What social media sentiment tells us about why customers churn

James Lappeman, Michaela Franco and Victoria Warner
School of Management Studies, University of Cape Town, Cape Town, South Africa, and
Lara Sierra-Rubia
Department of analytics, DataEQ, Cape Town, South Africa

Abstract
Purpose – This study aims to investigate the factors that influence South African customers to potentially switch from one bank to another. Instead of using established models and survey techniques, the research measured social media sentiment to measure threats to switch.

Design/methodology/approach – The research involved a 12-month analysis of social media sentiment, specifically customer threats to switch banks (churn). These threats were then analysed for co-occurring themes to provide data on the reasons customers were making these threats. The study used over 1.7 million social media posts and focused on all five major South African retail banks (essentially the entire sector).

Findings – This study concluded that seven factors are most significant in understanding the underlying causes of churn. These are turnaround time, accusations of unethical behaviour, billing or payments, telephonic interactions, branches or stores, fraud or scams and unresponsiveness.

Originality/value – This study is unique in its measurement of unsolicited social media sentiment as opposed to most churn-related research that uses survey- or customer-data-based methods. In addition, this study observed the sentiment of customers from all major retail banks across 12 months. To date, no studies on retail bank churn theory have provided such an extensive perspective. The findings contribute to Susan Keaveney's churn theory and provide a new measurement of switching threat through social media sentiment analysis.

Keywords Switching, Social media sentiment, Banking behaviour, Churn

Paper type Research paper

Introduction

The purpose of this study was to understand the factors that cause South African retail bank customers to express the desire (or threat) to switch banks. This behaviour, once executed, is known as *customer churn* and is a significant topic in services research (Njenga, 2010; Hejazinia and Kazemi, 2014; Teichert *et al.*, 2020). By breaking the convention of using cross-sectional surveys and data mining, the study used online social media sentiment as a tool by which to categorise and measure the reasons behind potential churn behaviour. The study was conducted over 12 months, using South Africa's five major banks, accounting for over 90% of the South African retail banking market (Lechela, 2018). To date, no published study on the banking sector has used social media sentiment to test conventional churn theory across an entire industry over a 12-month period.

The existence of online social media platforms allows customers to easily and openly speak about their brand experiences and opinions to multiple people. This form of communication has opened a new stream of 21st-century research called electronic word of mouth (eWOM) (Lorenzo-Romero *et al.*, 2013; O'Brien, 2011; Moran *et al.*, 2014). This elective online sharing of information aids marketers in understanding the factors shaping the customer experience but also poses a major threat to businesses. Specifically, if customers can easily influence a wider online community, they can stimulate churn (Johnson, 2015; Verhagen *et al.*, 2013; Lappeman *et al.*, 2018; Hennig-Thurau *et al.*, 2004). This behaviour may increase with improved access to information on competing companies (Chua and Banerjee, 2013).

Most research on customer churn uses either cross-sectional surveys (Vyas and Raitani, 2014a, 2014b; Mavri and Loannou, 2008; Gerrard and Barton Cunningham, 2004) or the mining of customer data (Raguseo, 2018; De Caigny *et al.*, 2020). While both methods have been proven to provide reliable results, calls have been made to seek alternative methods to better understand churn (Gerrard and Barton Cunningham, 2004; Vyas and Raitani, 2014a). In particular, data mining (with various statistical approaches) is the most common way for companies to predict the likelihood of a customer choosing to switch (Amin *et al.*, 2017; Prasad and Madhavi, 2012;

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/0736-3761.htm**

Journal of Consumer Marketing
39/5 (2022) 385–403
Emerald Publishing Limited [ISSN 0736-3761]
[DOI 10.1108/JCM-12-2019-3540]

Received 13 December 2019
Revised 19 June 2020
20 December 2020
7 October 2021
Accepted 20 April 2022

Chitra and Subashini, 2011). These data-driven techniques, however, are limited by existing demographic and behavioural data owned and accessible by a company. Although not used in churn analysis, online sentiment analysis has slowly become a prevalent method for assessing online opinions and has grown in use in consumer-behaviour research (Salampasis *et al.*, 2014). This approach allows companies to capture raw (voluntary) data to assess sentiment and gather unstructured opinions (Srivastava and Gopalkrishnan, 2015; O'Brien, 2011; Gamon *et al.*, 2005). This study makes use of social media sentiment analysis to explore customers' threats to switch retail banks in South Africa.

Globalisation, digitisation and an increased number of firms have created highly competitive markets in most industries around the world. The South African retail banking sector is no exception. According to The Banking Association South Africa (2018), the largest five banks in South Africa are Standard Bank (with 10.6 million customers), Capitec (9 million), Absa (8.65 million) and Nedbank and First National Bank (FNB) (both with 7.7 million customers). Between 2016 and 2017, Capitec and Nedbank both acquired customers (Capitec gained over 1 million) and Standard Bank and Absa Bank lost 550,000 customers between them. The intense competition between South African banks, as well as two significant new entrants (Discovery and TymeBank), has made customer retention an increasingly pressing concern. Thus, this study was guided by the following research question:

RQ1. What factors influence South African customers to potentially switch (churn) from their banks to alternative service providers?

Exploring the factors that influence South African customers to potentially switch from one banking service provider to another will aid marketers in developing campaigns that focus on retention, as well as campaigns that focus on acquiring new customers. This study tests Keaveney's (1995) theoretical model in the light of the continued advancement of bank marketing and research tools like sentiment analysis.

Literature review

Customer churn

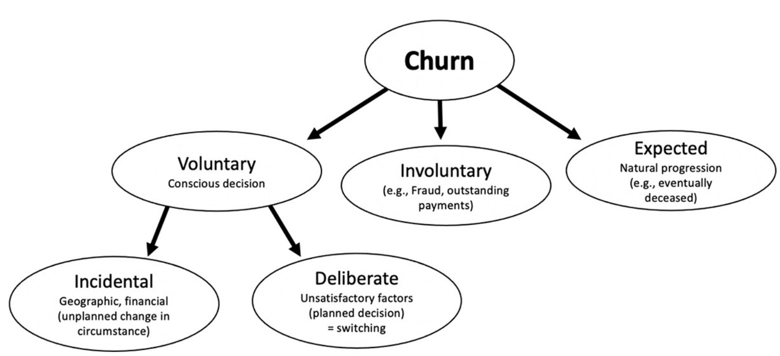
Customer churn (also known as *turnover*, *attrition* or *defection*) is used in business to describe the loss of customers (Marr, 2012). More specifically, churn occurs when a customer abandons a standing relationship with one company and switches their purchasing to another company (Hejazinia and Kazemi, 2014; Hadden *et al.*, 2005). The word *churn* embodies the notion of movement and is a composite of the words *change* and *turn* (Klepac, 2014). The customer churn rate is the probability that a customer will cease contact with a company (Madden *et al.*, 1999) and is often measured in industries like banking, telecommunications and insurance (Ahmad *et al.*, 2019).

Increased consumer knowledge and growing competition have caused firms to face heightened threats of losing customers (Nisand, 2017; Pires *et al.*, 2006). The means of attracting new customers is in constant balance with increasing concern about retaining existing customers, as the required strategies for these will differ (Njenga, 2010). Research has been conducted in various service industries in the hope of understanding phenomena like the waste of acquisition efforts, a reduction in long-term revenue and the narrowing of profit margins (Ascarza *et al.*, 2016; Hejazinia and Kazemi, 2014). In addition, churn management strategies have been designed to lower the risk of churn or win back defecting customers (Maga *et al.*, 2014).

Churn can be categorised into types presented in Figure 1. This diagram was adapted from Klepac *et al.*'s (2015) types of Churn model and divided churn into two main categories: voluntary and involuntary churn. Involuntary churn refers to customers being suspended by the firm or being forced to leave due to issues like outstanding payments or fraudulent behaviour (Klepac *et al.*, 2015). Companies can easily identify this type of churn and management theory has not focused on it.

Voluntary churn happens when a customer makes the conscious decision to leave their service provider (Hadden *et al.*, 2005). This type of churn is difficult to measure and is further broken down into two subcategories: deliberate and incidental churn. Incidental churn happens when a customer's circumstances change – usually financial or geographic circumstances – that leads to their no longer requiring or being able to access the service. Incidental churn only explains a small portion of a company's voluntary churn (Klepac *et al.*, 2015). Deliberate churn is a more significant portion of voluntary churn and is, therefore, prioritised by most churn-management solutions (Shaaban *et al.*, 2012). Deliberate churn takes place when a customer decides, for any number of reasons, to switch from their usual service provider to a competing provider

Figure 1 Holistic customer churn typology



Source: Adapted from Klepac (2014), Shaaban *et al*. (2012), Berry and Linoff (2004) and Reinartz and Kumar (2002)

(Amin *et al.*, 2017; Klepac *et al.*, 2015). In addition, expected churn occurs when a customer ceases to be a customer as part of a natural progression and can be accounted for at the beginning of the relationship (Berry and Linoff, 2004). For example, the death of a life insurance client or the growing up of a nappy-using baby are inevitable outcomes for the respective insurance and nappy companies.

Understanding the types of customers who churn and the reason these customers stop doing business with a company is imperative for any company seeking to manage its relationships with its customers and to improve its retention rates (Mamčenko and Gasimov, 2014; Irianto *et al.*, 2015). The factors that impact churn vary from sector to sector, which makes churn theory and measurement a broad subject for research (Bhatnagar *et al.*, 2019; Lovelock, 1983; Coulter and Ligas, 2000).

Churn measurement

While this study focused on social media sentiment, an understanding of existing churn measurement is important in identifying the gap that this research fills. Most churn-related literature focuses on churn prediction (Siemes, 2016) and is reported using statistical techniques to mine customer data. These data are often incorporated into sales forecasting (Trapero *et al.*, 2013; Fildes *et al.*, 2018), as well as customer relationship management (Huang, 2012; Audzeyeva *et al.*, 2012). In particular, customer-churn prediction (CCP) has grown in prominence as companies wish to know which customers are inclined to switch (De Caigny *et al.*, 2020; Ganesh *et al.*, 2000). This kind of churn prediction uses statistical modelling (of existing historical data) to create a scoring system to predict potential churn from existing customers (Ahmad *et al.*, 2019; Risselada *et al.*, 2010).

Adwan *et al.* (2014) noted that supervised machine-learning techniques are most widely investigated in the churn-prediction literature, but very few authors focus on which variables are most important for churn prediction (Faris, 2018). Neslin *et al.* (2006) compared different models of statistical prediction and confirmed that researchers must strive for better techniques and procedures in churn measurement. Specifically, alternative types of data were cited as a need for marketing-relevant variables that could help churn-reduction efforts rather than just identifying who is likely to churn. While looking at social media sentiment, although this study will not conclusively provide a list of new churn-related variables, the results do provide a window into the subject of variables that could lead to churn. Traditional statistical modelling in churn prediction tends to use models like logistic regression, survival models, neural networks and self-organising maps (Vafeiadis *et al.*, 2015; Verbeek, 2015; Klepac, 2014), all of which require existing customer data. One problem with using existing company data for churn prediction is that sometimes data are not time-stamped and, therefore, are not usable for tracking retention campaigns (Maldonado *et al.*, 2020). In addition, data must be available, and it is extremely unlikely that competitor or industry-wide company data will be made available for churn analysis.

One common use of predictive modelling is an early-warning system for customers who could switch. Approaches like the probit/logit approach and signalling approach are applied to multivariate models but need large samples (Klepac *et al.*, 2015). The former is often applied on a multivariate model, which allows testing of the statistical significance of explanatory variables. This type of model requires large samples and can only accommodate a limited number of explanatory variables so as to avoid multicollinearity (Klepac *et al.*, 2015). Statistical models can, however, lose their predictive power if competitors shift strategy (Klepac, 2014). This is where social media sentiment analysis can be useful as an exploratory guide to both company and industry sentiment.

Despite the growth in data mining, companies still struggle to extract meaningful information from textual data (Gandomi and Haider, 2015). Specifically, with churn prediction, the use of textual data is limited (De Caigny *et al.*, 2020) due to the challenges of preprocessing large quantities of textual data and also dimensionality (Schneider and Gupta, 2016). Examples of textual data algorithms in text-classification tasks include artificial, recurrent and convolutional neural networks (De Caigny *et al.*, 2018; Amiri and Daumé, 2016; West and Dellana, 2011). Models like Naïve Bayes and Support Vector Machine are also used in large-scale text classification analyses (Chumwatana and Wongkolkitsilp, 2019; Rashid, 2010). Other classification techniques, like hinge loss and logistic loss (Amiri and Daumé, 2016), as well as deep learning and logistic regression (El Kassem *et al.*, 2020), have also been used. Lexicon-based classifiers were used by Varsheney and Gupta (2014) and decision trees by Jahromi *et al.* (2014) among other rule-based decision-making techniques (Jahromi *et al.*, 2014).

Attitudinal measures have also been popular in measuring why customers switch (Bhatnagar *et al.*, 2019; Anderson and Sullivan, 1993). Preswitching behaviour, however, has been overlooked in the literature (Bhatnagar *et al.*, 2019; Vyas and Raitani, 2014a, 2014b), and this study addresses these antecedents. There have also been calls to better understand the role of attitudes before switching (Chadha and Bhandari, 2014). In addition, rich multi-company data for churn-related analyses are rare (De Caigny *et al.*, 2020; Huang *et al.*, 2012).

The study of social media, while not able to provide predictive modelling, is able to provide an industry-wide perspective on attitudes before switching. Furthermore, customers have different reasons for churning. Big data predictive models may be able to identify customers who are likely to switch, but not all churning customers do so for the same reasons and should not be treated in the same way (El Kassem *et al.*, 2020). There is a need for a model to predict churn customers and provide a strategy of retention depending on their churn factors. In addition, churn modelling should not be done using a one-size-fits-all approach but should be conceptualised according to business needs (Klepac, 2014).

Churn in the services sector

The service sector is generally characterised by long-term relationships between customers and companies. While loyalty levels vary depending on the type of service, customers have generally given their business to one service provider, with infrequent switching of providers (Hejazinia and Kazemi, 2014). The modern interconnected digital economy has, however, increased customer access to influential information about the companies they use and their competitors (O'Brien, 2011). This foundation has increased the threat of service-sector churn and increased research interest in the churn phenomenon.

To date, scholars have conducted research on churn within a range of service industries. Morgan and Dev (1994) researched customer churn in the hospitality industry, as did Soeini and Rodpysh (2012), who determined that a company's reputation plays a vital role in influencing customer churn. The hospitality industry was also considered by De Wit (2017), with a focus on business-to-business and non-contractual contexts. Njite *et al.* (2008) also researched churn in the hospitality industry with a focus on developing countries. Their research focused on factors driving churn and established that feedback loops, the external environment and internal systems were factors driving churn. Roos (1999) and Geppert (2002) studied churn in the US telecommunications sector. Similarly, Nimako (2012) studied telecommunications in Ghana and some large telecommunications churn studies have been conducted in Germany (Gerpott *et al.*, 2001), South Korea (Kim and Yoon, 2004) and China (Jia *et al.*, 2013). In Canada, auto repairs and hairstyling services were researched by Bansal *et al.* (2004) and Bansal *et al.* (2005), respectively. The studies found a range of factors that influenced churn, namely, affective commitment, continuance commitment and normative commitment. In addition, they observed push factors like low service quality, value, trust and commitment. Pull factors are mainly centred on the attractiveness of alternatives. Tavakoli *et al.* (2011) identified the reasons for customer churn by evaluating an insurance company's database. Marshall *et al.* (2011) compared churn in Eastern and Western financial markets. Colgate and Hedge (2001) addressed churn in the retail banking sectors of New Zealand and Australia. Both of these studies paid attention to the factors that influence customers to churn. Studies related to churn in retail banking are addressed separately in the subsection below.

Churn in retail banking

Customer switching in retail banking has been studied by various researchers. Stewart (1998) explained the process of customer exit, while Colgate and Hedge (2001) focused on the relevance of switching barriers in the banking sector. Keaveney (1995) was the first researcher to investigate a multi-industry context (Bhatnagar *et al.*, 2019). Naumann *et al.* (2010) used grounded theory to uncover motives for switching, and Chuang and Tai (2016) provided a macro-perspective on switching research by dividing switching research into three segments, namely, model exploration, model formation and model elaboration. Vyas and Raitani (2014a, 2014b) noted that churn in the financial-services sector has not received as much attention as other sectors (Friedman and Smith, 1993; Keaveney, 1995; Mittal and Lassar, 1998; Grace and O'Cass, 2003), although complaining is known to precede switching in retail banking (Stewart, 1998; Colgate and Hedge, 2001). Vyas and Raitani (2014a, 2014b) conducted their research in India and concluded that pricing, bank reputation, response to service failure, products, competition and service quality influenced churn.

Other notable studies include Sepehri *et al.* (2010), who conducted bank-related research on churn factors in Iran. The authors found that employees' attitudes affect customer churn directly. Colgate and Hedge (2001) investigated retail-banking churn in Australia and New Zealand and determined core-service failures, pricing and denial of services to be key factors of churn. Mavri and Loannou (2008) found that the quality of a bank's products and services – as well as a bank's reputation and credibility – were the most significant factors influencing customer churn. Gerrard and Barton Cunningham's (2004) research of Singaporean banks determined that pricing, service failures and inconvenience were the three most significant factors influencing customers to switch banks. Gerrard and Barton Cunningham (2004) also pointed out that churn in the banking sector was complex due to the locked-in effect. In Kenya, Ndung'u (2013) identified pricing, banking reputation and service quality as the most significant factors of customer churn.

Except for a study by Colgate and Hedge (2001), research on churn in the banking sector has mostly focused on which specific type of customer is most likely to churn. Oyeni and Adeyemi (2015), for example, focused their research on the Nigerian banking sector and used predictive data-mining techniques to determine the types of customers who would churn (as opposed to factors that influence churn). Research by Amin *et al.* (2017) and Prasad and Madhavi (2012) also made use of predictive data-mining techniques in the telecommunications and banking industries, respectively. Both studies emphasised CCP and argued that it is in a company's best interests to identify which customers are likely to churn since it can minimise costs and maximise profits. Similarly, Chitra and Subashini (2011) were able to determine which customers had a higher probability of churning in the banking sector by using predictive data-mining methods.

Theoretical underpinning

Several models have been developed to map the underlying factors that influence customers' switching behaviour (Bansal *et al.*, 2004; Colgate and Hedge (2001); Keaveney, 1995; Lees *et al.*, 2007) as well as their intentions to switch (Bansal *et al.*, 1999). Moreover, several studies have attempted to examine the constructs of churn (Mamčenko and Gasimov, 2014; Keramati *et al.*, 2008; Geppert, 2002); however, these determinants vary with each model, as shown in Table 1, which presents key aspects of each study.

Table 1 presents the key constructs of all the major existing churn/switching models in the services sector. To date, researchers have found over 20 factors to be influential for customer churn and switching behaviour. However, these constructs differ greatly depending on the model, location and industry. Historically, the most recurring factors are service quality, satisfaction, price and switching costs (Bansal, 2004; Geppert, 2002; Shin and Kim, 2008; Bansal *et al.*, 2005; Ahn *et al.*, 2006; Marshall *et al.*, 2011).

As the comprehensive review of the literature presented in Table 1 shows, there are no distinct churn or switching models focused solely on the banking industry. However, several studies have been carried out in the banking sector with the use of Keaveney's (1995) switching behaviour model, as represented in Figure 2.

The Keaveney model has been used in banking studies for over two decades, as shown in Table 2. Each of these studies was done in very different banking markets and all used cross-sectional survey methodologies (with widely varying sample sizes).

In spite of its prevalence, Vyas and Raitani (2014a, 2014b) noted that Keaveney's model did need further testing. In this

Journal of Consumer Marketing
Volume 39 · Number 5 · 2022 · 385–403

Table 1 Sample of existing theories and models of churning and switching

Theory/model	Source(s)	Industry	Constructs/factors
A synthesised model of consumer switching (SMCS)	Nimako (2012)	Telecommunications in Ghana	• Push pull mooring factors • Government policy
Prospect theory of consumer switching	Marshall *et al.* (2011)	Financial market in the Western and Eastern World	• Psychological and economic switching costs • Repeat purchase intention
Factors influencing switching intention	Shin and Kim (2008)	Mobile market in the USA	• Customer satisfaction • Service quality • Price • Perceived switching barriers: • Customer lock-in • Switching costs
Agency theory of customer switching	Abou Aish *et al.* (2008)	Advertising – agent relationship in Egypt	• Information asymmetry • Moral hazard • Diverse risk attitudes
General systems theory of consumer switching	Njite *et al.* (2008)	Hospitality industry in developed country	• Consumer • Regulatory subsystem • Feedback loop • External environment • Internal resources
Hanvali model customer churn	Keramati *et al.* (2008)	Mobile market in Iran	• Switching cost • Customer dissatisfaction • Customer status
A conceptual model for customer churn with mediation effects	Ahn *et al.* (2006)	Korean telecommunications service market	• Switching costs • Service usage • Customer dissatisfaction • Customer status • Customer-related variables
Push-pull mooring theory (PPM)	Bansal *et al.* (2005)	Auto-repairs and hairstyling services in Canada	• *Push factors*; low service quality, satisfaction, value, trust, commitment and high price • *Pull factors;* alternative attractiveness • *Mooring factors*; unfavourable attitude towards switching, unfavourable subjective norms, high switching cost, infrequent prior switching behaviour and low variety seeking
Three-component model of consumer commitment to service provider	Bansal *et al.* (2004)	Auto-repairs and hairstyling services in Canada	• Affective commitment • Continuance commitment • Normative commitment
The switching process model	Colgate and Hedge (2001)	Retail bank in Australia and New Zealand	• Core service failure • Pricing • Denied services
Service provider switching model (SPSM)	Bansal and Taylor (1999)	No specific industry	• Service quality • Perceived relevance • Satisfaction • Attitude towards switching • Subjective norms • Perceived switching cost • Switching intention

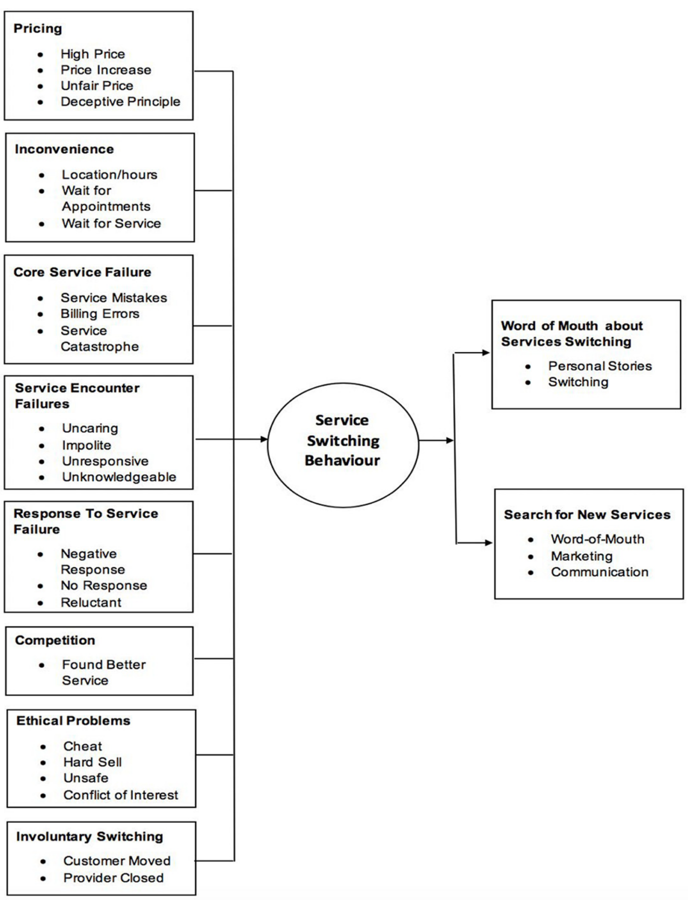
(***continued***)

Journal of Consumer Marketing
Volume 39 · Number 5 · 2022 · 385–403

Table 1

Theory/model	Source(s)	Industry	Constructs/factors
A catalytic switching model/SPAT	Roos (1999)	Telecommunications industry in the USA	• Triggers • Switching path • Switching determinants
Theory of planned behaviour (TPB)	Bansal and Taylor (1999)	No specific industry	• Service quality • Perceived relevance • Satisfaction • Attitude towards switching • Subjective norms • Perceived switching cost
A model of consumer's service switching behaviour	Keaveney (1995)	**Many industries**	• Inconvenience • Core service failure • Service encounter failures • Competition • Ethical problems • Involuntary switching • Pricing
Product importance model-based switching model	Morgan and Dev (1994)	**Hospitality in a developed country**	• Context • Control • Customer

Figure 2 Keaveney's (1995) conceptual framework to understand switching behaviour



Journal of Consumer Marketing
Volume 39 · Number 5 · 2022 · 385–403

Table 2 Past studies using Keaveney's (1995) model in the banking sector

Study component	Banking study 1	Banking study 2	Banking study 3	Banking study 4
Source	Mavri and Loannou (2008)	Gerrard and Barton Cunningham (2004)	Ndung'u (2013)	Vyas and Raitani (2014a)
Model	Keaveney model	Keaveney model	Keaveney model	Keaveney model
Location	Greece	Asia, Singapore	Kenya	India
Sample	NA	600	200	191
Method	Survey	Survey	Exploratory, survey, regression analysis	Survey
Timeframe	Cross-sectional study	Cross-sectional study	Cross-sectional atudy	Cross-sectional study
Most significant factors	Bank reputation and credibility, service quality, customer service	Pricing, service failures, inconvenience	Price, bank reputation, service quality	Pricing, bank reputation, response to service failure, service quality, competition

study, Keaveney's model is used as a theoretical underpinning, but the model was also tested by approaching the subject of churn through the lens of online sentiment gathered over a year, as opposed to cross-sectional survey data.

Methodology

This study involved a longitudinal 12-month analysis of consumer sentiment based on a sample of 1,726,543 social media posts. Specifically, 30,478 consumer posts that expressed the desire (or threat) of the post writer to leave their current bank were analysed for co-occurring topics to explore the unsolicited reasons behind the potential to churn. These reasons were coded and quantified to compare social media sentiment to churn theory derived from Keaveney's (1995) model. Unlike past research into antecedents of churn behaviour that mainly used a cross-sectional survey methodology, the purpose of this research was to understand churn purely by observing social media sentiment. This methodology is currently unique to the literature and heeds calls for non-survey methods and longer timeframes by which to study drivers of churn (Amiri and Daumé, 2016; Gerrard and Barton Cunningham, 2004; Vyas and Raitani, 2014a).

Sentiment analysis, also known as *opinion mining*, is the study of individuals' opinions, attitudes, feelings and sentiments on the basis of their written language (Liu, 2010; Cambria *et al.*, 2013). Sentiment analysis uses natural language processing (NLP), a form of artificial intelligence (AI) that helps computers interpret and understand human language (Manning and Schütze, 1999). NLP is widely studied in text mining, the process of studying large collections of written sources to acquire new information (Liu, 2010) and to identify and extract opinions from a text (Manning and Schütze, 1999). Tsytsarau and Palpanas (2012) differentiated between opinion mining and sentiment analysis by concluding that opinion mining is a method for identifying online texts that hold opinions, whereas sentiment analysis assesses texts for polarity. Liu, 2010 and Ravi and Ravi (2015) disagree and propose that the two concepts are synonymous. Since this analysis technique on modern media platforms is relatively new and there is no universally accepted definition for opinion mining, this study incorporates the essence of opinion mining within the umbrella of sentiment analysis (Cambria *et al.*, 2013).

Social media sentiment analysis is, therefore, the process of determining the source and polarity of an opinion (positive, negative or neutral opinion) and can also identify topics of online conversation (Liu, 2010). In this study, we use both aspects, whereby posts were initially sorted according to polarity (negative posts were used as the subsample for having the potential to be churn-related). By distinguishing positive and negative evaluations, opinions and emotions, a net sentiment score can be determined (Ravi and Ravi, 2015). This negative sentiment was then analysed for specific topics, which were coded and categorised. Data captured through mining has been found to be more affordable and reliable than survey data (Mirabeau *et al.*, 2013). In addition, using social media to gauge a targeted sample's expressed opinion is more authentic and voluntary, in contrast, to post hoc surveys (Klasnja *et al.*, 2017). A key reason for possible survey inaccuracy lies in the nature of predetermined questions that can potentially influence participants' answers and a lack of population representation (Golafshani, 2003). Further bias can occur within survey responses owing to memory or recall error (Buntain *et al.*, 2016).

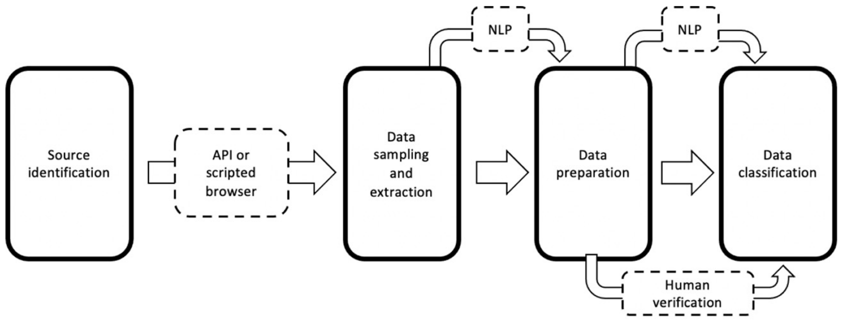
Huang *et al.* (2012) identified a systematic methodology for sentiment analysis related to churn prediction. Although our study is not directly a churn prediction modelling exercise, elements of this method are directly transferable. The authors identified phases in their approach, namely, data sampling, data preparation and classification. Data sampling randomly selects a set of customers with the relevant information, according to the definition of churn. This would align with a predictive model using existing customer data but not a sample of social media sentiment. The data preparation (preprocessing) phase includes data cleaning, feature extraction and normalisation steps. The prediction phase then predicts the potential behaviour of customers in the near future based on past behaviour (Huang *et al.*, 2012). Chumwatana and Wongkolkitsilp (2019) used a similar methodology in four steps, including source identification, data extraction, data preparation and data classification (Figure 3).

The methodology for this study was published after an extensive peer-review process (Lappeman *et al.*, 2020) and has been known to out-predict other methodologies (McKenzie and Swails, 2016). The specifics of the methodology are outlined below using a combination of the approaches of Huang *et al.* (2012) and Chumwatana and Wongkolkitsilp (2019).

Source identification, data sampling and extraction

The initial phases of social media sentiment analysis start with source identification and then the actual extraction of social

Figure 3 Methodology steps



Source: Adapted from Huang *et al.* (2012) and Chumwatana and Wongkolkitsilp (2019)

media posts from their relevant platforms. There are usually two options for this, namely, through an application programming interface (API) or by using a scripted browser (Chumwatana and Wongkolkitsilp, 2019). The support of opinion-mining firm Brandseye allowed us access to a paid subscription to the APIs of popular social media sites Twitter, Facebook and Instagram. The social networks were chosen due to their popularity as the largest social networking sites in South Africa. Facebook has approximately 16 million users in South Africa, followed by Twitter (8.3 million) and Instagram (6.6 million; Goldstuck, 2017). The chosen sample period was from 31 September 2017 to 1 August 2018 (Lassen *et al.*, 2014).

Using the API, keyword filters were used to identify posts that mentioned any of the five major banks included in this study. There is minimal consensus on how to choose a sample from these large-scale social media platforms (Lewis *et al.*, 2013). Proposals in the existing literature are to analyse as large a sample as needed to increase accuracy and representativity (Palguna *et al.*, 2015). Microblog posts that contained any of the words "Absa Bank", "First National Bank", "Standard Bank", "Capitec Bank" and "Nedbank" – as well as the respective acronyms and the bank's social media handles – were filtered from the three platforms and over 1,700,000 microblog posts were retrieved. This collection of microblog posts became the full sample data set (Tumasjan *et al.*, 2010; BrandsEye, 2018a, 2018b). This methodology is aligned with other research needing social media sentiment retrieval (Kim *et al.*, 2013; Nakov *et al.*, 2016; Sarlan *et al.*, 2014; Priyanka and Senthilkumar, 2016).

Data preparation and classification

Once the social media data were extracted, the data preparation and text classification process was possible (Chumwatana and Wongkolkitsilp, 2019). The first phase in data preparation was to test for polarity. This process allowed the sample to be narrowed down to a more focused group of negative blogs, which could be further assessed for churn-related themes. Negative micro-blog posts have been described as having "churny" content (Amiri and Daumé, 2016, p. 2566). The measure of polarity involved two stages to increase accuracy. Firstly, a similar process to the initial sampling was used, whereby negative and positive terms were used to isolate which posts contained negative sentiment (for example, *hate, angry* and *frustrated*). Neutral posts were discarded. Another round of analysis then isolated terms related to possible churn within the negative sample (for example, *I'm leaving*, *never again* and *goodbye*). The next stage involved micro-sampling through crowdsourced human raters (Lappeman *et al.*, 2020). Owing to NLP's known inability to fully understand the nuances of human conversation, manual validation techniques are sometimes used (Bifet and Frank, 2010; Tumasjan *et al.*, 2010). Specifically, NLP cannot easily process humour and sarcasm (Agarwal *et al.*, 2011; Wilson *et al.*, 2005). When this accuracy-enhancer uses large-scale human analysis to validate the results of large-scale machine analysis, it is known as crowdsourcing (Castle, 2018). Crowdsourcing was used in combination with NLP as done in the analysis of other major service brands like Spotify, Waze and Yelp (Castle, 2018). Kim *et al.* (2013) further emphasised the value of crowdsourcing when working with large data set validity, as have other authors (Lappeman *et al.*, 2020; Ghiassi *et al.*, 2013; Turney, 2002). The crowdsourcing process used trained human raters (from BrandsEye's database) who were sent a random sample of microblog posts extracted from the main sample. The raters then analysed the content according to three criteria: relevance (yes/no), sentiment (positive/neutral/negative) and topic, enabling them to assess both the posts and the actual outputs of the NLP algorithm, thus validating their accuracy or fixing classification errors. This aspect-level (topic) analysis took the sentiment analysis beyond just polarity and uncovered topics of conversation (Rana and Cheah, 2016; Agarwal *et al.*, 2011; Wilson *et al.*, 2005). Aspect-level sentiment analysis is usually divided into three main subtasks: aspect and opinion extraction, sentiment lexicon and opinion summarisation (Hu and Liu, 2004) to extract opinion-topics of interest from online text (Liu, 2010; Rana and Cheah, 2016). By using human raters, we were able to build a codebook of terms related to churn (also known as *aspect extraction*; Bifet and Frank, 2010; Miles, 1979). Posts were coded until data saturation (Dey, 1993; Crabtree and Miller, 1992; Miles and Huberman, 1994). This type of analysis is also called *intention mining* (Faed *et al.*, 2016). By developing the codebook, the NLP lexicon was taught by the human raters (Ghiassi *et al.*, 2013; Turney, 2002). More specifically, it allowed for identifying the sampled posts that were particular to the aspect of churn. In addition, co-occurring themes could also be identified and quantified.

Figure 4 shows how human raters would first-rate a post nominally by a *yes* or *no* score (Steps 1 and 2). The crowdsourcing provided verification with a 99% confidence level of the data analysis, creating a 1.75% margin for error and up to a 97% total accuracy rate (BrandsEye, 2018a, 2018b). The aspect extraction (Step 3) allowed for co-occurring topics to be coded based on the codebook.

Results

Once the data preparation and classification phase was complete, a total of 417,500 customers were assessed as having expressed a negative sentiment towards banks during the sample period. This is compared to the 286,606 expressing a positive sentiment (Table 3). This finding excludes repeat mentions by the same customer but assumes that each customer post is by a unique author and not the same author using a separate account.

The classification process allowed for the measurement of consumers who expressed the desire or intention to leave their bank (churn). Phrases like "keep it up and you won't see me anymore" (Figure 4) fell into this category. In total, 30,478 (7.3%) of the negative-sentiment-bearing posts were categorised as having a churn-related theme (Figure 5).

The aspect-level sentiment analysis revealed seven umbrella topics that were established from the codebook. These topics were customer service, ethics or reputation, pricing, staff or human resources, bank facilities, banking products and customer acquisition or retention. Within each umbrella topic, further categorisation was used to produce a more granular view of the sentiment expressed. The co-occurrence of the churn theme with another subtopic was now possible. For example, the following microblog post represents a customer with the intention to leave: "The transfer fees at Standard bank are way too high, I'm going to Capitec". This would be categorised under the umbrella topic "Price", thus, falling under the co-occurring churn topic of "Affordability". A total of 69 banking topics of conversation were identified and listed in the codebook. These are represented in Appendix.

Once the full set of subtopics were distilled, any mentions related to customer acquisition and retention were then connected to the subtopics. The result for each bank is displayed in Table 4.

Given that the negative-bearing churn posts can be assigned multiple topics, percentages do not equal 100%. Table 5 explains how many times a certain co-occurring topic was mentioned among the total churn posts, thereby presenting the industry averages for each topic.

Table 3 Volume of positive and negative sentiment-bearing posts

Total mentions	Positive mentions	Negative mentions
1,726,543	288,606	417,500

Figure 5 Number of microblog posts having potential churn-related conversations

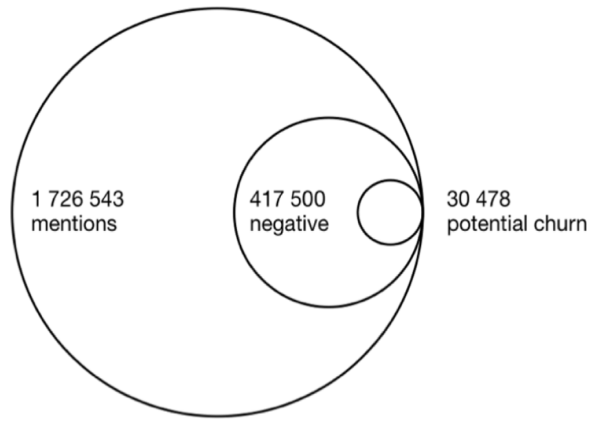


Figure 4 Illustrating the crowdsourcing interface

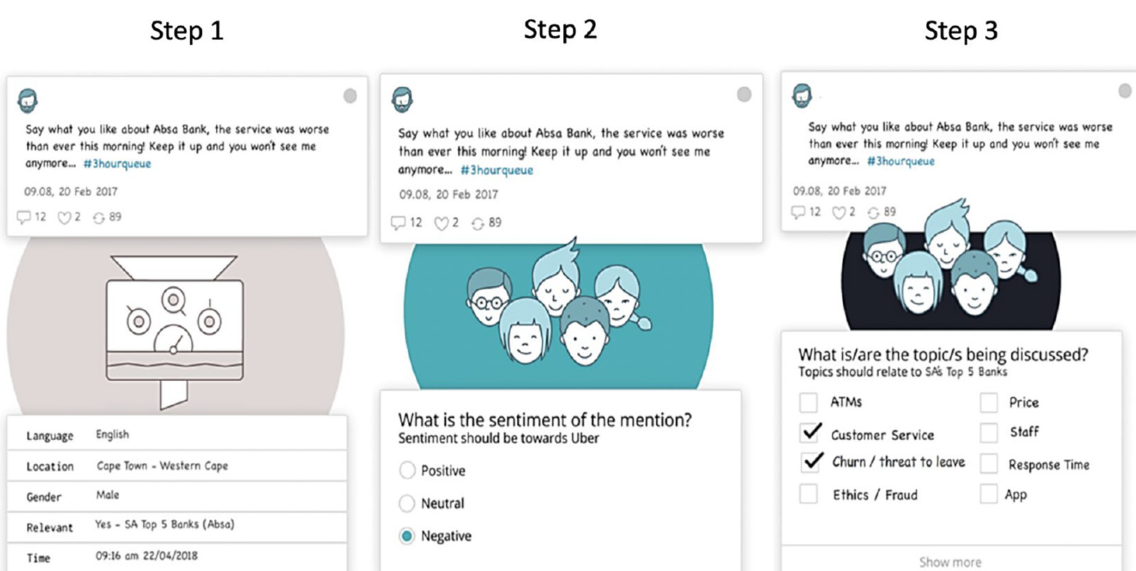


Table 4 Percentages of significance between co-occurring topics and churn within each bank

Co-occurring topics of churn	Absa (%)	Capitec (%)	FNB (%)	Nedbank (%)	Standard bank (%)
Turnaround time	34.1	24.2	33.9	25.4	33.9
Accusations of unethical behaviour	20.4	24.0	23.8	40.2	18.3
Billing or payments	25.8	24.0	27.5	16.0	23.4
Telephonic interactions	18.5	11.0	20.8	15.0	17.6
Branches or stores	19.1	17.7	11.2	16.1	17.5
Fraud, corruption or scams	15.5	20.5	20.4	3.9	11.0
No response received	14.0	10.5	12.2	9.1	12.1

Table 5 Most significant co-occurrence between negative sentiment and churn

Co-occurring topics of churn	Industry average (%)
Turnaround time	30.3
Accusations of unethical behaviour	25.3
Billing or payments	23.3
Telephonic interactions	16.6
Branches or stores	16.3
Fraud or scams	14.3
No response received from service provider	11.6
Turnaround time	30.3
Accusations of unethical behaviour	25.3
Billing or payments	23.3
Telephonic interactions	16.6
Branches or stores	16.3
Fraud or scams	14.3
No response received from service provider	11.6

The top three topics co-occurring with consumers threats to cancel were turnaround time, accusations of unethical behaviour and billing or payments. The factor that comprised the largest churn driver within South Africa's banking sector is turnaround time (30.3%). Turnaround time refers to the bank's ability to provide a banking product or service and complete customer requests within a reasonable time. The second most significant topic was accusations of unethical behaviour (25.3%). Most issues regarding this topic pertained to consumers alleging that banks intentionally allowed unauthorised debit orders to occur, as well as the undisclosed deduction of fees. Consumer perception was that the banks were doing this to secure reversal charges.

The third-largest topic was billing or payments (23.3%). Incorrect billing and debit orders prompted customers to threaten to leave their banks. These results highlight the customers' intentions to churn because they were unsatisfied with their bank's billing or payment methods. The fourth topic, telephonic interactions (16.7%), referred to unprofessional handling of clients' needs and a lack of adequate help over the telephone. Branches or stores (16.3%) was the fifth topic and included store layout and dissatisfaction with assistance provided within branches. Fraud or scams (14.3%) was sixth and came from the bank's unethical behaviour and lack of customer consideration. If a customer is a victim of fraud, there could be a natural desire to switch to a competing bank. Unresponsiveness (11.6%) was also found to be a significant influence on online churn conversations and reiterates how customers' intention to churn is significantly influenced by whether they get a response from their bank.

All other co-occurring topics of churn comprised less than 10% of negative churn conversations online. These less significant topics can be seen in the Figure 6, which is a visual representation of the results. In the rope diagram, the significance of each co-occurring topic has been linked back to the primary topic and the significance of each topic's relation to churn is represented by the band's thickness.

The results of the analysis answer the research question calling for an exploration into the factors that influence South African customers to switch banks.

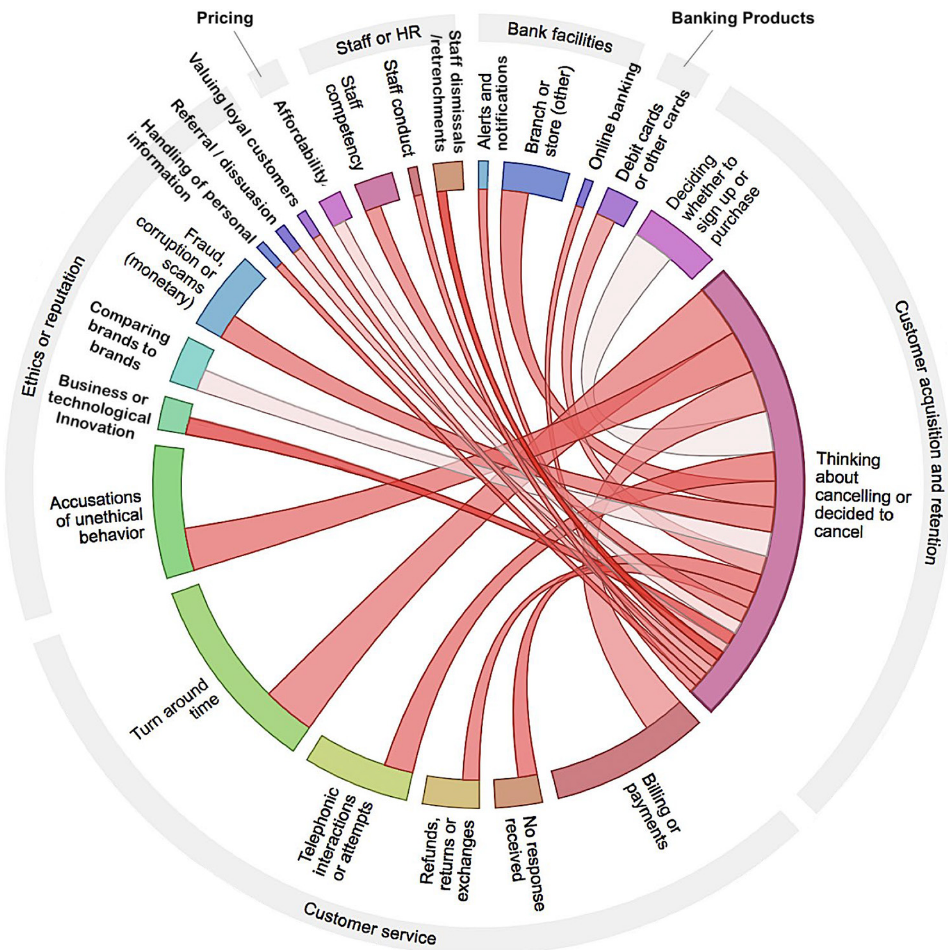
Discussion and implications

The results of this study directly challenge the churn factors presented by Keaveney's (1995) switching behaviour model. While the Keaveney model is a few decades old, it is still the most widely used basis for understanding churn in the services sector. Figure 7 illustrates the findings of the study and shows three significant improvements to Keaveney's model. Firstly, the churn factors were distilled through mining the sentiment of thousands of data points and not predetermined criteria in either a survey or predictive model. Secondly, the relative importance of each factor was assessed by volume and theme co-occurrence, and this enabled a more representative industry perspective to be quantified. Finally, the relative importance of each factor was assessed on a more granular level and provided a new mix of factors that most likely lead to churn.

By comparing the themes that co-occur with potential churn, a model derived from social media sentiment could then be compared to Keaveney's (1995) model. For example, Keaveney found core service failure (44.3%), inconvenience (34.1%) and pricing (29.9%) to be the three most significant factors. In contrast, this study did not find pricing to be within the top seven most significant factors, accounting for less than 10% of the samples' churn reasoning. Our results also showed turnaround time (30.3%), accusations of unethical behaviour (25.3%) and billing or payments (23.3%) to be the three largest contributors to churning (Keaveney's model resulted in only 7.5% of the sample depicting ethical problems as a reason for leaving).

We then compared the results of this study to the other studies that used Keaveney's (1995) churn model in various settings. This study's findings partially align with Mavri and Loannou (2008), who found that the quality of a bank's products and services – as well as a bank's reputation and credibility – were the most significant factors influencing customer churn. While there were some similarities, this study

Figure 6 Rope diagram of co-occurring churn topics

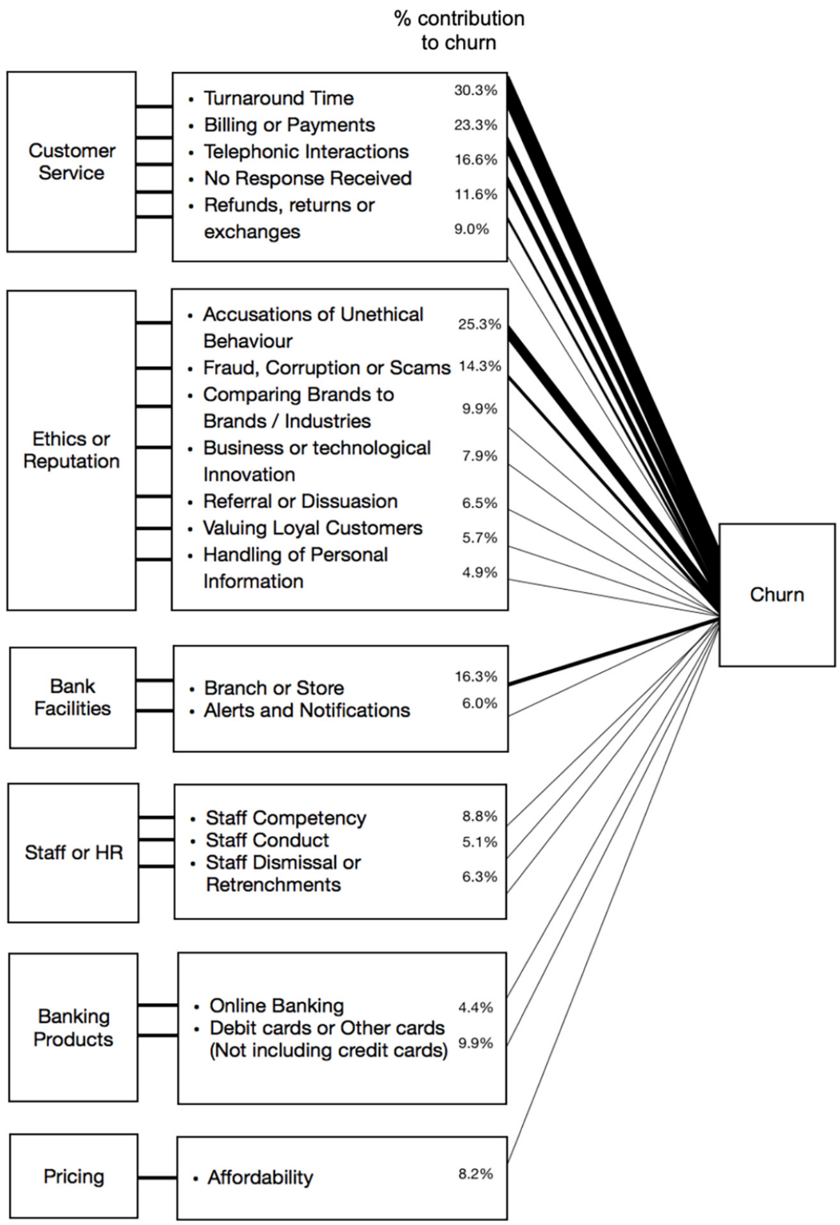


provided further detail about what aspects of these factors drove churn. From this study, it is clear that service quality and ethics or reputation, which Mavri and Loannou (2008) discovered as the most significant factors, were made up of a number of contributing subtopics or umbrella topics. This study provided greater detail into what part of a bank's customer service was driving churn, with turnaround time being the biggest influence (Figure 6). Accusations of unethical behaviour contributed the most to the factor bank reputation. Gerrard and Barton Cunningham (2004) determined that pricing, service failures and inconvenience were the three most significant factors influencing customers to churn in Singapore. Pricing proved to be the most influential factor when considering a bank's fees and switching costs. Ndung'u (2013) identified pricing as the most significant factor influencing customer churn in Kenya. Additionally, Ndung'u (2013) also identified that bank reputation and service quality of a banks' products and services were significant. Finally, Vyas and Raitani (2014a, 2014b) concluded that pricing, bank reputation, response to service failure, products, competition and service quality were factors that influenced churn in the Indian banking sector. However, pricing – as concluded repeatedly – was found to be the most significant factor overall.

In summary, according to the four key studies in question, the three most common factors influencing customers to churn were pricing, bank reputation (credibility) and service quality. In this analysis, the three most influential factors were turnaround time, accusations of unethical behaviour and billing or payments. These differences are material and provide a differentiated perspective on churn by showing that aspect-level sentiment analysis can significantly contribute to understanding the sentiment behind churn behaviour (Cambria *et al.*, 2013; Napitu *et al.*, 2018). While most churn-related studies use surveys and companies' existing customer data, this study took unsolicited social media sentiment from customers across an entire industry over 12 months. The results of this study prove that current churn-related research in retail banking has limitations and adding social media sentiment to churn prediction can help with functions like early identification of churn risks and risk-mitigation campaigns (Kumar and Yadav, 2020). The results brought into question a number of long-held beliefs about why customers leave their banks. This has implications for both strategic and customer experience management.

In a practical sense, banks could use this. As the methodology continues to improve, it can also allow banks to identify and intervene in online firestorms at an early stage to

Figure 7 Main themes co-occurring with threat to churn



mitigate damage (Lappeman *et al.*, 2021). Banks can also identify and observe general service-delivery trends that threaten to cause customer churn to both forecast attrition and also provide preventative structural measures in areas of weakness. In addition, banks can use this methodology to create an industry overview and competitor benchmarking in a way not possible before social media. Being able to see industry drivers not only benefits the pursuit of competitive advantage but also enables the building of better churn-prediction models with existing customer data. A more sophisticated and real-time view of the market can also be compared year-on-year. For retail banks, the relative importance of factors like turnaround time, reputation management (when accusations of unethical behaviour are made) and billing and payments must be stressed. Of note was the comparison to previous findings in the banking sector, which found pricing to be a large contributor to customer churn. In our study, pricing emerged as one of the least significant factors driving churn.

These findings have key implications for the banking sector in South Africa. Firstly, there is a major customer retention risk for South Africa's largest five banks, as shown by the volume of churn-related complaints. Equal to this, however, is the opportunity to acquire customers from competitors. Secondly, the main factors driving customers to churn in South Africa are turnaround time, accusations of unethical behaviour, billing and payments, telephonic interactions, branches or stores and fraud or scams. This particular perspective has established a unique window into the retail banking sector. As customer experience is a growing field of study, these findings are likely to have some multi-sector generalisations, like the importance of turnaround time, convenience and ethics, in consumers' decisions to switch. Subscription-based services, like many telecommunication services and media subscriptions, are likely to see similar trends, although further research is needed to confirm this.

Future research

Future research for this study can be grouped into a set of key avenues. Firstly, although this study has a large sample, there is currently no clear way to assess representativity. Not all banking customers use social media platforms and, of those that do, not all express their intention to switch on social media before doing so. The silent majority who leave the company without expressing dissatisfaction is outside the scope of this study, and its reasons are not uncovered. This, however, may be captured in exit interviews and other more standard churn-measurement surveys in parallel to a sentiment analysis. A larger sample over a longer (multi-year) time period may provide different results. In addition, understanding the context of a microblog post will also provide information. For example, posting in response to a promotion, service failure or retention campaign would yield a more nuanced picture. This, however, would need more innovative measurement techniques to be employed since a context is not immediately available from social media posts themselves.

Secondly, research into the role of key social media influencers in the study of churn is needed. Influencers are known to have the ability to override other microbloggers and can dominate social media conversations (Conover *et al.*, 2011). The potential to override the negative sentiment of other microbloggers or to increase the volume of eWOM has yet to be fully explored.

Thirdly, future studies can build in a demographic view of a microblogging sample to assess churn in a more refined way. Mavri and Loannou (2008) considered key demographics – age, lifestyle and gender – when determining the influential factors of churn. Big data prediction that profiles those most likely to churn uses these variables extensively (Prasad and Madhavi 2012), and the merging of sentiment and profiling is a stream for future analysis as predictive tools improve. This will initially be challenging since social media data are not always as clean as customer and survey data, as many users have pseudonyms. Social media sentiment analysis, however, can provide demographic detail like location, with more sophisticated tools becoming available to researchers periodically as the ecosystem develops.

Fourth, future researchers could analyse the role of timing and whether the threat to churn is part of the bandwagon effect during an nWOM firestorm or an unprovoked intention to switch. In short, researchers should consider customers' vulnerability to being influenced online and whether this vulnerability impacts what customers do and do not converse about online.

Finally, differences in region and sector are an important potential expansion of this research. For example, customers in a developing nation may consider the layout of branches and stores to be of greater significance due to the lower adoption of online banking. Similarly, poorer regions may find affordability to be a stronger driver of churn. As discussed in the conclusion, a similar analysis of multiple sectors will yield a nuance to the findings of this study and highlight differences in customer experience between sectors.

References

Abou Aish, E.M., Kortam, W.A. and Hassan, S.S. (2008), "Using Agency Theory in Understanding Switching Behavior in B2B Service Industries "I"", Working Paper No. 6, Faculty of Management Technology, German University of Cairo, Cairo.

Adwan, O., Faris, H., Jaradat, K., Harfoushi, O. and Ghatasheh, N. (2014), "Predicting customer churn in telecom industry using multilayer perceptron neural networks: modeling and analysis", *Life Science Journal*, Vol. 11 No. 3, pp. 75-81.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O. and Passonneau, R. (2011), "Sentiment analysis of twitter data", Proceedings of the Workshop on Language in Social Media, available at: https://dl.acm.org/citation.cfm?id=2021114 (accessed 20 April 2019).

Ahmad, A.K., Jafar, A. and Aljoumaa, K. (2019), "Customer churn prediction in telecom using machine learning in big data platform", *Journal of Big Data*, Vol. 6 No. 1, p. 28, doi: 10.1186/s40537-019-0191-6.

Ahn, J.H., Han, S.P. and Lee, Y.S. (2006), "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry", *Telecommunications Policy*, Vol. 30 Nos 10/11, pp. 552-568.

Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A. and Huang, K. (2017), "Customer churn prediction in the telecommunication sector using a rough set approach", *Neurocomputing*, Vol. 237, pp. 242-254, doi: 10.1016/j.neucom.2016.12.009.

Amiri, H. and Daumé, H. (2016), "Short text representation for detecting churn in microblogs", *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, pp. 2566-2572.

Anderson, E. and Sullivan, M. (1993), "The antecedents and consequences of customer satisfaction for firms", *Marketing Science*, Vol. 12 No. 2, pp. 125-143.

Ascarza, E., Iyengar, R. and Schleicher, M. (2016), "The perils of proactive churn prevention using plan recommendations: evidence from a field experiment", *Journal of Marketing Research*, Vol. 53 No. 1, pp. 46-60, doi: 10.1509/jmr.13.0483.

Audzeyeva, A., Summers, B. and Schenk-Hopp, K.R. (2012), "Forecasting customer behaviour in a multi-service financial organisation: a profitability perspective", *International Journal of Forecasting*, Vol. 28 No. 2, pp. 507-518.

Bansal, H.S. and Taylor, S.F. (1999), "The service provider switching model (spsm) a model of consumer switching behavior in the services industry", *Journal of service Research*, Vol. 2 No. 2, pp. 200-218.

Bansal, H.S., Irving, P.G. and Taylor, S.F. (2004), "A three-component model of customer to service providers", *Journal of the Academy of Marketing Science*, Vol. 32 No. 3, pp. 234-250, doi: 10.1177/0092070304263332.

Bansal, H.S., Taylor, S.F. and St. James, Y. (2005), "Migrating to new service providers: toward a unifying

framework of consumers' switching behaviors", *Journal of the Academy of Marketing Science*, Vol. 33 No. 1, pp. 96-115, doi: 10.1177/0092070304267928.
Berry, M.J.A. and Linoff, G.S. (2004), *Data Mining Techniques Second Edition – for Marketing, Sales, and Customer Relationship Management*, John Wiley & Sons, New York, NY.
Bhatnagar, S.B., Mishra, J.K. and Syed, A.A. (2019), "Customer disloyalty in retail banking services: attitudinal and behavioural dimensions", *Asia-Pacific Journal of Business Administration*, Vol. 11 No. 1, pp. 46-67.
Bifet, A. and Frank, E. (2010), "Sentiment knowledge discovery in twitter streaming data", *International Conference on Discovery Science*, Vol. 6332, pp. 1-15, available at: https://link.springer.com/chapter/10.1007/978-3-642-16184-1_1 (accessed 1 April 2019).
BrandsEye (2018a), "AI and human intelligence work together to achieve industry-leading sentiment accuracy", available at: www.brandseye.com/how-it-works/ (accessed 29 April 2019).
BrandsEye (2018b), "Topics: understand the issues driving public opinion", available at: www.brandseye.com/how-it-works/topics/ (accessed 29 April 2019).
Buntain, C., McGrath, E., Golbeck, J. and LaFree, G. (2016), "Comparing social media and traditional surveys around the Boston marathon bombing", *#Microposts*, Vol. 1691, pp. 34-41, available at: http://ceur-ws.org/Vol-1691/paper_02.pdf (accessed 4 March 2019).
Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013), "New avenues in opinion mining and sentiment analysis", *IEEE Intelligent Systems*, Vol. 28 No. 2, pp. 15-21, doi: 10.1109/MIS.2013.30.
Castle, N. (2018), "Crowd-sourced data and AI: a perfect pair", available at: www.datascience.com/blog/crowdsourced-data-ai (accessed 29 April 2019).
Chadha, S. and Bhandari, N. (2014), "Determinants of customer switching towards mobile number portability", *Paradigm*, Vol. 18 No. 2, pp. 199-219.
Chitra, K. and Subashini, B. (2011), "Customer retention in the banking sector using predictive mining technique", ICIT 2011, *The 5th International Conference on Information Technology*, available at: http://icit.zuj.edu.jo/icit11/PaperList/Papers/E-Commerce/612_Chitra.pdf (accessed 1 March 2019).
Chua, A.Y. and Banerjee, S. (2013), "Customer knowledge management via social media: the case of Starbucks", *Journal of Knowledge Management*, Vol. 17 No. 2, doi: 10.1108/13673271311315196.
Chuang, Y.F. and Tai, Y.F. (2016), "Research on customer switching behavior in the service industry", *Management Research Review*, Vol. 39 No. 8, pp. 925-939.
Chumwatana, T. and Wongkolkitsilp, K. (2019), "Using classification technique for customer relationship management based on Thai social media data", *Proceedings of 2019 11th International Conference on Computer and Automation Engineering (ICCAE 2019)*, pp. 7-11.
Colgate, M. and Hedge, R. (2001), "An investigation into the switching process in retail banking services", *International Journal of Bank Marketing*, Vol. 19 No. 5, pp. 201-212, doi: 10.1108/02652320110400888.
Conover, M.D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F. and Flammini, A. (2011), "Political polarization on twitter", *Fifth International AAAI Conference on Weblogs and Social Media*, available at: www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2847/3275 (accessed 28 August 2019).
Coulter, R.A. and Ligas, M. (2000), "The long good-bye: the dissolution of customer-service provider relationships", *Psychology and Marketing*, Vol. 17 No. 8, pp. 669-695.
Crabtree, B.F. and Miller, W.L. (1992), "A template approach to text analysis: developing and using codebooks", *Research Methods for Primary Care*, Vol. 3 No. 1, pp. 93-109, available at: https://psycnet.apa.org/record/1992-97742-005 (accessed 31 July 2019).
De Caigny, A., Coussement, K. and De Bock, K.W. (2018), "A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees", *European Journal of Operational Research*, Vol. 269 No. 2, pp. 760-772.
De Caigny, A., Coussement, K., De Bock, K.W. and Lessmann, S. (2020), "Incorporating textual information in customer churn prediction models based on a convolutional neural network", *International Journal of Forecasting*, Vol. 36 No. 4, pp. 1563-1578.
De Wit, L. (2017),"An Analysis of Non-Contractual Churn in the B2B Hotel Industry". Master's dissertation, Tilburg University Data Science: Business and Governance, Tilburg, available at: http://arno.uvt.nl/show.cgi?fid=144398 (accessed 10 September 2020).
Dey, I. (1993), *Qualitative Data Analysis: A User-Friendly Guide for Social Scientists*, Routledge & Kegan Paul, London.
El Kassem, E.A., Hussein, S.A., Abdelrahman, A.M. and Alsheref, F.K. (2020), "Customer churn prediction model and identifying features to increase customer retention based on user generated content", *International Journal of Advanced Computer Science and Applications*, Vol. 11 No. 5, pp. 522-531.
Faed, A., Chang, E., Saberi, M., Hussain, O.K. and Azadeh, A. (2016), "Intelligent customer complaint handling utilising principal component and data envelopment analysis (PDA)", *Applied Soft Computing*, Vol. 47, pp. 614-630.
Faris, H. (2018), "A hybrid swarm intelligent neural network model for customer churn prediction and identifying the influencing factors", *Information*, Vol. 9 No. 11, p. 288.
Fildes, R., Goodwin, P. and Önkal, D. (2018), "Use and misuse of information in supply chain forecasting of promotion effects", *International Journal of Forecasting*, Vol. 35 No. 1, pp. 144-156.
Friedman, M.L. and Smith, L.J. (1993), "Consumer evaluation processes in a service setting", *Journal of Services Marketing*, Vol. 2 No. 7, pp. 47-61.
Gamon, M., Aue, A., Corston-Oliver, S. and Ringger, E. (2005), "September pulse: mining customer opinions from free text", *International Symposium on Intelligent Data Analysis*, Springer, Berlin, Heidelberg, pp. 121-132.
Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.
Ganesh, J., Arnold, M.J. and Reynolds, K.E. (2000), "Understanding the customer base of service providers: an

examination of the differences between switchers and stayers", *Journal of Marketing*, Vol. 64 No. 3, pp. 65-87.

Geppert, C. (2002), "Customer churn management: retaining high-margin customers with customer relationship management techniques", *KPMG & Associated Yarhands Dissou Arthur/Kwaku Ahenkrah/David Asamoah*, available at: www.kpmg.com.au/Portals/0/ChurnMgmt-whitepaper_0226.pdf (accessed 3 March 2019).

Gerpott, T.J., Rams, W. and Schindler, A. (2001), "Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market", *Telecommunications Policy*, Vol. 25 No. 4, pp. 542-562, doi: 10.1016/S0308-5961(00)00097-5.

Gerrard, P. and Barton Cunningham, J. (2004), "Consumer switching behavior in the Asian banking market", *Journal of Services Marketing*, Vol. 18 No. 3, pp. 215-223, doi: 10.1108/08876040410536512.

Ghiassi, M., Skinner, J. and Zimbra, D. (2013), "Twitter brand sentiment analysis: a hybrid system using n-gram analysis and dynamic artificial neural network", *Expert Systems with Applications*, Vol. 40 No. 16, pp. 6266-6282, doi: 10.1016/j.eswa.2013.05.057.

Golafshani, N. (2003), "Understanding reliability and validity *in qualitative research*", *Quantitative, Qualitative, Comparative and Historical Methodologies Commons*, Vol. 8 No. 4, pp. 597-606, available at: https://nsuworks.nova.edu/cgi/viewcontent.cgi?referer=https://scholar.google.co.za/scholar?hl=en&as_sdt=0%2C5&q=A+reason+for+survey+data+inaccuracy+due+to+predetermined+questions+that+can+potentially+influence+participants'+answers&btnG=&httpsredir=1&article=1870&context=tqr/ (accessed 6 October 2019).

Goldstuck, A. (2017), "SA social media landscape 2017", (Research Report), World Wide Worx, South Africa.

Grace, D. and O'Cass, A. (2003), "Child care services: an exploratory study of choice, switching and search behaviour", *European Journal of Marketing*, Vol. 37 Nos 1/2, pp. 107-132.

Hadden, J., Tiwari, A., Roy, R. and Ruta, D. (2005), "Computer assisted customer churn management: state-of-the-art and future trends", *Computers & Operations Research*, Vol. 34 No. 10, pp. 2902-2917, doi: 10.1016/j.cor.2005.11.007.

Hejazinia, R. and Kazemi, M. (2014), "Prioritizing factors influencing customer churn", *Interdisciplinary Journal of Contemporary Research in Business*, Vol. 5 No. 12, pp. 227-236, available at: https://journal-archieves36.webs.com/227-236apr14.pdf (accessed 1 March 2019).

Hennig-Thurau, T., Gwinner, K.P., Walsh, G. and Gremler, D.D. (2004), "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?", *Journal of Interactive Marketing*, Vol. 18 No. 1, pp. 38-52, doi: 10.1002/dir.10073.

Hu, M. and Liu, B. (2004), "Mining and summarizing customer reviews", *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, ACM*, Vol. 1, pp. 168-177, doi: 10.1145/1014052.1014073.

Huang, C.Y. (2012), "To model, or not to model: forecasting for customer prioritization", *International Journal of Forecasting*, Vol. 28 No. 2, pp. 497-506.

Huang, B., Kechadi, M.T. and Buckley, B. (2012), "Customer churn prediction in telecommunications", *Expert Systems with Applications*, Vol. 39 No. 1, pp. 1414-1425.

Irianto, H., Haryono, T., Haryanto, B. and Riani, A.L. (2015), "The model of consumer's switching intention from conventional food to organic food: an experimental design study", *Mediterranean Journal of Social Sciences*, Vol. 6 No. 3, p. 588, doi: 10.5901/mjss.2015.v6n3s2p588.

Jahromi, A.T., Stakhovych, S. and Ewing, M. (2014), "Managing B2B customer churn, retention and profitability", *Industrial Marketing Management*, Vol. 43 No. 7, pp. 1258-1268.

Jia, Y.B., Zhang, Q., Ding, Q.Q. and Liu, D.L. (2013), "The study and realization of customer-churn model based on date mining in Telcom", *Applied Mechanics and Materials*, Vol. 336, pp. 2229-2232.

Johnson, J. (2015), "Influence of parents, peers, internet product search and visual social media on college students' purchase behavior: a mixed methods study", PhD. Dissertation, University of Nebraska, available at: http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1007&context=textilesdiss (accessed 4 March 2019).

Keaveney, S.M. (1995), "Customer switching behavior in service industries: an exploratory study", *Journal of Marketing*, Vol. 59 No. 2, pp. 71-82, doi: 10.1177/002224299505900206.

Keramati, A., Seyedin Ardebili, S.M. and Sohrabi, B. (2008), "Analysis churn customers: check the status of Iran's mobile operators with data mining techniques", *Journal of Management Sciences in Iran*, Vol. 14, pp. 63-91.

Kim, H.S. and Yoon, C.H. (2004), "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market", *Telecommunications Policy*, Vol. 28 Nos 9/10, pp. 751-765, doi: 10.1016/j.telpol.2004.05.013.

Kim, A.E., Hansen, H.M., Murphy, J., Richards, A.K., Duke, J. and Allen, J.A. (2013), "Methodological considerations in analyzing Twitter data", *Journal of the National Cancer Institute Monographs*, Vol. 47, pp. 140-146.

Klašnja, M., Barberá, P., Beauchamp, N., Nagler, J. and Tucker, J. (2017), "Measuring public opinion with social media data", *The Oxford Handbook of Polling and Survey Methods*, Oxford, Oxford.

Klepac, G. (2014), *Developing Churn Models Using Data Mining Techniques and Social Network Analysis*, IGI Global, Hershey, PA.

Klepac, G., Kopal, R. and Mrsic, L. (2015), "Early warning system framework proposal based on structured analytical techniques, SNA, and fuzzy expert system for different industries", *Handbook of Research on Artificial Intelligence Techniques and Algorithms*, IGI Global, Hershey, PA, pp. 763-796, doi: 10.4018/978-1-4666-7258-1.ch025.

Kumar, H. and Yadav, R.K. (2020), "Rule-based customer churn prediction model using artificial neural network based and rough set theory", *Advances in Intelligent Systems and Computing*, Vol. 1053, pp. 97-108.

Lappeman, J., Clark, R., Evans, J. and Sierra-Rubia, L. (2021), "The effect of nWOM firestorms on South African retail banking", *International Journal of Bank Marketing*, Vol. 39 No. 3, pp. 455-477, doi: 10.1108/IJBM-07-2020-0403.

Lappeman, J., Clark, R., Evans, J., Sierra-Rubia, L. and Gordon, P. (2020), "Studying social media sentiment using human validated analysis", *MethodsX*, Vol. 7, p. 100867, doi: 10.1016/j.mex.2020.100867.

Lappeman, J., Patel, M. and Appalraju, R. (2018), "Firestorm response: managing brand reputation during an nWOM firestorm by responding to online complaints individually or as a cluster", *Communicatio*, Vol. 44 No. 2, pp. 67-87, doi: 10.1080/02500167.2018.1478866.

Lassen, N.B., Madsen, R. and Vatrapu, R. (2014), "Predicting iPhone sales from iPhone tweets", *2014 IEEE 18th International Enterprise Distributed Object Computing Conference*, pp. 81-90, doi: 10.1109/EDOC.2014.20.

Lechela, N. (2018), "Five banks ranked amongst SA's strongest banks", *Fin24*, 24 May 2018, available at: www.fin24.com/Companies/Financial-Services/five-banks-ranked-among-sas-strongest-brands-20180524 (accessed 1 March 2019).

Lees, G., Garland, R. and Wright, M. (2007), "Switching banks: old bank gone but not forgotten", *Journal of Financial Services Marketing*, Vol. 12 No. 2, pp. 146-156, doi: 10.1057/palgrave.fsm.4760070.

Lewis, S.C., Zamith, R. and Hermida, A. (2013), "Content analysis in an era of big data: a hybrid approach to computational and manual methods", *Journal of Broadcasting & Electronic Media*, Vol. 57 No. 1, pp. 34-52, doi: 10.1080/08838151.2012.761702.

Liu, B. (2010), "Sentiment analysis and subjectivity", *Handbook of Natural Language Processing*, Vol. 2 No. 10, pp. 627-666, available at: http://people.sabanciuniv.edu/berrin/proj102/1-BLiu-Sentiment%20Analysis%20and%20Subjectivity-NLPHandbook-2010.pdf (accessed 1 March 2019).

Lorenzo-Romero, C., Constantinides, E. and Alarcón-del-Amo, M.D.C. (2013), "Web aesthetics effects on user decisions: impact of exposure length on website quality perceptions and buying intentions", *Journal of Internet Commerce*, Vol. 12 No. 1, pp. 76-105, doi: 10.1080/15332861.2013763695.

Lovelock, C.H. (1983), "Classifying services to gain strategic marketing insights", *Journal of Marketing*, Vol. 47 No. 3, pp. 9-20.

McKenzie, D. and Swails, B. (2016), "Psychic software? How a small social media company predicted Donald Trump's victory", *CNN*, available at: http://psychic8.blogspot.com/2016/11/psychic-software-how-small-social-media_41.html (accessed 12 May 2019).

Madden, G., Savage, S.J. and Coble-Neal, G. (1999), "Subscriber churn in the Australian ISP market", *Information Economics and Policy*, Vol. 11 No. 2, pp. 195-207, doi: 10.1016/S0167-6245(99)00015-3.

Maga, M., Canale, P. and Bohe, A. (2014), U.S. Patent No. 8,712,828, Accenture Global Services Ltd: Churn Prediction and Management System.

Maldonado, S., López, J. and Vairetti, C. (2020), "Profit-based churn prediction based on minimax probability machines", *European Journal of Operational Research*, Vol. 284 No. 1, pp. 273-284.

Mamčenko, J. and Gasimov, J. (2014), "Customer churn prediction in mobile operator using combined model", *Proceedings of the 16th International Conference on Enterprise Information Systems*, pp. 233-240, doi: 10.5220/0004896002330240.

Manning, C.D. and Schütze, H. (1999), *Foundations of Statistical Natural Language Processing*, MIT Press, Cambridge, MA.

Marr, B. (2012), *Key Performance Indicators (KPI): the 75 Measures Every Manager Needs to Know*, Pearson, London.

Marshall, R., Huan, T.C.T., Xu, Y. and Nam, I. (2011), "Extending prospect theory cross-culturally by examining switching behavior in consumer and business-to-business contexts", *Journal of Business Research*, Vol. 64 No. 8, pp. 871-878, doi: 10.1016/j.jbusres.2010.09.009.

Mavri, M. and Loannou, G. (2008), "Customer switching behaviour in Greek banking services using survival analysis", *Managerial Finance*, Vol. 34 No. 3, pp. 186-197, doi: 10.1108/03074350810848063.

Miles, M.B. (1979), "Qualitative data as an attractive nuisance: the problem of analysis", *Administrative Science Quarterly*, Vol. 24 No. 4, pp. 590-601, doi: 10.2307/2392365.

Miles, M.B. and Huberman, A.M. (1994), *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed., Sage, Thousand Oaks, CA.

Mirabeau, L., Mignerat, M. and Grange, C. (2013), "The utility of using social media networks for data collection in survey research", *Thirty Fourth International Conference in Information Systems*, available at: https://pdfs.semanticscholar.org/e743/a15edede1f70085f58813178a6078acc3827.Pdf (accessed 29 April 2019).

Mittal, B. and Lassar, W.M. (1998), "Why do customers switch? The dynamics of satisfaction versus loyalty", *Journal of Services Marketing*, Vol. 12 No. 3, pp. 177-194.

Moran, G., Muzellec, L. and Nolan, E. (2014), "Consumer moments of truth in the digital context: how 'search' and 'e-word of mouth' can fuel consumer decision making", *Journal of Advertising Research*, Vol. 54 No. 2, pp. 200-204, doi: 10.2501/JAR-54-2-200-204.

Morgan, M.S. and Dev, C.S. (1994), "An empirical study of brand switching for a retail service", *Journal of Retailing*, Vol. 70 No. 3, pp. 267-282, doi: 10.1016/0022-4359(94)90036-1.

Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F. and Stoyanov, V. (2016), "Sentiment analysis in twitter", *Proceedings of the 10th international workshop on semantic evaluation, SemEval-2016*, pp. 1-18, doi: 10.18653/v1/S16-1001.

Napitu, F., Bijaksana, M.A., Trisetyarso, A. and Heryadi, Y. (2018), "Twitter opinion mining predicts broadband internet's customer churn rate", *2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), Phuket*, 2017, pp. 141-146, doi: 10.1109/CYBERNETICSCOM.2017.8311699.

Naumann, E., Haverila, M., Sajid Khan, M. and Williams, P. (2010), "Understanding the causes of defection among satisfied B2B service customers", *Journal of Marketing Management*, Vol. 26 Nos 9/10, pp. 878-900.

Ndung'u, A.W. (2013), "Modeling of churn behavior of bank customers using logistic regression", Masters dissertation.

Neslin, S.A., Gupta, S., Kamakura, W., Lu, J. and Mason, C.H. (2006), "Defection detection: measuring and understanding the predictive accuracy of customer churn models", *Journal of Marketing Research*, Vol. 43 No. 2, pp. 204-211.

Nimako, S.G. (2012), “Linking quality, satisfaction and behaviour intentions in Ghana’s mobile telecommunication industry”, *European Journal of Business and Management*, Vol. 4 No. 7, pp. 1-17, available at: www.researchgate.net/profile/Simon_Nimako/publication/267802960_Linking_Quality_Satisfaction_and_Behaviour_Intentions_in_Ghana’s_Mobile_Telecommunication_Industry/links/54aff1e00cf220c63ccdc95d.pdf (accessed 4 March 2019).
Nisand, P. (2017), “How does globalization impact through retailing? A case study of Kerala”, *International Journal of Advanced Research in Management and Social Sciences*, Vol. 6 No. 2, pp. 12-26.
Njenga, K. (2010), “The increasing focus on managing relationships and customer retention”, *Journal of Language, Technology & Entrepreneurship in Africa*, Vol. 2 No. 1, pp. 85-92, available at: www.ajol.info/index.php/jolte/article/view/51992/40627 (accessed 4 March 2019).
Njite, D., Kim, W.G. and Kim, L.H. (2008), “Theorizing consumer switching behavior: a general systems theory approach”, *Journal of Quality Assurance in Hospitality & Tourism*, Vol. 9 No. 3, pp. 185-218, doi: 10.1080/15280080802412701.
O’Brien, C. (2011), “The emergence of the social media empowered consumer”, *Irish Marketing Review*, Vol. 21 Nos 1/2, pp. 32-40.
Oyeniyi, A.O. and Adeyemo, A.B. (2015), “Customer churn analysis in banking sector using data mining techniques”, *African Journal of Computing & ICT*, Vol. 8 No. 3, pp. 165-174.
Palguna, D.S., Joshi, V., Chakaravarthy, V., Kothari, R. and Subramaniam, L.V. (2015), “Analysis of sampling algorithms for twitter”, *Twenty-Fourth International Joint Conference on Artificial Intelligence*, available at: www.aaai.org/ocs/index.php/IJCAI/IJCAI15/paper/view/10690 (accessed 27 April 2019).
Pires, G., Stanton, J. and Rita, P. (2006), “The internet, consumer empowerment and marketing strategies”, *European Journal of Marketing*, Vol. 40 Nos 9/10, pp. 936-949, doi: 10.1108/03090560610680943.
Prasad, U.D. and Madhavi, S. (2012), “Prediction of churn behavior of bank customers using data mining tools”, *Business Intelligence Journal*, Vol. 5 No. 1, pp. 96-101, available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.456.4094&rep=rep1&type=pdf#page=100 (accessed 1 March 2019).
Priyanka, Y. and Senthilkumar, R. (2016), “Sampling techniques for streaming dataset using sentiment analysis”, *Recent trends in information technology (ICRTIT) 2016 International Conference*, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569580.
Raguseo, E. (2018), “Big data technologies: an empirical investigation on their adoption, benefits and risks for companies”, *International Journal of Information Management*, Vol. 38 No. 1, pp. 187-195.
Rana, T.A. and Cheah, Y.N. (2016), “Aspect extraction in sentiment analysis: comparative analysis and survey”, *Artificial Intelligence Review*, Vol. 46 No. 4, pp. 459-483, doi: 10.1007/s10462-016-9472-z.
Rashid, T. (2010), “Classification of churn and non-churn customers for telecommunication companies”, *International Journal of Biometrics and Bioinformatics*, Vol. 3 No. 5, pp. 82-89.
Ravi, K. and Ravi, V. (2015), “A survey on opinion mining and sentiment analysis: tasks, approaches and applications”, *Knowledge-Based Systems*, Vol. 89, pp. 14-46.
Reinartz, W. and Kumar, V. (2002), “The mismanagement of customer loyalty”, *Harvard Business Review*, Vol. 80 No. 7, pp. 86-94.
Risselada, H., Verhoef, P.C. and Bijmolt, T.H.A. (2010), “Staying power of churn prediction models”, *Journal of Interactive Marketing*, Vol. 24 No. 3, pp. 198-208.
Roos, I. (1999), “Switching processes in customer relationships”, *Journal of Service Research*, Vol. 2 No. 1, pp. 68-85, doi: 10.1177/109467059921006.
Salampasis, M., Paltoglou, G. and Giachanou, A. (2014), “Using social media for continuous monitoring and mining of consumer behaviour”, *International Journal of Electronic Business*, Vol. 11 No. 1, pp. 85-96.
Sarlan, A., Nadam, C. and Basri, S. (2014), “Twitter sentiment analysis”, *Proceedings of the 6th International Conference on Information Technology and Multimedia*, Vol. 1, pp. 212-216, doi: 10.1109/ICIMU.2014.7066632.
Schneider, M.J. and Gupta, S. (2016), “Forecasting sales of new and existing products using consumer reviews: a random projections approach”, *International Journal of Forecasting*, Vol. 32 No. 2, pp. 243-256.
Sepehri, M.M., Norozi, A., Teymorpur, B. and Chubdar, S. (2010), “Customer churn reasons of banking services by combining data mining and survey methods”, *Research in Management in Iran*, Vol. 14 No. 15.
Shaaban, E., Helmy, Y., Khder, A. and Nasr, M. (2012), “A proposed churn prediction model”, *International Journal of Engineering Research and Applications*, Vol. 2 No. 4, pp. 693-697.
Shin, D.H. and Kim, W.Y. (2008), “Forecasting customer switching intention in mobile service: an exploratory study of predictive factors in mobile number portability”, *Technological Forecasting and Social Change*, Vol. 75 No. 6, pp. 854-874, doi: 10.1016/j.techfore.2007.05.001.
Siemes, T. (2016), “Churn prediction models tested and evaluated in the Dutch indemnity industry”, PhD. dissertation, Open University of the Nederland, available at: https://core.ac.uk/download/pdf/80496548.pdf (accessed 3 March 2019).
Soeini, R.A. and Rodpysh, K.V. (2012), “Applying data mining to insurance customer churn management”, *International Journal of Computer Science and Information Technologies*, Vol. 30, pp. 82-92, available at: https://pdfs.semanticscholar.org/33be/72961a970b52516f494f9ef26712e0790e3b.pdf (accessed 8 March 2019).
Srivastava, U. and Gopalkrishnan, S. (2015), “Impact of big data analytics on banking sector: learning for Indian banks”, *Procedia Computer Science*, Vol. 50, pp. 643-652, doi: 10.1016/j.procs.2015.04.098.
Stewart, K. (1998), “An exploration of customer exit in retail banking”, *International Journal of Bank Marketing*, Vol. 16 No. 1, pp. 6-14.
Tavakoli, A., Mortazavi, S., Kahani, M. and Hosseini, Z. (2011), “Data mining for customer churn prediction in insurance”, *Journal of Business Management Perspective*, Vol. 4

No. 37, pp. 55-41, available at: www.sid.ir/En/Journal/ViewPaper.aspx?ID=220653 (accessed 23 March 2019).

Teichert, T., Rezaei, S. and Correa, J.C. (2020), "Customers' experiences of fast food delivery services: uncovering the semantic core benefits, actual and augmented product by text mining", *British Food Journal*, Vol. 122 No. 11, pp. 3513-3528, doi: 10.1108/BFJ-12-2019-0909.

The Banking Association South Africa (2018), "Annual report", available at: www.banking.org.za/wp-content/uploads/2019/08/BASA-Annual-Report-2018.pdf (accessed 2 April 2019).

Trapero, J.R., Pedregal, D.J., Fildes, R. and Kourentzes, N. (2013), "Analysis of judgmental adjustments in the presence of promotions", *International Journal of Forecasting*, Vol. 29 No. 2, pp. 234-243.

Tsytsarau, M. and Palpanas, T. (2012), "Survey on mining subjective data on the web", *Data Mining and Knowledge Discovery*, Vol. 24 No. 3, pp. 478-514.

Tumasjan, A., Sprenger, T.O., Sandner, P.G. and Welpe, I.M. (2010), "Predicting elections with twitter: what 140 characters reveal about political sentiment", *Fourth International AAAI Conference on Weblogs and Social Media*, Vol. 10 No. 1, pp. 178-185, available at: www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewPaper/1441 (accessed 8 March 2019).

Vafeiadis, T., Diamantaras, K., Sarigiannidis, G. and Chatzisavvas, K. (2015), "A comparison of machine learning techniques for customer churn prediction", *Simulation Modelling Practice and Theory*, Vol. 55, pp. 1-9.

Verbeek, M. (2015), *A Guide to Modern Econometrics*, John Wiley & Sons, New York, NY.

Verhagen, T., Nauta, A. and Feldberg, F. (2013), "Negative online word-of-mouth: behavioral indicator or emotional release?", *Computers in Human Behavior*, Vol. 29 No. 4, pp. 1430-1440, doi: 10.1016/j.chb.2013.01.043.

Vyas, V. and Raitani, S. (2014a), "An exploratory study of factors influencing the e-loyalty of online banking consumers", *IUP Journal of Bank Management*, Vol. 13 No. 3, pp. 34-47.

Vyas, V. and Raitani, S. (2014b), "Drivers of customers' switching behaviour in Indian banking industry", *International Journal of Bank Marketing*, Vol. 32 No. 4, pp. 321-342.

West, D. and Dellana, S. (2011), "An empirical analysis of neural network memory structures for basin water quality forecasting", *International Journal of Forecasting*, Vol. 27 No. 3, pp. 777-803.

Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E. and Patwardhan, S. (2005), "OpinionFinder: a system for subjectivity analysis", *Proceedings of HLT/EMNLP 2005 Interactive Demonstrations*, pp. 34-35, available at: www.aclweb.org/anthology/H05-2018 (accessed 27 April 2019).

Further reading

Keaveney, S.M. and Parthasarathy, M. (2001), "Customer switching behaviour in online services: an exploratory study of the role of selected attitudinal, behavioral, and demographic factors", *Journal of the Academy of Marketing Science*, Vol. 29 No. 4, pp. 374-390, doi: 10.1177/03079450094225.

Varshney, N. and Gupta, S.K. (2014), "Mining churning factors in Indian telecommunication sector using social media analytics", *International Conference on Data Warehousing and Knowledge Discovery*, pp. 405-413.

Yu, Y., Duan, W. and Cao, Q. (2013), "The impact of social and conventional media on firm equity value: a sentiment analysis approach", *Decision Support Systems*, Vol. 55 No. 4, pp. 919-926, doi: 10.1016/j.dss.2012.12.028.

Appendix

Figure A1 Topic code book

Corresponding author

James Lappeman can be contacted at: j.lappeman@uct.ac.za