Furthermore, single metrics alone cannot predict customer loyalty and consequently are unlikely to deliver actions to managers. Customers' loyalty-based behaviors are multidimensional and therefore a better measurement tool is required.

Keiningham et al. (2007) examined different customer satisfaction and loyalty metrics and evaluated their relationship to customer retention, recommendation and share of wallet. Again, this study confirmed that the recommended intention single measure alone is not sufficient to assess customer loyalty. Rather, the authors suggest that a multidimensional metric would perform better in predicting customer loyalty. As such, the studies of Keiningham et al. (2007) demonstrate that there are no simple solutions for turning loyalty into business growth and that firms should balance and manage different aspects of the customer experience simultaneously to improve customer loyalty. Failing to do so would be misguided, resulting in the associated resources and actions being misallocated.

Furthermore, Grisaffe (2007) scrutinized the conceptual foundations of the NPS based on a practitioner's perspective and principles of social science and marketing methodology. Grisaffe argued that the NPS is not sufficient as an approach to customer loyalty measurement and management because recommendations alone are unable to drive business success. For example, customers could give high NPS scores (promoters), but a firm could lose a percentage of its customer base. Thus, the metric misinforms managers and diverts them away from marketing actions such as retaining the existing customer base and understanding their lifetime value. The research also emphasized that a single diagnostic measure is not comprehensive, and thus it is essential for firms to undertake detailed investigations. Indeed, the authors suggest that a "total system/multi-dimensional" approach in terms of indicators capturing the customer experience (both attitude and response to offering) would help companies to succeed in the long run, and to understand the root cause of customers' problems in order to identify the best strategic actions. In particular, the NPS does not provide such a prescription for firms to diagnose the underlying causal factors.

Pingitore et al.'s (2007) study used customer feedback data across automotive industry, insurance, full-service investment, airlines and car rentals to examine the strengths and limitations of the NPS from a practitioner's perspective. The research contrasted four net VOC metrics (net delighted, net committed, net satisfied, net promoter) with four standard continuous satisfaction and loyalty VOC metrics (overall satisfaction, customer satisfaction index, commitment, likelihood to recommend). They examined the impact of multiple VOC metrics on industry-dependent financial performance metrics and compared the NPS scale to its four-point likelihood-to-recommend scores. The study found that the intention-to-recommend question is not critical and, furthermore, that the NPS is not the only net customer feedback that correlates with financial performance. Other questions such as net satisfied, net delighter and net committed provide equally robust correlations to different financial outcomes. Clearly, firms should look to different measurements to assess customer loyalty.

Pollack and Alexandrov (2013) provided insights into the validity of the Net Promoter Index (NPI) to measure customer loyalty and predict financial performance. The study investigated the net promoter question as an alternative to the traditional word-of-mouth measure. The findings did not support the assertion of NPI's superiority over other voice of customer metrics. Furthermore, the findings suggest that a combination of metrics is best for predicting actual loyalty behaviors, thereby confirming the study by Keiningham et al. (2007). There is no empirical evidence to support the predictive ability of NPI on growth and financial performance. The authors recommend that firms do not use NPI as a standalone diagnostic tool, or as a predictor of growth or financial performance.

Morgan and Rego (2006) empirically investigated which customer satisfaction and loyalty metrics (average customer satisfaction score and net promoter scores) are most useful in predicting future business performance. The study employed the American Customer Satisfaction Index (ACSI) data and evaluated the link between six loyalty metrics and firm performance. The findings revealed a negative correlation between net promoters and gross margin performance. The results clearly demonstrated that firms should not focus solely on customer feedback systems and customers' recommendation intentions and behavior metrics. Furthermore, the study suggests that increasing the number of net promoters will fail to improve a firm's performance and promoters are subsequently not buying significantly more and/or that they may not influence potential new customers. The authors studied the impact of customer complaints' behavior on business performance and the analysis suggested that monitoring customer complaints does provide insights into satisfaction and is valuable for predicting future business performance.

Another study by Kristensen and Eskildsen (2011), which questions the NPS claim and asserted that NPS is a very poor predictor of customer loyalty and satisfaction, investigated the superiority claim of the NPS as a loyalty measure for firm growth. The research used a 2006 customer satisfaction survey of the entire insurance sector in Denmark. This study was different to previous studies in the sense that it conducted an experiment in the Danish insurance industry to answer specific questions concerning NPS. The paper suggested that using the NPS as a key element in managerial decision-making might mislead organizations in measuring customer loyalty. The findings in this paper are consistent with previous research (Grisaffe 2007; Keiningham, Cooil, Andreassen et al. 2007; Keiningham et al. 2008). However, the analyses were based on business-to-consumer data from Denmark. More research is needed on business-to-business contexts and in different countries. Another study using the same data set as Eskildsen and Kristensen (2011) demonstrated that demographics also need to be investigated because the NPS suffers from distinct gender differences. In particular, the distribution of females and males is different when it comes to promoters and detractors, but the same for passives. Therefore, there are serious issues related to the robustness with the NPS because females rate higher than males within the promoter category. However, female promoters tend to have a smaller effect on the NPS score. Thus, fusing survey-based metrics alone is challenging because organizations rely on the respondents' memory of a service process or a transaction, which may not always be a correct representation of the actual occurrence (Kristensen and Eskildsen 2014).

Big Data Techniques

We argue that market researchers will serve both their organizations and customers better if they play an active role in updating the customer loyalty measurement by using big data techniques. Firms could benefit from the use of more sophisticated and advanced modelling approaches, which have the potential to uncover patterns in customer data and to link with business results (Aksoy 2013). Approaches such as data mining could be incorporated to evaluate customer loyalty rather than relying solely on survey-based measurement and statistical techniques. For example, Hosseini, Maleki and Gholamian (2010) proposed a data-mining model to assess customer loyalty and customer lifetime value based on recency, frequency and monetary (RFM) attributes and the K-means clustering method. The study assessed the degree of customer loyalty to maximize the profits of B2B organizations. The proposed model demonstrated a significant improvement in the accuracy of customer loyalty measurement. However, this study has limitations; the model relied on combining data-mining techniques alone and did not consider the impact of textual data such as verbatim comments from customers, which are now generated at many touchpoints in the customer's journey. Specifically, online chatter, or user-generated content, constitutes an excellent emerging source for marketers to mine meaning at high temporal frequency (Tirunillai and Tellis 2014), and to gain insights into specific points of friction.

Baumann, Elliott and Burton (2012) studied the factors that help to explain customers' behavioral intentions in retail banking. In particular, the research investigated the associations between customer satisfaction, perceived service quality, recent and current consumer behavior and long-term intentions to remain a customer. This research contributes to the customer loyalty literature by adding other loyalty predictors, such as customers' perceptions of market conditions (perceived switching costs and benefits), and customer characteristics, such as demographic factors, which have received less attention in the marketing literature. The study found that customer perceptions, market conditions and some customer characteristics are unique predictors of behavioral intentions. The full model explained 56.9 per cent of behavioral intentions in retail banking.

Customer relationship and churn management literature has discussed the importance of using data-mining techniques (Neslin et al. 2006; Hopmann and Thede 2005; Wubben 2008; Wübben and Wangenheim 2016) to predict customer loyalty. Furthermore, Tamaddoni, Stakhovych and Ewing (2015) compared different churn prediction techniques in non-contractual settings and assessed their performance. The study used simulated data from two online retailers and evaluated the performance of different customer churn classifiers (Support Vector Machine (SVM) and one probability model (Pareto/NBD)) against logistic regression and the RFM model. The findings suggested that campaigns that use the boosting technique have the highest profits compared to campaigns designed using SVM, logistic regression, Pareto/NBD or RFM methods. Moreover, campaigns using the Pareto/NBD model have the lowest profit. Furthermore, the Pareto/NBD model is unable to outperform logistic regression.

Another study by Coussement and Poel (2009) investigated the impact of emotionality indicators on churn behavior and compared three classification techniques, namely, Logistic Regression, SVM and Random Forests, to distinguish churners from non-churners. They found that adding emotions expressed in client/company emails increases the model's performance. The study suggested that the use of Random Forests improves predictive performance when compared to SVMs and Logistic Regression.

Text mining is another big data technique that has been used recently to extract customer opinions from unstructured comments. For example, Pang and Lee (2008) utilized text mining to extract sentimental insights from customer data to improve customer loyalty measurement. Tirunillai and Tellis (2014) used text analytics to understand the dimensions of product quality in order to gain insights into brand positioning. Using longitudinal data on product reviews across firms and markets, their study extracts specific latent dimensions of quality, and the valence, labels, validity, importance, dynamics and heterogeneity of those dimensions. Villarroel Ordenes et al. (2014) utilized text analytics to analyze customer feedback in order to capture key aspects of the customer experience. In particular, the text-mining model captured customer activities and resources, company activities and resources, and customer sentiment (compliments, complaints or suggestions) from the customer satisfaction data. However, their research did not combine the qualitative data and quantitative data to assess and predict customer loyalty.

Together, this nascent stream of research suggests that there is a strong need to combine data mining and text mining to better analyze and predict customer loyalty. Clearly, this review highlights a critical research gap and important research opportunity. To date the methods have been confined to either qualitative or quantitative techniques. Our paper seeks to fill this gap, and to further our knowledge of customer loyalty measurement using a data-driven approach. Additionally, most firms do not measure both the attitudinal and behavioral components of loyalty, instead measuring either one or the other (Aksoy 2013). Although the marketing literature stresses the importance of repurchase behavior as a loyalty assessor, the most popular performance measure, NPS, completely disregards it. To address this, the NPS was compared to actual customer spending patterns in our model. This enables a fuller understanding of actual customer behavior, not just customers' self-proclaimed loyalty referral intentions. Another contribution is that we built a customer loyalty model that combines a multitude of data sources representing both loyalty measures: attitudinal and behavioral (Aksoy 2013; Uncles, Dowling, and Hammond 2003). In addition, the "voice of customer" was taken into account through the extraction of customer sentiment within feedback survey comments and incorporated into our proposed loyalty assessor model. Specifically, we built on and extended the study of Villarroel Ordenes et al. (2014) using a linguistic-based text-mining approach that combines qualitative data (text analytics) on specific customer experience touchpoints with quantitative company data. In addition, our model offers the predictive ability of determining whether customers are likely to churn, thereby increasing the model's functionality.

METHODOLOGY

Transforming customer loyalty data into systematic customer insights requires a formal system by which measures are included as part of a data collection and analytics process. Our aim was to develop a new loyalty assessment model that combines both text and predictive analytics. Rather than relying on survey-based NPS measurement, we applied big data techniques to customer data gathered from multiple sources in order to evaluate customer loyalty. In addition, we demonstrate the use of these data sources are essential to confirm that the self-proclaimed referral intention of the single measurement does not necessarily translate into behavior and purchasing patterns. Importantly, our study utilized customers' textual feedback to understand *why* customers are churning or inactive, and quantified the generated textual categories for use in our customer loyalty model. To do so, several stages were required following the Cross Industry Process for Data Mining (CRISP-DM) (Chapman et al. 2000). The CRISP-DM process is the "de facto standard for developing data mining and knowledge discovery projects" (Marbán et al. 2009).

As shown in Figure 1, the multi-staged research program involves five key steps: 1) business understanding; 2) data source understanding; 3) predictive variables; 4) model construction; and 5) model evaluation. We undertook a pre-study, comprising interviews conducted to understand the participant organization's current practices. We used a longitudinal customer data set covering attitudinal data (customer surveys, which include NPS ratings and qualitative customer comments), behavioral data (transactional data) and demographic data across multiple touchpoints for a large international B2B service organization. The data sources were limited to a three-year timeframe (2012–15), with a total of 1,044,512 transaction records. As such, it provided an excellent setting in which to develop B2B customer loyalty, as there were multiple touchpoints with both the same and multiple customers, across multiple time points. We employed a variety of analytic approaches to develop, test and validate our customer loyalty model. We used the recency, frequency and monetary (RFM) technique to transform customer transactional data into profitability scores, facilitating the categorization of customers based on purchasing behavior. Furthermore, the K-means clustering technique was used to identify customers who had *churned* and were no longer conducting business with the organization. We extended the linguistic text-mining approach introduced by Villarroel Ordenes et al. (2014) to determine the complaint status of each customer by mining textual survey feedback that divides customers into groups of complainer, neutral or satisfied. Finally, the prediction model used neural network and Bayesian network algorithms to predict customer loyalty. This methodology enables managers to transform data into information about the customer and assists them with making informed decisions.

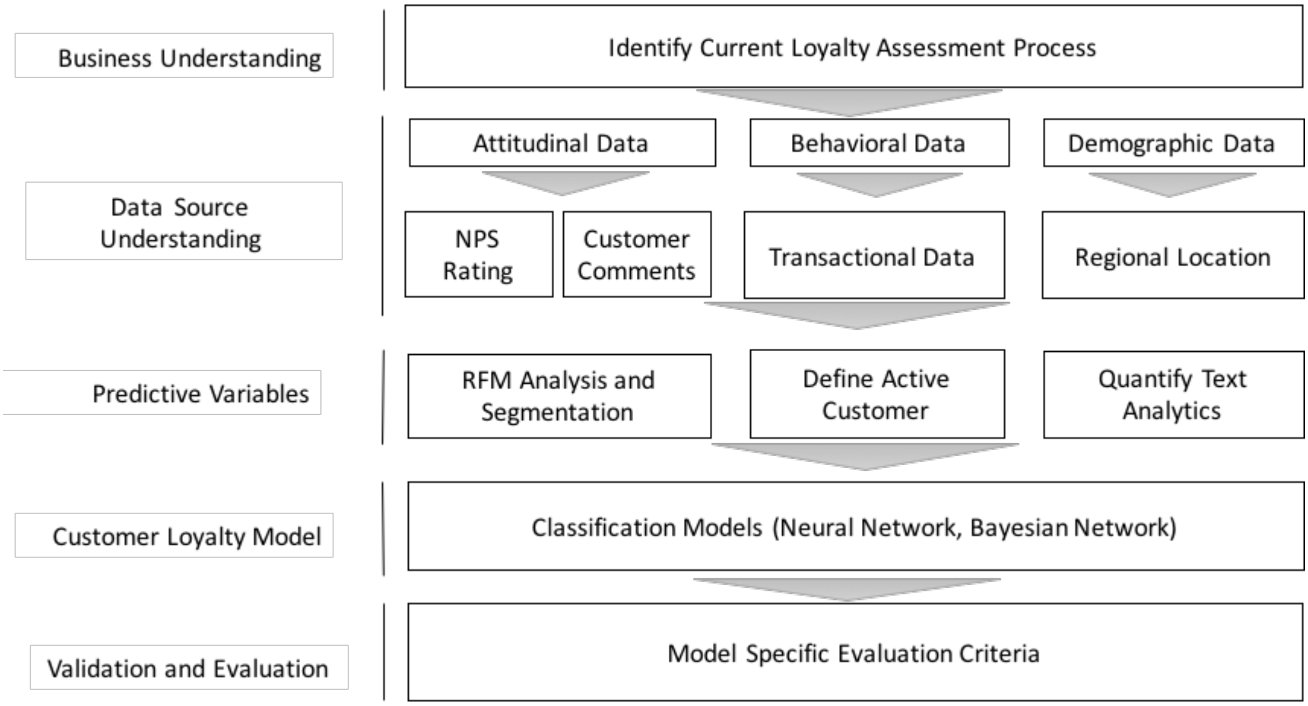


Figure 1. Customer Loyalty Assessment and Prediction Framework

Business Understanding

The participant company is a UK-based asset-heavy organization, where there are many customer–firm touchpoints along a customer's journey. The organization has a widespread customer base, with each customer having multiple branches. The company has many inter-company relationships, which may include a partnership with the original manufacturer, the company's customers and the company's competitors. The firm offers both products and services to B2B customers, making it a suitable selection for the intent of exploring customer loyalty assessment measures existing within a complex service industry. It also offers an opportunity to address the research gaps found within the literature concerning the lack of research into the validity of NPS in both the B2B and service settings (Kristensen and Eskildsen 2014; Pollack and Alexandrov 2013).

At this stage, interviews were conducted with key informants, including the participant organization's General Manager of Service Excellence across the UK, the Customer Experience Manager, and the Strategic Customer Performance Analyst in charge of the customer feedback process. The aim of the interviews was to understand the current overall customer loyalty process to identify both the strengths and weaknesses of the current measurement system. The questions describe the company's customer experience strategy and customer loyalty measurement process. Questions were asked about what the experience is like, the customer experience strategy, what are the pain points of customers, why they believe this occurs, the customer loyalty (NPS) measurement process, and how the organization is currently using these insights.

The main customer experience strategy of the participant company focuses on making customers feel appreciated, giving them a sense of security and offering them an effortless experience when dealing with the company. The organization is committed to highlighting areas of the business that need improvement or are operating efficiently, and to identifying the relevant practices. The organization focuses on five pillars. First, engagement is emphasized to ensure that all business segments within the company are aligned with the customer experience strategy and to understand its effect on the business's performance. Second, a protocol is designed for dealing with customers and methods of communication that specify who interacts with a customer, how they interact and when this takes place. It is also important to note that this step is currently in development. Third, structure and governance aim to ensure that launched projects are tracked effectively and that each business segment has a voice in the business. They are also used to set a precise definition of each business segment's role in the overall business cycle. The fourth pillar is a measurement to assess the company's current performance regarding customer experience. It is administered by a third party through a telephone survey conducted on a weekly basis with the aim of capturing the real voice of the customer. The final pillar is continuous improvement based on the insights captured in the measurement step to improve the indicated business segment.

The focus of our study resides in the measurement pillar, specifically understanding the current method and suggesting improvements. The customer satisfaction measurement is based on a customer survey, which is administered on a monthly basis (including structured and unstructured data). The survey includes questions on overall satisfaction, repurchase, referral, resource availability, responsiveness, communication, service completion duration, preparation, service quality, invoice timeliness and invoice accuracy. Customers rate each question from 1 to 10, where 10 is "Very Satisfied" and 1 is "Very Dissatisfied." The final question is open-ended: "Do you have any other comments or suggestions on how (NAME) could improve this service." This allows the company to gather real-time comments from customers and to obtain information about the service that is not captured elsewhere.

When the company receives the customer feedback, manual analysis of the data is performed on a monthly basis to identify customer concerns. Currently, if a score of 7 or lower is given then an "Alert" is triggered, and every month an alerts report is generated using the monthly survey data. All customers, regardless of the type of transaction, are asked the NPS question, "*How likely is it that you would recommend the company to a friend or colleague?"*, from which customer loyalty is assessed. According to the Customer Experience Manager, the company uses "*the industry standards' NPS to measure customer loyalty and that it's the one metric that interests the company."* Moreover, the company has no clear method of determining customer inactivity as a result of the lack of a full CRM system.

Data Source Understanding

To acquire a holistic view of customer loyalty, we integrated data across multiple systems. The data was classified into three categories: attitudinal, behavioral and demographics. The attitudinal data was collected from the customer survey, which includes structured (NPS rating) and unstructured data (verbatim comments). The behavioral data was collected from the financial system. This data consists of sales (new, used, lease), product support (parts and service transaction types) and customer service agreement (CSA) transactions (parts and service transaction types). Two groups of customers were identified: those who have a maintenance contract with the company, referred to as Customer Service Agreement (CSA) customers; and those who deal in a transactional setting, referred to as Product Support (PS) customers. Demographic data, which contains the regional locations of customers, was included. In total, we collected 1,044,512 transaction records over a three-year period.

We undertook many interactions with the company to ensure data integrity and quality. These data sources were essential for the following reasons: 1) customers' loyalty-based behaviors are multidimensional, and for this reason the attitudinal and behavioral data are essential components of measuring customer loyalty; 2) to establish a link between a customer's self-proclaimed referral intention and their actual spending patterns, accomplished by comparing the NPS score with customer transactional data; and 3) to analyze the customer textual feedback found in surveys in order to identify customer pain points and to unmask the root causes that may affect customer loyalty.

Predictive Variables of Customer Loyalty Model

This section outlines the techniques we used as predictive variables for our customer loyalty model. The recency, frequency and monetary (RFM) technique is a well-known method for determining a customer's purchasing behavior based on historical transactional data. Originated in 1997 by Cullinan, it was later adopted for direct database marketing purposes by Hughes (1996). RFM variables were found to be important predictors of future customer life value (CLV) and customer behavior and churn (Ballings, Van Den Poel, and Verhagen 2012; Bardauskaite 2014; Tamaddoni, Stakhovych, and Ewing 2015). Notably, recency is considered to have the highest predictive influence (Coussement and Van den Poel 2008). Since its development, it has been widely adopted as a measure of behavioral loyalty.

The RFM model is essential because it transforms customer transactional data into profitability scores, facilitating the categorization of customers based on their purchasing behavior. In order to perform the RFM analysis, customers' purchasing patterns are observed over a predefined period. This is done by first sorting customers based on their recency values and then dividing the customer base into five groups. Each customer then receives a rank, where customers with the lowest recency values receive a score of (5), and customers with the highest recency values receive a score of (1). This is because a customer who has made a recent transaction is considered more valuable to the company. This step is repeated for both the frequency and monetary values, with the difference between the highest values being assigned a score of (5). Finally, each customer receives an overall RFM score, which is a combination of each RFM variable rank. Table 1 shows the RFM predictive variable scores of some customers in 2012. For example, the customer with ID (#04478) has an overall RFM score of 445, because he had made a recent purchase and was a frequent buyer (97 times), with £71,948.82 monetary value. In contrast, customer ID (#01992) has an overall RFM score of 111 because his last purchase was 174 days ago, he bought four times from the participant company, and with a low monetary value of £220.

Customer_ID	Recency	Frequency	Monetary (£)	Recency_Score	Frequency_Score	Monetory_Score	RFM_Score
01425	92	13	170389.47	1	2	5	125
01761	14	26	105758.16	3	3	5	335
02543	11	14	9325.05	4	2	3	423
01992	174	4	220.00	1	1	1	111
01394	21	10	841.10	2	2	1	221
03381	76	15	8906.77	1	2	3	123
04250	14	38	71348.02	3	3	5	335
00272	12	61	92463.13	3	4	5	345
00804	40	4	31083.00	2	1	4	214
02511	46	7	35254.93	1	1	4	114
02874	11	18	21441.22	4	2	4	424
02176	18	96	176817.70	3	4	5	345
03391	13	75	86651.25	3	4	5	345
04478	11	97	71948.82	4	4	5	445

Table 1. RFM Predictive Variables

The K-means clustering model is a data-mining technique used to segment data points into groups, each containing data points similar to one another and dissimilar to data points in other groups (Lloyd 1982). Considered the most widely used technique in a commercial context (Parvaneh, Abbasimehr, and Jafar 2012; Qin, Zheng, and Huang 2010), it uses an iterative algorithm that continuously readjusts until the best possible segmentation is achieved. In each iteration of the algorithm, each record is assigned to the cluster whose center is closest.

In our study, we divided customers into 11 groups based on their RFM scores. This was accomplished using the K-means segmentation algorithm. Each resulting K-means cluster has an average RFM score, which is the mean of RFM scores of all customers belonging to that group. A total of 11 clusters was chosen to mimic the 11-point scale of the NPS, which would allow the classification of customers based on NPS's promoter, passive and detractor categories. The 11 customer groups were then sorted based on their average RFM score and given a corresponding NPS scale. For example, customers belonging to cluster number (1) had the highest average RFM score of (555) and received an NPS score of (10). However, customers grouped into cluster number (5) with an average RFM of (111) received an NPS score of (0) on the 11-point scale. This process was repeated for each of the (11) customer groups, as shown in Table 2. NPS categories were then formulated for the developed corresponding NPS scale, according to the classification of promoters (9, 10), passives (6, 7, 8) and detractors (0 to 5).

Cluster Number	Average RFM	Corresponding NPS Scale	New_RFM_NPS Categories
5	111	0	Detractor
9	112	1	
8	221	2	
4	222	3	
10	223	4	
11	332	5	
3	333	6	Passive
7	334	7	
6	343	8	
2	344	9	Promoter
1	555	10	

Table 2. The New RFM_NPS Categories Using RFM and K-means Clustering Model

Active Customers. We have identified active customers as those whose time in days since their most recent transaction does not exceed the average number of days of the difference between the two most recent consecutive transactions across all customers. Active customers were considered to be those who had not churned and had company dealings in the form of transactions, as shown in Table 3. Each customer's recency value is compared to the average number of days of the difference between the two most recent consecutive transactions. If the value exceeds the average, the customer is labeled a churner; otherwise, the customer is considered loyal. In our case, the average of all calculated customers' is (88.09907). Any customer with a recency value of greater than (88 days) was considered to no longer be conducting business with the participant company and thus inactive.

Customer_ID	Recency	Recency 2	Diff_Recency	Loyalty_Status
0000111	680	701	21	Churner
000014C	552	604	52	Churner
0000824	33	34	1	Loyal
0000207	48	63	15	Loyal
0001117	15	29	14	Loyal
0000224	197	282	85	Churner
000036P	688	702	14	Churner
0003803	78	153	75	Loyal
0003874	68	78	10	Loyal
0000413	2	37	35	Loyal
0000435	287	322	35	Churner
0000438	554	706	152	Churner

Table 3. Active Customer Predictive Variables

Demographic Variables. We used the geographical location of customers, which appears within a transaction record. This resulted in limiting demographic input to the predictive model to customers' region. Customers are distributed across seven regions in the UK: south-east, south-west, Scotland, central-east, central-west, Northern Ireland and the Republic of Ireland.

Text-mining Model. We built on and extended Villarroel Ordenes et al.'s (2014) study using a linguistic-based text-mining model to analyze the open-ended customer comments in the survey. We employed sentence-level analysis to extract insightful information about the customer experience. We developed a set of linguistic patterns to analyze the comments sentence by sentence. In total, 213 linguistic patterns were developed across service processes (e.g. field, parts, sales, control center, invoicing, finance and credit, and overall satisfaction). Then each sentence was categorized as a complaint, a compliment or a suggestion. The complaint category is defined as customer-initiated expressions of dissatisfaction with the company (Landon 1980). A compliment is defined as the "expressions of personal praise that indicate the degree to which someone or something is liked" (Straight 1989, p. 37). Furthermore, Kraft and Martin (2001) defined positive feedback as a compliment, taking the form of acknowledgment or an expression of gratitude. Customer suggestion is defined as an idea offered by the customer for improving the service (Villarroel Ordenes et al. 2014). Our model provides a greater level of sophistication than the standard text analytics, with a high level of accuracy at 94 per cent (Rust and Cooil 1994; Villarroel Ordenes et al. 2014). A sentiment score for each comment was then calculated. Then we divided customers into 11 groups based on their average sentiment score. This was accomplished using the K-means segmentation algorithm. Again, a total of 11 clusters was chosen to mimic the 11-point scale of the NPS, which enabled the classification of customers' NPS promoter, passive and detractor categories. Similar to the newly calculated RFM_NPS score, the 11 customer groups were then sorted based on their average sentiment score. Satisfied customers were assigned to clusters 3 or 10. Neutral customers were clustered under 5, 7 and 11. Complainer customers were mapped to clusters 1, 2, 4, 6, 8 and 9. The percentage of customers falling into each category is based on the three years of data.

Clusters

Input (Predictor) Importance
1.0 0.8 0.6 0.4 0.2 0.0

Cluster	1	3	4	5	6	11	7	9	8	2	10
Label	Complainer	Satisfied	Complainer	Neutral	Complainer	Neutral	Neutral	Complainer	Complainer	Complainer	Satisfied
Description											
Size	37.6% (774)	21.3% (439)	14.2% (292)	7.4% (152)	5.0% (103)	4.2% (86)	3.8% (79)	3.4% (71)	1.3% (26)	1.0% (21)	0.8% (17)
Inputs	Avg_CommentScore -0.00	Avg_CommentScore 1.00	Avg_CommentScore -1.00	Avg_CommentScore 0.50	Avg_CommentScore -0.50	Avg_CommentScore 0.34	Avg_CommentScore 0.21	Avg_CommentScore -0.28	Avg_CommentScore -0.68	Avg_CommentScore -2.00	Avg_CommentScore 0.67

Figure 2. Average Sentiment Score Clusters

Customer Loyalty Model

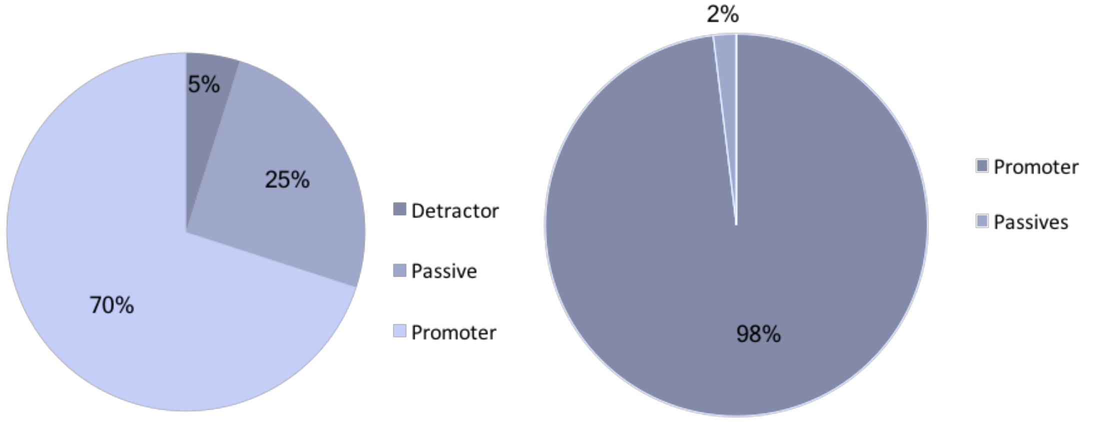
In classification modelling, sometimes referred to as supervised learning in data-mining and machine-learning communities, predictive variables are used to predict the values for a target field. In our study we used the predictive variables generated from the RFM model (recency, frequency and monetary scores), K-means model (the new RFM_NPS score), active customer category, demographic variable and the text-mining model (average sentiment score) to predict customer loyalty, which is our target variable. In particular, our customer loyalty model enabled us to classify customers as either churners or loyal customers based on these predictive indicators. "Churn" is a marketing term; however, there is some ambiguity when defining churn in the literature. A churner may be defined as a consumer who is either inactive in non-contractual settings (Rust, Lemon, and Zeithaml 2004; Tamaddoni Jahromi et al. 2010; Venkatesan and Kumar 2004) or has defected to the competitor in contractual settings (Glady, Baesens, and Croux 2009). Our study set out to utilize behavioral, demographic and attitudinal data to predict customer loyalty.

Our predictive model was built using neural and Bayesian network classification techniques. In marketing and service research, neural networks are generally viewed as powerful modeling techniques in customer retention (Datta, Pal, and Pal 2000). A Bayesian network is a graphical model that displays fields in a data set, and the probabilistic, or conditional, independencies between them. For example, a Bayesian network was used to predict customer life value (Baesens et al. 2004). We split the data into two steps: the creation or building of the model itself, and the actual usage of the model for future classification. In the context of this study we investigated the performance of the two classification algorithms (neural networks, Bayesian networks) for model generation. The accuracy results of these two algorithms were compared. The steps employed in the model's construction are three-fold. First, a training set, 60 per cent of the entire data, was used to develop a training model. Then we tested the model against new data in the construction stage, which formed 30 per cent of the original data. Here the model was fine-tuned to decrease the error of false predictions. Finally, the model was validated against the remaining 10 per cent of customer data. These three steps are performed to ensure the repeatability and validity of the prediction model (Tamaddoni, Stakhovych, and Ewing 2015).

RESULTS

Descriptive Analytics

This section focuses on analyzing the current loyalty measure (NPS) and identifying trends within the customers' feedback survey data. We analyzed the status of customer NPS values across the entire three years of data. This process was split into a series of analytic tasks that were performed on the attitudinal, behavioral and demographic data. First, we analyzed the loyalty status of customers with unchanged NPS values for the three years. The total number of customers surveyed over the three years was approximately 3,000 customers. The split between the NPS categories was 70 per cent promoters, 25 per cent passives and 5 per cent detractors across the three years (see Figure 3). This is typically interpreted as an excellent loyalty score for the company, with over half of the customers considered to be completely satisfied. Furthermore, the status of customers who had unchanged loyalty throughout the three-year period was also analyzed. Approximately 98 per cent of customers with unchanged loyalty are company promoters, while the other 2 per cent are passives and viewed as being indifferent to the company (see figure 3).



The Overall NPS status Across Three Years

The NPS Categories of Customers with Unchanged Loyalty Status

Figure 3. The Status of Customers with Unchanged NPS Values for Three Years was Observed

Second, our longitudinal analysis shows changes that occurred in the NPS over time, between two consecutive years (2012–13 and 2013–14). We analyzed those who had different NPS scores from one year to the next and split them into positive changes (from passive or detractor categories to promoter category), and negative changes (from promoter category to passive or detractor categories). The analysis shows that the positive change, which was 9 per cent at the end of 2013, decreased to 6 per cent at the end of 2014. Furthermore, the negative change increased from 5 per cent at the end of 2013 to 8 per cent at the end of 2014.

Third, based on the behavioral data, the company offers different services to its customers, which consists of sales (new, used, lease), product support (parts and service transaction types (workshop, field) and customer service agreement (CSA) transactions (parts and service transaction types). Two types of customer were identified: those who have a maintenance contract with the company, referred to as Customer Service Agreement (CSA) customers; and those who deal in a non-contractual setting, referred to as Product Support (PS) customers.

We found that customers who had made a sales transaction simultaneously made either a PS or CSA transaction. All CSA customers had made a PS transaction, with the exception of the case of one customer in 2014, who made a CSA transaction only. Moreover, there are very few customers who could be viewed as a “missed opportunity,” particularly as the number of customers having only sales transactions decreased over the years to almost a negligible fraction. Moreover, the customers with a PS transaction have been further broken down to show the type of transaction itself, namely whether a customer ordered parts or the delivery of a service. This can be seen in

Figure 4, where an ongoing trend of parts exceeding service transactions is observed. This is particularly interesting considering that the company operates within the services context.

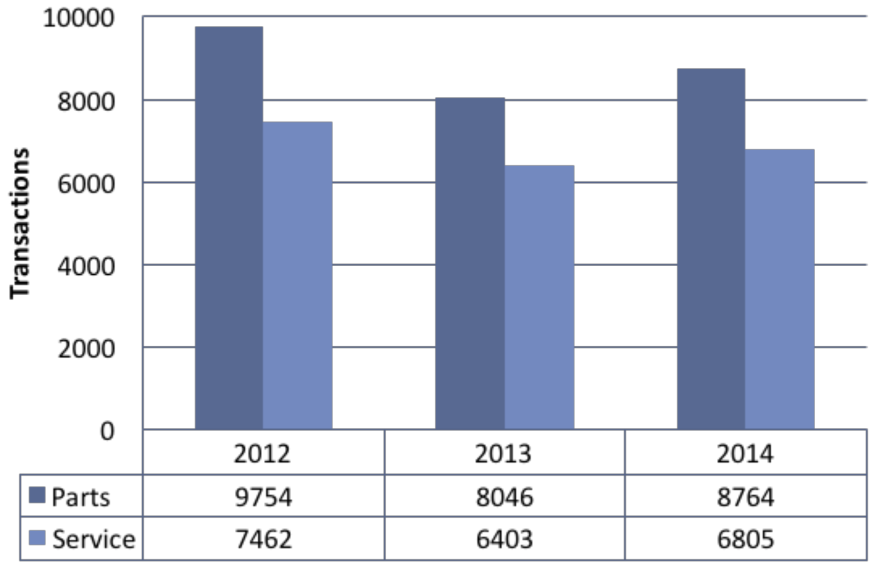


Figure 4. The Split Between PS Parts and Service for Sales Customers

Predictive Analytics

As previously discussed, the company does not link customers' NPS with their actual purchasing patterns. However, NPS was compared to actual customer spending patterns in our customer loyalty model. This makes it possible to understand *actual* customer behavior, instead of just customers' self-proclaimed loyalty referral intention. We limited our analysis to the PS customers as a result of the unavailability of survey data for CSA customers.

Following the RFM model, customers were divided into 11 groups based on their RFM scores. This was accomplished using the K-means segmentation algorithm, as previously discussed. Having developed an RFM score categorization similar to that of the Net Promoter Score (NPS), it then became possible to link the true underlying customer purchasing behavior with their self-proclaimed referral intention present in the survey's NPS category. This process entailed comparing customers' resulting *New_NPS_RFM*, shown in Table 2, with their actual NPS categories present in the customer feedback survey data, to determine whether or not they matched. In this analysis, we identified the misclassified NPS categories for the years 2012, 2013 and 2014, as shown in Table 4. For instance, in the *Promoter→Detractor* category, over the three years there were 291, 522 and 524 customers respectively with a survey NPS category of "promoter", when in fact their purchasing behavior was that of a "detractor." Similarly, in the *Detractor→ Promoter* category, there were 10 customers in 2012, 8 customers in 2013 and 3 customers in 2014, who were misclassified as "detractors" based on the survey and whose underlying spending pattern was that of a "promoter." In total, the percentage of NPS misclassification in 2012 was 72 and 85 per cent in 2013 and 82 per cent in 2014. The number of negative misclassifications exceeded the amount of positive ones. Moreover, the percentage of customers whose NPS score and actual spending pattern were dissimilar was over half of the entire customer base for each of the three years of data. Thus, this analysis suggests that using survey-based single-item metrics alone is inadequate. Customers' loyalty-based behaviors are multidimensional, and for this reason the attitudinal and behavioral data are essential components of measuring customer loyalty.

Furthermore, organizations should establish this link between customers' self-proclaimed referral intentions and their *actual* spending patterns by comparing the NPS rating with customer transactional data, as the analysis suggests.

Misclassified NPS	2012	2013	2014
Promoter → Detractor	291	522	524
Passive → Detractor	126	185	218
Promoter → Passive	135	147	104
Passive → Promoter	44	18	32
Detractor → Promoter	10	2	3
Detractor → Passive	9	8	10
% of customer misclassified	72%	85%	82%

Table 4. NPS Misclassification (New_NPS_RFM Vs. Survey NPS)

Another analysis based on the RFM model and active customer predictor was performed to confirm that the NPS does not necessarily correspond with behavior and purchasing patterns. In this analysis, we compared each customer's recency with the Average Diff_Recency. If the value exceeded the average, the customer was labeled a churner; otherwise, the customer was viewed as being loyal. In total, we found that customers' NPS score and purchasing patterns did *not* match across the three years. While customer loyalty based on the NPS score classified customers into 70 per cent promoters, 25 per cent passives and 5 per cent detractors, the customers' purchasing behavior analysis shows that only 54 per cent of customers were loyal and still buying from the company or using its services, and 46 per cent of the company's customers were churners. Thus, the NPS metric misinforms managers, diverting them away from marketing actions because organizations rely on a sample of respondents' memories of a service transaction. *Text-mining Model.* Our text-mining model automatically extracted customer opinions from the customer survey data. Our approach enabled us to capture and analyze the root causes of customer complaints expressed in verbatim comments. In addition, organizations can understand *why* customers are churning using details about the sources of friction. The model uncovers potentially vulnerable customers that NPS would have considered loyal and not requiring intervention strategies. Every comment was analyzed automatically using the developed linguistic patterns and then assigned to one of the seven high-level root-cause categories: communication, capability, parts, price value, process adherence, quality and service capacity. Root causes help managers to better understand and analyze the key reasons for friction. We found that 63 per cent of the company's customers were complainers, 22 per cent were satisfied and 15 per cent were neutral.

Top Churners Versus Sales Volume. Our customer loyalty predictive model shows how top churners are promoters in the NPS category. We selected 20 top churners based on their monetary indicator to evaluate their customer loyalty. Customers with high spending levels are more likely to stay, and changing patterns in spending may be an indicator of churn. As shown in Table 5, for example, the top churner is a customer (#408720) who used to buy products or services to the tune of £1,333,736. Similarly, our model reveals that the company was losing £6,073,696 by relying on the NPS alone to indicate whether the customer was still active. This is further confirmation that NPS is not a useful measure of customer loyalty. Furthermore, our model enables firms to identify the source of friction that has led to churning. For example, customer (#408720) had a problem with parts delivery communication, as noted: "Generally, it's a long way with getting feedback and part order and delivery…" Customer #53511, who bought £949,016 of services and products, is also a churner and complained about technician prices: *"Reduce the hourly rate a bit. I would like it to be towards 40 pounds per hour."*

Customer_ID	Monetary	Original NPS	Predicted Loyalty Status	Additional Comments	Customer Evaluation	Root Cause
408720	1333736.33	Promoter	Churner	generally, it's a long way with getting feedback and part order and delivery sometimes.	Compalints	Communication
53511	949016.22	Promoter	Churner	Reduce the hourly rate a bit. I would like it to be towards 40 pounds per hour.	Compalints	Price Value
343268	425586.3	Passive	Churner	Not really. They normally come when we ring them etc. They're always polite.	Neutral	NA
391588	363748.15	Promoter	Churner	No.	Neutral	
391511	361819.81	Detractor	Churner	They should improve on-site job training, improve quality control.	Compalints	Quality
298076	286064.93	Promoter	Churner	Not at the moment.	Neutral	NA
438714	281165.08	Passive	Churner	reduce the prices	Compalints	Price Value
53576	275402.57	Passive	Churner	Compay is too expensive	Compalints	Price Value
278502	230234.94	Passive	Churner	A bit more organised in some cases.	Suggestions	Process Adherence
298003	200174.48	Promoter	Churner	No it's [Company] just always very helpful, parts come through very quickly, which is good.	Compliments	Parts
019706Q	194904.28	Promoter	Churner	No, all good.	Compliments	NA
445635	187114.41	Detractor	Churner	No.	Neutral	NA
34076	177503.61	Promoter	Churner	No nothing	Neutral	NA
343237	170369.63	Promoter	Churner	Not really. I think experience with buying parts from other manufacturers, I don 't think they come close to your experience with the service. I suppose we deal with Company and, if not the best, they are one of the best with parts service and availa	Compalints	Parts
142518	169424.12	Promoter	Churner	We have dealt with them for a number of years and have never had a complaint with our local dealer.	Compliments	NA
343235	158213.02	Promoter	Churner	No, [Company] always do my work and im always very satisfied	Compliments	NA
047254C	155259.87	Promoter	Churner	pricing, expensive. Especially parts	Compalints	Price Value
400200	153958.76	Promoter	Churner	delivery service to go back to old system.	Compalints	Parts

Table 5. Top Churners Versus Sales Volume

Customer Loyalty Model Evaluation

We examined the classification power of the neural and Bayesian network techniques in the three phases of analysis (training, testing and validation). Table 6 shows the predictive accuracy of each of the training, testing and validation phases of the Bayesian and neural network classifiers algorithm. Our model is accurate across all three stages. While there is a slight increase in the percentage of error in the validation phase compared to the training phase, on average the model predicts correctly in 98 per cent of cases. This is a remarkably high percentage of accuracy and confirms the validity of the developed model with the suggested predictive variables as inputs.

Model	Max profit occurs in(%)	Lift {top 30%}	Overall accuracy training data set (60%)	Overall accuracy testing data set (30%)	Overall accuracy validation data set (10%)	No. of predictive fields used	Area under curve
Bayesian Network	35%	2.655	98.24%	98.21%	96.19%	6	99%
Neural Network	35%	2.655	97.98%	97.51%	97.62%	6	98%

Table 6. Comparing the Outperformed Classifier Techniques (Neural Network and Bayesian Network)

Typically, in big data projects, an evaluation is needed to compare the selected predictive models in order to choose the best model. Evaluation charts show how models perform in predicting particular outcomes. They work by sorting records based on the predicted value and confidence of the prediction, splitting the records into groups of equal size (quantiles) and then plotting the value of the business criterion for each quantile, from highest to lowest. Multiple models are shown as separate lines in the plot. In our case, we used a cumulative gains chart to evaluate and compare the performances of the neural and Bayesian networks (Burez and Van den Poel 2007; Linoff, Berry, and Bery 2004). For cumulative charts, higher lines indicate better models. In many cases, when comparing multiple models, the lines will cross, so that one model will be higher in one part of the chart and another will be higher in a different part of the chart. As shown in Figure 5, cumulative gains charts start at 0 per cent and end at 100 per cent, going from left to right. The gains chart in our model, using the two algorithms, rises steeply toward 100 per cent and then levels off.

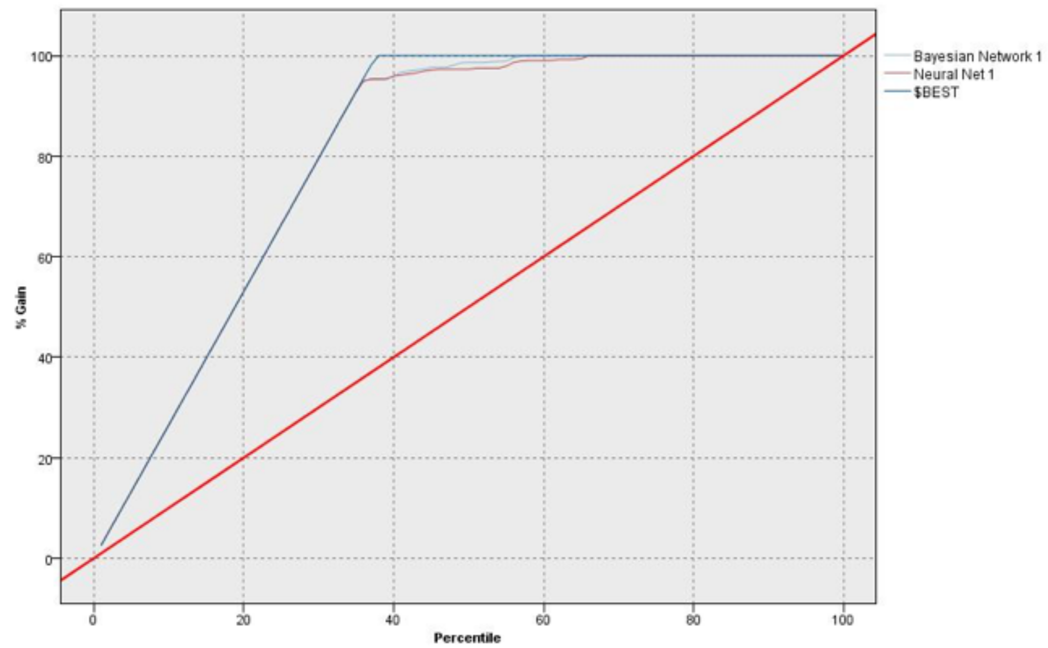


Figure 5. Cumulative Gains Chart of Neural Network and Bayesian Network

In our model, the Bayesian network provides insights into the probabilities of loyal customers and churners being in specific geographical areas. This identifies where they should focus their efforts to retain customers who are likely to leave the company. In our case, as shown in Table 7, the probabilities are that 22 per cent of south-western customers, 22 per cent in the south-east, 19 per cent in Scotland, 17 per cent in the central-east and 16 per cent of west-central customers will churn. By contrast, the model predicts that 20 per cent of customers in the central-west, 19 per cent in Scotland, 19 per cent in the central-east, 17 per cent in the south-west and 15 per cent in the south-east will be loyal.

Parents	Probabilities						
Loyalty Status	Central-East	Central West	South-East	Republic of Ireland	Scotland	South-West	North Ireland
Churner	17%	16%	22%	1%	19%	22%	3%
Loyal	19%	20%	15%	4%	19%	17%	6%

Table 7. Bayesian Network Customer Loyalty Probabilities for the Geographical Areas

DISCUSSION

Creating and maintaining customer loyalty is imperative for any firm's long-term success, which is fueled by managers' recognition of the benefits that a loyal customer base brings (Aksoy 2013). Loyal customers are much more likely to be retained, devote a higher share of spending with the firm, tend to be more satisfied (Oliver 1999) and engage in positive word-of-mouth about the firm (Reichheld and Sasser 1990). All of these desired behaviors are ultimately expected to translate into consistent cash flows and increased revenue for the business. However, this requires formal customer experience programs that monitor performance and guide improvement efforts (Homburg, Jozić, and Kuehnl 2015). Typically, this is obtained through customer surveys that track measures such as satisfaction, repurchase intention and word-of-mouth intention (Morgan and Rego 2006).

Managers still believe and use the survey-based measures as leading indicators of customers' behavior (e.g. retention, share-of-wallet allocation, word-of-mouth) (Aksoy 2013). Although these measures are important, they do not provide practical insights into *why* customers are churning and what managers should capture from customers' voices from textual comments to start to better allocate resources and improve customer loyalty. One of the central arguments in this paper is that organizations should begin capitalizing on the data points extracted from customer systems to simultaneously measure the attitudinal and behavioral components of loyalty. The novelty of our proposed customer loyalty model stems from addressing this specific gap and demonstrates that a data-driven approach to assessing and predicting customer loyalty is superior to relying on single-item metrics such as NPS. In particular, our model shows how firms can compare NPS scores with repurchasing behavior as a loyalty assessor, using predictive variables such as the RFM model, demographics, active customers and textual customer complaints, while the most popular performance measure, the NPS, completely disregards this. Clearly, customer loyalty measurement requires holistic loyalty-tracking initiatives. Based on the amalgamation of data sources, the actual underlying customer loyalty can be fully assessed and interpreted through the use of big data analytics. Furthermore, our model has the predictive ability to determine whether customers are likely to churn, thereby increasing the model's functionality. Over the three years, if the organization used NPS as an indicator for customer loyalty, over half of the customers were considered to be completely satisfied (70% promoters, 25% passives and 5% detractors). However, using our analysis, *New NPS_RFM*, as shown in Table 2, we identified many misclassified NPS categories, which has misled the company. For example, in 2013 and 2014 we found that approximately 500 customers were considered to be *detractors* when they were classified as *promoters* according to the NPS classification. Furthermore, the customers' purchasing behavior analysis shows that only 54 per cent of customers were loyal and still buying from the company and 46 per cent of the company's customers were churners (see Figure 6).

Thus, the NPS alone is not sufficiently accurate. If organizations want to understand *why* their customers churn, the answer could come from the verbatim comments provided in survey data or social media. Our text-mining model enables us to analyze the root causes of customer complaints, which expressed in free verbatim comments, uncover potentially vulnerable customers that the NPS would have considered loyal and not requiring intervention strategies. Furthermore, we transferred the complaint status of survey comments into an *average sentiment score* (satisfied, complainer, neutral) using the K-means clustering technique (satisfied, complainer, neutral). We found that 63 per cent of the company's customers were complainers, 22 per cent were satisfied and 15 per cent were neutral. The selected 20 top churners (Table 5) show how *churners* with high spending levels are *promoters* in the NPS category. These findings demonstrate that measuring customer loyalty is complex and our predictive analytics model can help organizations to measure it precisely.

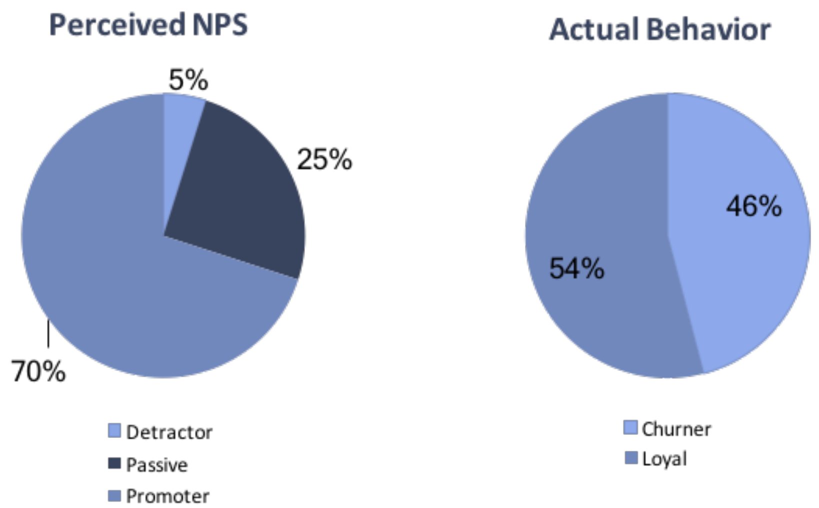


Figure 6. Customers' NPS Scores and Purchasing Patterns Do Not Match Across Three Years

MANAGERIAL IMPLICATIONS

Our paper has several important implications for managers. First, we provide a framework with which to integrate a multitude of customer data, by including demographic, behavioral and attitudinal customer data to assess and measure customer loyalty. The combination of these multiple data sources, along with the proposed techniques, measures customer loyalty more accurately than a single loyalty metric. Indeed, our study highlights the unreliability of the NPS or overall satisfaction as a single loyalty measure within B2B complex service organizations. Second, our predictive analytics model uses novel big data techniques to predict customer loyalty and to identify customers who have "churned" or are no longer conducting business with the organization. In particular, our approach underscores the benefits of setting up systematic multi-methods using a big data approach to capture and analyze customers' data across multiple customer systems. Third, the model incorporates a linguistic text-mining approach to determine the complaint status of each customer by mining survey textual feedback that divides customers into groups of complainers, neutral or satisfied. If organizations want to understand why customers churn, the answer is likely to be found in verbatim comments. Integration of the verbatim comments can identify and address customers' pain points within critical touchpoints of the service and unmask the root causes that are likely to affect customer loyalty. Fourth, the research provides a new method for organizations to begin capitalizing on their data points extracted from customer systems to simultaneously measure the attitudinal and behavioral components of loyalty. At the very least, organizations should compare their customers' self-proclaimed referral intention (NPS rating) with their actual spending patterns.

CONCLUSION

Single-question customer metrics such as the Net Promoter Score are popular tools that are used widely in practice in an attempt to measure customer loyalty. The literature suggested that growing reliance on a simple single customer metric is a very risky trend and that companies should adopt a more nuanced multidimensional approach to predicting customer behavior. This paper addresses this research gap and contributes to the assessment and measurement of customer loyalty and satisfaction using big data techniques within B2B complex service organizations. We developed a new model employing the recency, frequency and monetary (RFM) technique to transform customer transactional data into profitability scores, facilitating the categorization of customers based on purchasing behavior. Furthermore, the K-means clustering technique was also used to identify customers who have "churned" and are no longer conducting business with the organization. We extended the linguistic text-mining approach introduced by Villarroel Ordenes et al. (2014) to determine the complaint status of each customer by mining textual survey feedback dividing customers into groups of "complainers," "neutral" or "satisfied." Finally, our predictive analytics model used neural networks and Bayesian networks to predict customer loyalty. The results of the study highlight the gross inadequacies of using the NPS as a single loyalty metric, thereby supporting the criticism of the NPS in the literature. The result of the prediction phase is a model that is capable of correctly predicting 98 per cent of customers that are likely to churn, based on the results of the validation data set. This is a remarkably high percentage in terms of accuracy and ensures the validity of the developed model, with the suggested data parameters as predictive variables. On this basis, firms are discouraged from using the NPS as a single metric of loyalty assessment. Importantly, this novel approach offers practitioners a new way to utilize data more effectively and to provide rich insights and specific actions to improve customer experiences.

REFERENCES

- Aksoy, Lerzan (2013), "How do you measure what you can't define?," *Journal of Service Management*, 24 (4), 356–81.
- Baesens, Bart, Geert Verstraeten, Dirk Van den Poel, Michael Egmont-Petersen, Patrick Van Kenhove, and Jan Vanthienen (2004), "Bayesian network classifiers for identifying the slope of the customer lifecycle of long-life customers," *European Journal of Operational Research*, 156 (2), 508–23.
- Ballings, Michel, Dirk Van Den Poel, and Emmanuel Verhagen (2012), "Improving customer churn prediction by data augmentation using pictorial stimulus-choice data," *Advances in Intelligent Systems and Computing*, 171, 217–26.
- Bardauskaite, Indre (2014), "Loyalty in the Business-to-Business Service Context: A Literature Review and Proposed Framework," *Journal of Relationship Marketing*, 13 (1), 28–69.

- Baumann, Chris, Greg Elliott, and Suzan Burton (2012), “Modeling customer satisfaction and loyalty: survey data versus data mining,” *Journal of Services Marketing*, 26 (3), 148–57.
- Bitner, Mary Jo (1990), “Evaluating service encounters the effects of physical surroundings and employee responses,” *Journal of Marketing*, 54 (2), 69–82.
- Bolton, Ruth N and James H Drew (1991), “A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes,” *Journal of Marketing*, 55 (1), 1–9.
- Briggs, Elten, Timothy D. Landry, and Patricia J. Daugherty (2007), “Patronage in continually delivered business service contexts,” *Journal of Business Research*, 60 (11), 1144–51.
- Buckinx, Wouter and Dirk Van den Poel (2005), “Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting,” *European Journal of Operational Research*, 164 (1), 252–68.
- Burez, J. and D. Van den Poel (2009), “Handling class imbalance in customer churn prediction,” *Expert Systems with Applications*, 36 (3), 4626–36.
- Burez, Jonathan and Dirk Van den Poel (2007), “CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services,” *Expert Systems with Applications*, 32 (2), 277–88.
- Chapman, Pete, Julian Clinton, Randy Kerber, and Thomas Khabaza (2000), “CRISP-DM 1.0 Step-by-step Data Mining Guide,” [available at ftp://ftp.software.ibmcom/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf]
- Coussement, Kristof and Dirk Van den Poel (2008), “Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques,” *Expert Systems with Applications*, 34 (1), 313–27.
- Coussement, Kristof and Dirk Van den Poel (2009), “Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers,” *Expert Systems with Applications*, 36 (3), 6127–34.
- Cronin, J. Joseph and Steven A. Taylor (1992), “Measuring service quality: a reexamination and extension,” *Journal of Marketing*, 56 (3), 55–68.
- Cullinan, G. J. (1997), Picking them by their batting averages’ recency-frequency-monetary method of controlling circulation. Manual Release 2103, New York: Direct Mail/Marketing Association.
- Dagger, Tracey S, Jillian C Sweeney, and Lester W Johnson (2007), “A Hierarchical Model of Health Service Quality: Scale Development and Investigation of an Integrated Model ,” *Journal of Service Research* , 10 (2), 123–42.
- Datta, A., S. Pal, and N. R. Pal (2000), “A connectionist model for convex-hull of a planar set,” *Journal of Neural Networks*, 13 (3), 377–84.
- Eugene W. Anderson, Claes Fornell, Sanal K. Mazvancheryl (*2004*) Customer Satisfaction and Shareholder Value. *Journal of Marketing:* 68 (4),172-185.
- Eugene W. Anderson, Claes Fornell, and Rust, Roland T. (1997), “Customer Satisfaction , Productivity , and Profitability : Differences Between Goods and Services Customer Satisfaction , Productivity , and Profitability: Differences Between Goods and Services,” *Marketing Science*, 16 (2), 129–45.
- Eklof, Jan A., Peter Hackl, and Anders Westlund (1999), “On measuring interactions between customer satisfaction and financial results,” *Total Quality Management*, 10 (4), 514-522.
- Eskildsen, Jacob K. and Kai Kristensen (2011), “The gender bias of the Net Promoter Score,” in 2011 IEEE International Conference on Quality and Reliability, IEEE, 254–58.

- Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of marketing*, 56 (1), 6–21.
- Frederick Reichheld (2006), "The Microeconomics of Customer Relationships", *MIT Sloan Management Review*, 47 (2), 73-78.
- Glady, Nicolas, Bart Baesens, and Christophe Croux (2009), "Modeling churn using customer lifetime value," *European Journal of Operational Research*, 197 (1), 402–11.
- Grisaffe, Douglas B. (2007), "Questions about the ultimate question: Conceptual Considerations in Evaluating Reichheld's Net Promoter Score (NPS)," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 20, 36–53.
- Gordon H.G. McDougall and Terrence Levesque (2000), "Customer satisfaction with services: putting perceived value into the equation," *Journal of Services Marketing*, 14 (5), 392–410.
- Gupta, Sunil, D. Hanssens, B. Hardie, W. Kahn, V. Kumar, N. Lin, N. Ravishanker, and S. Sriram (2006), "Modeling Customer Lifetime Value," *Journal of Service Research*, 9 (2), 139–55.
- Gupta, Sunil and V. Zeithaml (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, 25 (6), 718–39.
- Hartmann, Philipp Max, Mohamed Zaki, Niels Feldmann, and Andy Neely (2016), "Capturing value from big data - a taxonomy of data-driven business models used by start-up firms," *International Journal of Operations & Production Management*, 36 (10), 1382-1406.
- Hogan, John E, Katherine N. Lemon, and Roland T. Rust (2002), "Customer Equity Management: Charting New Directions for the Future of Marketing ," *Journal of Service Research* , 5 (1), 4–12.
- Homburg, Christian, Danijel Jozić, and Christina Kuehnl (2015), "Customer experience management: toward implementing an evolving marketing concept," *Journal of the Academy of Marketing Science*, online, forthcoming.
- Hopmann, Jo ̈rg and Anke Thede (2005), "Applicability of Customer Churn Forecasts in a Non-Contractual Setting," in Innovations in Classification, Data Science, and Information System*s* Proceedings of the 27th Annual Conference of the Gesellschaft fu ̈r Klassifikation, Brandenburg University of Technology, Cottbus,Baier Daniel and Wernecke, Berlin, 330–37.
- Hosseini, Seyed Mohammad Seyed, Anahita Maleki, and Mohammad Reza Gholamian (2010), "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty," *Expert Systems with Applications*, 37 (7), 5259–5264.
- Huang, Ming-Hui and Roland T. Rust (2013), "IT-Related Service: A Multidisciplinary Perspective," *Journal of Service Research,*16 (3), 251-258
- Hughes, Arthur Middleto (1996), "Boosting response with RFM," *Marketing Tools*, 3 (3), 4.
- Ittner, C.D. and Larcker, D.F. (1996), "Measuring the impact of quality initiatives on firm financial performance," *Advances in the Management of Organizational Quality*, 1, 1–37.
- Kamakura, Wagner A, Vikas Mittal, Fernando de Rosa, and José Afonso Mazzon (2002), "Assessing the Service-Profit Chain," *Marketing Science*, 21 (3), 294–317.
- Kandampully, J.a, T.C.a Zhang, and A.b Bilgihan (2015), "Customer loyalty: A review and future directions with a special focus on the hospitality industry," *International Journal of Contemporary Hospitality Management*, 27 (3), 379–414.
- Keiningham, Timothy L., Bruce Cooil, Tor Wallin Andreassen, and Lerzan Aksoy (2007), "A Longitudinal Examination of Net Promoter and Firm Revenue Growth.," *Journal of Marketing*, 71 (3), 39–51.
- Keiningham, Timothy L, Lerzan Aksoy, Bruce Cooil, Tor Wallin Andreassen, and Luke Williams (2008), "A holistic examination of Net Promoter," *Journal of Database Marketing & Customer Strategy*

Management, 15 (2), 79–90.

- Klaus, Philipp and Stan Maklan (2013), “Towards a better measure of customer experience.,” *International Journal of Market Research*, 55 (2), 227–46.
- Kraft, Frederic B and Charles L. Martin (2001), “Customer compliments as more than complementary feedback,” *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 14 (1).
- Kristensen, Kai and Jacob Eskildsen (2011), “The validity of the Net Promoter Score as a business performance measure,” in 2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, IEEE, 970–74.
- Kristensen, Kai and Jacob Eskildsen (2014), “Is the NPS a trustworthy performance measure?,” *The TQM Journal*, 26 (2), 202–14.
- Lam, Shun Y., Venkatesh Shankar, M. Krishna Erramilli, and Bvsan Murthy (2004), “Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration From a Business-to-Business Service Context,” *Journal of the Academy of Marketing Science*, 32 (3), 293–311.
- Landon, E. Laird (1980), “The Direction of Consumer Complaint Research,” *Advances in Consumer Research*, 7, 335–38.
- Levesque, Terrence and Gordon H.G. McDougall (1996), “Determinants of customer satisfaction in retail banking,” *International Journal of Bank Marketing*, 14 (7), 12–20.
- Linoff, Gordon S., Michael J. Berry, and Gordon S.Linoff Michael J.A.Bery (2004), Data Mining Techniques for Marketing ,Sales and Customer Relationship Management, Management, Wiely, Indiana.
- Lloyd, Stuart P (1982), “Least Squares Quantization in PCM,” *IEEE Transactionson Information Theory*, 28 (2),129–37.
- Marbán, Oscar, Javier Segovia, Ernestina Menasalvas, and Covadonga Fernández-Baizán (2009), “Toward data mining engineering: A software engineering approach,” *Information Systems*, 34 (1), 87–107.
- McColl-Kennedy, Janet R., Anders Gustafsson, Elina Jaakkola, Phil Klaus, Zoe Radnor, Helen Perks, and Margareta Friman (2015), “Fresh Perspectives on Customer Experience,” *Journal of Services Marketing*, 29 (9), 430-435.
- Moorradian, Todd A and James M Olver (1997), “I Can’t Get No Satisfaction:" The Impact of Personality and Emotion on Postpurchase Processes,” *Psychology & Marketing*, 14 (4), 379–93.
- Morgan, Neil A., Eugene W. Anderson, and Vikas Mittal (2005), “Understanding firms’ customer satisfaction information usage,” *Journal of Marketing*, 69 (3), 131–151.
- Morgan, Neil A and Lopo Leotte Rego (2006), “The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance,” *Marketing Science*, 25 (5), 426–39.
- Nash, Dave, Doug Armstrong, and Michael Robertson (2013), “Customer Experience 2.0: How Data, Technology, and Advanced Analytics are Taking an Integrated, Seamless Customer Experience to the Next Frontier,” *Journal of Integrated Marketing Communications, 1-39*.
- Neslin, S. A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, and P. C. Verhoef (2006), “Challenges and Opportunities in Multichannel Customer Management,” *Journal of Service Research*, 9 (2), 95-112.
- Oliver, Richard L. and Linder, G. (1981), “Effect of satisfaction and its antecedents on consumer preference and intention,” *Advances in Consumer Research*, 8 (1), 88–93.
- Oliver, Richard L., and John E. Swan.(1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach." *Journal of Marketing* 53 (2), 21-35.

- Oliver, Richard L. (1999), "Whence Consumer Loyalty ?," *Journal of Marketing*, 63, 33–44.
- Oliver, Richard L. (1980), "A cognitive model of the antecedents and consequences of satisfaction decisions," *Journal of marketing research*, 17 (4), 460–70.
- Ostrom, Amy L., A Parasuraman, David E Bowen, Lia Patrício, and Christopher A Voss (2015), "Service Research Priorities in a Rapidly Changing Context," *Journal of Service Research* , 18 (2), 127–59.
- Pang, Bo and Lillian Lee (2008), "Opinion Mining and Sentiment Analysis," *Foundations and Trends® in Information Retrieval*, 2 (1–2), 1–135.
- Parasuraman, A. (2006), "Modeling Recovery Opportunities Customer-Managed," *Marketing Science*, 25 (6), 590–93.
- Parvaneh, Amin, Hossein Abbasimehr, and Mohammad Jafar (2012), "Integrating AHP and Data Mining for Effective Retailer Segmentation Based on Retailer Lifetime Value," *Journal of Optimization in Industrial Engineering* 11, 25–31.
- Pingitore, Gina, Neil A. Morgan, Lopo L. Rego, Adriana Gigliotti, and Jay Meyers (2007), "The single-question trap," *Marketing Research*, 19 (2), 8–13.
- Pollack, Birgit Leisen and Aliosha Alexandrov (2013), "Nomological validity of the Net Promoter Index question," *Journal of Services Marketing*, 27 (2), 118–29.
- Qin, Xiaoping, Shijue Zheng, and Ying Huang (2010), "Optimizated K-means algorithm and application in CRM system," *International Symposium on Computer, Communication, Control and Automation Optimizated* (3), 3–6.
- Rawson, Alex, Ewan Duncan, and Conor Jones (2013), "The Truth About Customer Experience: Interaction.," *Harvard Business Review*, 91 (11), 26.
- Reichheld, Frederick F.and Teal, T. (1996), *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*, Boston, MA: Harvard Business School Press.
- Reichheld, Frederick F. (1993), "Loyalty-Based Management," *Harvard business review*, 71 (2), 64-73.
- Reichheld, Frederick F. (2003), "The one number you need to grow.," *Harvard business review*, 88 (12), 46–54.
- Reichheld, Frederick F. and W.E. Earl Sasser (1990), "Zero Defections Quality Comes to Services," *Harvard Business Review,* 68 (5), 105–111.
- Rust, Roland T. and Bruce Cooil (1994), "Reliability Measures for Qualitative Data: Theory and Implications," *Journal of Marketing Research*, 31 (1), 1–14.
- Rust, Roland T., Katherine N. Lemon, and Valarie A Zeithaml (2004), "Return on Marketing," *Using Customer Equity to Focus Marketing Strategy*, 68 (1), 109–27.
- Rust, Roland T. and Ming-Hui Huang (2014), "The Service Revolution and the Transformation of Marketing Science," *Marketing Science*, 33 (2), 206–21.
- Rust, Roland T. and J Zahorik (1993), "Customer Satisfaction , Customer Retention , and Market Share," *Journal of Retailing*, 69 (2), 193–215.
- Sivadas, Eugene and Jamie L Baker-prewitt (2000), "An examination of the relationship between service quality , customer satisfaction , and store loyalty," *International journal of retail & Distribution Management*, 28 (2), 73–82.
- Spreng, Richard A, Gilbert D Harrell, and Robert D Mackoy (1995), "Service recovery : impact on satisfaction and intentions," *Journal of Service Marketing*, 9 (1), 15.
- Straight, H Stephen (1989), "Compliment-Rejection Versus Compliment-Avoidance : Listener-Based Versus Speaker-Based Pragmatic," *Language & Communication,* 9 (1), 35–47.
- Tamaddoni, Ali, Stanislav Stakhovych, and Michael Ewing (2015), "Comparing Churn Prediction

Techniques and Assessing Their Performance: A Contingent Perspective," *Journal of Service Research,* 19 (2), 123-141.

- Tamaddoni Jahromi, Ali, Mohammad Mehdi Sepehri, Babak Teimourpour, and Sarvenaz Choobdar (2010), "Modeling customer churn in a non-contractual setting: the case of telecommunications service providers," *Journal of Strategic Marketing*, 18 (7), 587–98.
- Taylor, Steven A. and Thomas L.Baker (1994), "An Assessment of the Relationship Between Service Quality and Customer Satisfaction in the Formation of Consumers' Purchase Intentions," *Journal of Retailing*, 70 (2), 163–78.
- Tirunillai, Seshadri and Gerard J Tellis (2014), "Mining Marketing Meaning from Online Chatter : Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51 (4), 463-479.
- Uncles, MD Mark D., Grahame R. Dowling, and Kathy Hammond (2003), "Customer loyalty and customer loyalty programs," *Journal of Consumer Marketing*, 20 (4), 294–316.
- Verhoef, Peter C., Katherine N Lemon, A Parasuraman, Anne Roggeveen, Michael Tsiros, and Leonard A Schlesinger (2009), "Customer Experience Creation : Determinants , Dynamics and Management Strategies," *Journal of Retailing,* 85 (1), 31–41.
- Venkatesan Rajkumar, V. Kumar (*2004*) A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy. *Journal of Marketing,* 68, (4),106-125.
- Villarroel Ordenes, Francisco, Babis Theodoulidis, Jamie Burton, Thorsten Gruber, and Mohamed Zaki (2014), "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach," *Journal of Service Research*, 17 (3), 278-295.
- Wiesel, T., Verhoef, Peter C., and Haan, Evert de (2012), "There Is No Single Best Measure of Your Customers," Harvard Business Review Blog, [Available at https://hbr.org/2012/07/there-is-no-one-best-measure-o]
- Wubben, Markus (2008), *Analytical CRM: Developing and Maintaining Profitable Customer Relationships in Non-contractual Settings*, New York, NY: Springer.
- Wübben, Markus and Florian v. Wangenheim (2016), "Instant Customer Base Analysis: Managerial Heuristics Often 'Get It Right,'" *Journal of Marketing*, 72 (3), 82–93.
- Yu, Xiaobing, Shunsheng Guo, Jun Guo, and Xiaorong Huang (2011), "An extended support vector machine forecasting framework for customer churn in e-commerce," *Expert Systems with Applications*, 38 (3), 1425–30.
- Yu, Yi-Ting and Alison Dean (2001), "The contribution of emotional satisfaction to consumer loyalty," *International Journal of Service Industry Management*, 12 (3/4), 234.
- Zeithaml, Valerie A., Leonard L. Berry, and A. Parasuraman (1996), "The behavioral consequences of service quality," *Journal of Marketing*, 60 (2), 31–46.