Predict employee attrition by using predictive analytics

Ramakrishnan Raman and Sandeep Bhattacharya
Symbiosis Institute of Business Management,
Symbiosis International University, Pune, India, and
Dhanya Pramod
Symbiosis Centre for Information Technology,
Symbiosis International University, Pune, India

Received 31 March 2018
Revised 26 April 2018
27 April 2018
Accepted 30 April 2018

Abstract

Purpose – Research questions that this paper attempts to answer are – do the features in general email communication have any significance to a teaching faculty member leaving the business school? Do the sentiments expressed in email communication have any significance to a teaching faculty member leaving the business school? Do the stages mentioned in the transtheoretical model have any relevance to the email behaviour of an individual when he or she goes through the decision process leading to the decision to quit? The purpose of this paper is to study email patterns and use predictive analytics to correlate with the real-world situation of leaving the business school.

Design/methodology/approach – The email repository (2010–2017) of 126 teaching faculty members who were associated with a business school as full-time faculty members is the data set that was used for the research. Of the 126 teaching faculty members, 42 had left the business school during this time frame. Correlation analysis, word count analysis and sentiment analysis were executed using "R" programming, and sentiment "R" package was used to understand the sentiment and its association in leaving the business school. From the email repository, a rich feature set of data was extracted for correlation analysis to discover the features which had strong correlation with the faculty member leaving the business school. The research also used data-logging tools to extract aggregated statistics for word frequency counts and sentiment features.

Findings – Those faculty members who decide to leave are involved more in external communication and less in internal communications. Also, those who decide to leave initiate fewer email conversations and opt to forward emails to colleagues. Correlation analysis shows that negative sentiment goes down, as faculty members leave the organisation and this is in contrary to the existing review of literature. The research also shows that the triggering point or the intention to leave is positively correlated to the downward swing of the emotional valence (positive sentiment). A number of email features have shown change in patterns which are correlated to a faculty member quitting the business school.

Research limitations/implications – Faculty members of only one business school have been considered and this is primary due to cost, privacy and complexities involved in procuring and handling the data. Also, the reasons for exhibiting the sentiments and their root cause have not been studied. Also the designation, roles and responsibilities of faculty members have not been taken into consideration.

Practical implications – Business schools all over India always have a challenge to recruit good faculty members who can take up research activities, teach and also shoulder administrative responsibilities. Retaining faculty members and keeping attrition levels low will help business schools to maintain the standards of excellence that they aspire. This research is immensely useful for business school, which can use email analytics in predicting the intention of the faculty members leaving their business school.

Originality/value – Although past studies have studied attrition, this study uses predictive analytics and maps it to the intention to quit. This study helps business schools to predict the chance of faculty members leaving the business school which is of immense value, as appropriate measures can be taken to retain and restrict attrition.

Keywords Attrition, Intention to leave, Predictive analytics, Attrition in B schools, Email communication, Modelling attrition

Paper type Research paper

1. Introduction

Using emails for official communication has become the norm of the day (Moecke and Volkamer, 2013). With the usage of mobile phones for data and voice, communicating using technology has become the integral part of daily official communication (Cousins and Robey, 2015). Also many day-to-day actions and decisions that happen in the real world are

Benchmarking: An International Journal
Vol. 26 No. 1, 2019
pp. 2-18

1463-5771
DOI 10.1108/BIJ-03-2018-0083

now carried out using online platforms. Using social media platforms, short messages services and email communication as a formal means of communication to share information with the real world is particularly increasing in business schools. Increasingly, a number of business schools now rely on social platforms such as Facebook and Twitter to share information. As Gmail and G Suite for education are free of cost, many educational institutions and business schools across the world use these as a platform for communication (Mason, 2010). As real-world actions manifest into online data modelling, online behaviour in relation to the real-world behaviour has gained significant importance.

Using predictive analytics to predict real-time behaviours based on online behaviour can be of great value to organisations. There have been previous research studies which have tried to address specific issues related to the real-time behaviour of people and its relation to online behaviour. Also, research work has been carried out to explore if one's actions are influenced by online interactions, and whether such manifestation and influences can be modelled and predicted (Ramanathan and Purani, 2014; Helvie-Mason, 2011; Hua and Haughton, 2009). There have been several approaches proposed and experimented including, network analysis of social network graphs (Coons and Chen, 2014; Jay, 2005), user perception and awareness related to online privacy (Raman and Pamod, 2015), analysis of unstructured data (Seng and Yang, 2017; Zhang *et al.*, 2016; Xue and Deng, 2012), analysis of tweets (Ozturkcan *et al.*, 2017; Hallward and Armstrong, 2016; Haustein *et al.*, 2014), likes and posts on social media (Khobzi and Teimourpour, 2014; Chauhan and Pillai, 2013; Su *et al.*, 2015) and their correlation to the real-time behaviour of people.

In this paper, the online business related interactions that happen using emails in a business school have been studied. The formal email communication has been analysed using tools, and predictive modelling has been used to correlate to the real-world behaviour (intention to quit the job). The privacy issues have been taken care, as teaching faculty members who use network recourses and email opted-in by agreeing to the terms and conditions of usage, which allows for such study to be conducted with the consent of the individual. The paper studies email patterns and uses predictive analytics to correlate with the real-world situation of leaving the organisation. Also, the email communication has been analysed to predict specific features and sentiments that can be used to detect changes in workplace behaviour and the intention to leave the business school.

2. Literature review

2.1 Attrition and its impact

Issues related to privacy have made it very difficult to model churn in any organisation. The non-availability of factual data related to the real reason for an employee to leave the organisation makes the data analysis almost impossible. Apart from the challenges related to the availability of factual data, modelling and predicting churn in a business school is difficult even from the dimension of data analysis. The data collected in the exit interviews or using online forms are not fully reliable and consistent. Collecting the data from the exit interviews related to social context and using an online form to capture the emotional aspects related to the root cause of the decision to leave are almost impossible. The real reason for an employee to leave an organisation from a data perspective is "noisy" as the reasons attributing for the decision to leave could be plenty. The employee might be leaving due to personal reasons or due to the lack of job satisfaction.

The review of literature indicates that a lot of research work has been done in the area related to job satisfaction. Also, there have been several conceptual and operational definitions of the job satisfaction construct. Oshagbemi (2003) brings evidence that the general job satisfaction depends on several aspects of the job, and many a time, the nature of the job itself could be the prime reason. Oshagbemi also describes that the many reasons including the co-workers effect on the person which could affect the job satisfaction, or the supervision in the job, the pay offered, the working environment, the policies and procedures followed in the organisation and the opportunities for growth in the organisation can affect the job satisfaction, which, in turn, can have an impact on the decision of an employee to leave the organisation. Also, the reasons for a person to leave an organisation could be due to health or due to a decision taken by the management to lay-off employees. Hence, the exact underpinning reason for an employees' decision to leave is sometimes unobservable. In higher educational institutions, even the lack of participation by the students might cause job dissatisfaction which could cause attrition (Ruhul and Nafeez, 2003). All this makes modelling and predicting employee's intention to leave all the more difficult.

There have been several studies related to attrition and job satisfaction. Varied aspects like role of technology, role of emotional intelligence (EI), change-oriented leadership, organisational culture, individual-, group-, environmental- and organisational-level variables effect on attrition have been studied. Researchers have also studied job characteristic model (JCM) through the motivating potential scores (MPSs) and their impact on attrition. Gupta *et al.* (2018) empirically test the effect of cloud-based enterprise resource planning services on the performance of an organisation.

A research study conducted amongst information technology professionals in Pune revealed the clarity of role and adequacy of resources, nurturing employee loyalty and organisational inspiration as the key drivers against turnover intentions (Raman *et al.*, 2013). Literature indicates that those individuals with high scores on EI get strong job performance ratings, which can possibly control attrition. Change-oriented leadership has a positive and significant direct effect on planned change (Al-Ali *et al.*, 2017), and there is a linkage between leader, team, perceived organisational support and organisational culture that is mediated by employee motivation (Al Mehrzi and Singh, 2016). A framework of employee turnover intentions was developed, which suggests the interplay of individual-, group-, environmental- and organisational-level variables on employee turnover intentions (Harhara *et al.*, 2015). A research study conducted in United Arab Emirates studied the impact of various dimensions of the JCM through the MPSs on professionals residing in the UAE. The results reveal the impact of increasing age on high MPSs, higher MPSs of white-collared jobs than the blue-collared jobs, owing to high skill variety in white-collared jobs. It was revealed that women scored higher on MPSs than men, and the men of Indian origin showed high motivation as compared to their counterparts from other countries.

In spite of all the numerous reasons affecting the decision to leave, a faculty member decides to leave due to job dissatisfaction. This has been validated by the social science literature (Fields *et al.*, 2011), which has explored the predictors of job change and supported the concept that churn may be modelled in a better fashion as a decision not only to leave a job, but also to move into a different work situation. Jiang *et al.* (2012) in their study have derived that job embeddedness has a negative correlation to turnover intentions, and this is true only after controlling all factors that affect job satisfaction.

The attrition of teachers is a perennial problem in schools and educational institutions (Macdonald, 1999). Generally, individuals in the teaching professions tend to job hop in their early years of their career (Ingersoll, 2003). The attrition of teaching faculty members if predicted and controlled can help an educational institution to do better, as the quality of teaching faculty members is one of the most important determinants of students' achievement and gain (Rivkin *et al.*, 2005; Sanders and Horn, 1994; Rice, 2003; Wayne and Youngs, 2003; Wilson and Floden, 2003; Wilson *et al.*, 2001).

There have been some studies related to the attrition of faculty members working in developed countries. A study in the USA related to teacher attrition explores the characteristics of individuals who are more likely to leave their jobs. The findings of the study show that maths and science teachers at the school level who are non-minority females are more likely to leave (Borman and Dowling, 2008; Guarino, 2006; Johnson *et al.*, 2005). A framework created to understand the role of factors affecting students' academic achievement found that school leadership and climate together affect the academic achievement of the students, and were mediated by the involvement of the parents of the students (Alhosani *et al.*, 2017).

Literature also reveals that the turnover intentions of teachers decrease with experience and increase significantly as they approach the retirement age (Boe *et al.*, 1997; Harris and Adams, 2007; Kirby and Grissmer, 1993; Luekens *et al.*, 2004; Murnane *et al.*, 1989). At the postgraduate level, attracting and retaining teachers is of paramount importance. Berliner (2004) identifies the complexities in identifying and recruiting expert teachers. Tschannen-Moran and Woolfolk Hoy (2001) in their research record the unrelenting measurement problems related to the study of teachers' effectiveness. Research portrays that teachers with less experience are less effective in delivering expected student satisfaction, in comparison with teachers with higher work experience (Hanushek *et al.*, 2004; Guarino *et al.*, 2006). Selecting the right teaching faculty members who have the right knowledge and ability to research and teach is of paramount importance for business schools. The agony is those teachers who have high scholastic aptitude are less likely to take up teaching as a profession and in case they do so, they are more likely to leave (Lankford *et al.*, 2002; Murnane and Olsen, 1989; Murnane *et al.*, 1991; Podgursky *et al.*, 2004; Stinebrickner, 2001). The selectivity of the university attended by a faculty member has a positive influence on students' achievement (Ehrenberg and Brewer, 1994; Summers and Wolfe, 1977; Ferguson and Ladd, 1996).

In India, attracting faculty members from the institutes of repute is a challenge. Several factors act as deterrents including the quality of life and the pay parity when compared to institutions in the USA, Europe and other countries. While a faculty member adds immense value to the institution, there is very limited literature that inspects the relationship between teacher value added and attrition (Boyd *et al.*, 2007; Goldhaber *et al.*, 2007; Hanushek *et al.*, 2005). Higher education institutions are becoming responsive to the needs of the academic staff, especially the faculty members (Chen *et al.*, 2006; Edwards *et al.*, 2009; Sahney *et al.*, 2008). Business schools have started focusing on reducing turnover and are finding means to improve job satisfaction and retention (Efraty *et al.*, 1991; Fuller, 2006; Sirin, 2009; Worrall and Cooper, 2006). Papis (2006) suggests that for an organisation to realise its objectives, understanding the workforce and being proactive to their needs is vital. Institutions are also taking measures to increase employee involvement and engagement in socially responsible activities to possibly find solutions to arrest attrition (Razaq *et al.*, 2011). For any organisation, effective workforce leads to high productivity, which, in turn, helps organisations to remain effective in a highly competitive business world (Horn and Fichtner, 2003). Similarly, a strong faculty strength with low attrition helps in branding and high productivity in terms of research output and ranking of the business school. While there have been measures taken by several Indian business schools to control attrition, they have done it by considering leaving of employees as an independent event. The general strategies to control attrition are either by providing better incentives or by imposing penalties that the cost of next best alternatives becomes unaffordable. In reality, faculty members do not take the decision to quit independently. There is a hidden, yet powerful social factor behind a user's decision-making processes. Drawing a parallel with the telecommunication industry, many a time a customer decides to switch from one telecom provider to another based on the decisions or suggestions given by his/her social friends. This social relationship modelling and predicting the chance of customer churn has been studied in the telecom industry (Dasgupta *et al.*, 2008; Ferreira *et al.*, 2004; Huang *et al.*, 2010; Hung *et al.*, 2006; Masand *et al.*, 1999; Mozer *et al.*, 2000; Tsai and Lu, 2009; Zhao *et al.*, 2005) in the banking and finance industry (Coussement and Den Poel, 2008; Nie *et al.*, 2009) and also in the insurance industry (Herrera and Znati, 2007), but there is no literature or model available for attrition being influenced by hidden, yet powerful social factor for teaching faculty members working in business schools.

2.2 Understanding the transtheoretical model and its influence on change in behaviour

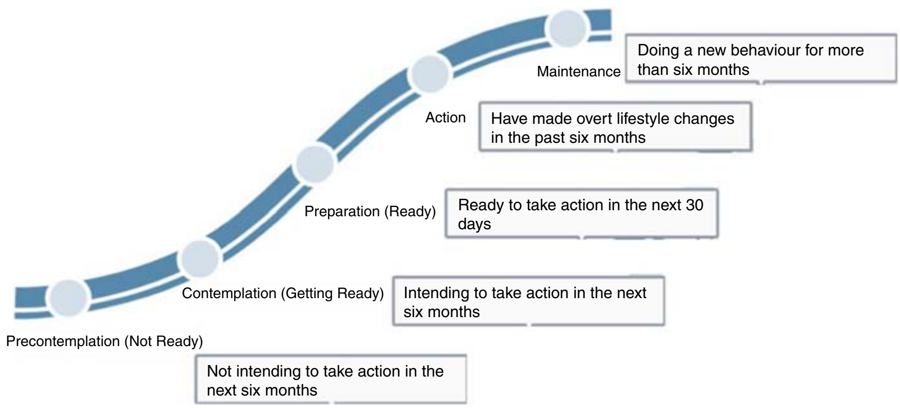
The transtheoretical model (TTM) (Prochaska and DiClemente, 1982; Prochaska *et al.*, 1992, 2002, 2008) gives an integrated approach to conceptualise the process of intentional behaviour change at the individual level. This model takes a bio-psychosocial approach to explain the intentional behaviour change of an individual, and hence it includes and integrates constructs from several theories which model change behaviour. The TTM can be applied to several behaviours and settings to explain the behaviour of an individual. The change stages are the core of TTM. The TTM explains that individuals go through a series of stages while they modify their behaviour and take decisions. The TTM explains change as a process that reveals over time, involving progress through a series of stages. While the "stage" represents a temporal dimension, change is a phenomenon that occurs over a period of time. According to the TTM, different stages that an individual goes through before a change are precontemplation, contemplation, preparation, action and maintenance.

Referring to Figure 1, precontemplation is the stage in which an individual does not plan to make or take decision related to any change, i.e., the individual is not ready for the change. Contemplation is the stage where the individual is getting ready for the change and the change or the decision related to change could happen in a period of six months. Preparation is the stage where the individual is ready for the change and it is most likely that the change or the decision to change happens in the near future, and the time period for the change is generally one month. Action stage is where the individual has taken decision for a change and the same is observed in the overall behaviour, which can be equated to the action or actions of the individual. Maintenance stage is the one in which the individual has made specific overt modifications to his/her behaviour and continues to live with the actions and behaviours practised during the action phase. The decision to leave an organisation is also a part of the change process for an individual.

Based on the literature, the following seven hypotheses have been formulated:

H1. General email communication features have no significant role in predicting attrition.

H2. Specific email communication parameters have no significant role in predicting attrition.



Source: Based on Prochaska and DiClemente (1982)

Figure 1.
The transtheoretical model (TTM)

H3. After office hour communication has no significant role in predicting attrition.

H4. Communication inside the organisation has no significant role in predicting attrition.

H5. Sentiment expressed in email communication has no significant role in predicting attrition.

H6. Being expressive in email communication has no significant role in predicting attrition.

H7. The stages mentioned in TTM have no relevance to email behaviour of an individual when he or she goes through the decision process leading to the decision to quit.

3. Methodology

The study understands and analyses online social interactions in official email communication and uses the same to predict the intention of a faculty member to quit the business school. We mine the official email repository of teaching faculty members during a specific time period. This mining of mail repository has been done for teaching faculty members who left a business school and also those who did not leave the business school. Privacy issues have been taken care, as all faculty members who use network recourses and email opted-in by agreeing to the terms and conditions of usage, which allows for such study to be conducted with the consent of individuals.

3.1 Participants

The email repository of 126 full-time faculty members who were associated with a business school was used for the research study. The data set was taken for six years starting from 2010 to 2017. From the 126 teaching faculty members, 42 had left the business school during this time frame. This email repository had members who were associated with the business school and also those who had left the institution. Explicit consent was taken from all the faculty members for accessing their email repository, which could enable querying of the content of the mail.

3.2 Measures and summary of procedure and analysis

To analyse, the data dump of emails in excel sheets was created. Numeric identifiers were used to identify and differentiate people who had left the business school from those who had not. Numeric identifiers also ensured that personally identifiable information was not accessed. Data were extracted from the email repository and were pushed into different columns of excel sheet. The columns had date, time of email and other metadata along with the set of words from the email message placed in a column. This was done using automated tools which were commercially available, and R programming was used for sentiment analysis. The content in the excel sheets was used to pull out aggregate content. Correlation analysis, sentiment analysis and word count analysis were executed using "R" programming, and sentiment R package was used to understand and predict the intention to leave the business school. From the email repository, a rich feature set of data was extracted for correlation analysis to discover the features which had strong correlation with the person leaving the business school. To predict the pattern and to understand the intentions to leave, subsets of the time frames were created.

Given an employee, we considered his/her employment period and segmented it into three phases. The initial period (Phase A) is the first one month of employment. Emails in this phase were not considered as they did not represent any significant regular work. The Phase A is the time frame when the employee becomes familiar with the organisation culture. The notice period (Phase C) is the last one month of employment when the employee had decided to quit the organisation. The activities during Phase C are usually wrapping up the pending tasks and handing over the work to appropriate person who will take things ahead. The working period (Phase B) is the period from the second month to the last one month, which is the actual working period of the employee. The data set during this period was taken for predicting the pattern which can trigger the intention to quit. The Phase B was further divided into three phases. Phase 1 was the first part of the productive working period, Phase 2 was the second part of the productive working period and Phase 3 was the third and last part of the working period. As shown in Figure 2, the research used data-logging tools to extract aggregated statistics for word frequency counts and sentiment features during Phase B. All analyses were performed on the data related to Phase B.

4. Findings

Table I gives the results of the correlation coefficient of several parameters which are taken from emails. The features have been grouped such as general email communication, specific email communication, after office hour communication, communication inside the organisation, sentiment and expressive communication. Specific parameters have been considered under each of the features. Correlation coefficient has been calculated for each parameter for those faculty members who left the business school and also for those who did not. Correlation coefficient has been computed as faculty members move from Phase 1 to Phase 2 and also from Phase 2 to Phase 3 of their employment period.

A positive value of the correlation coefficient denotes a general increase of parameter value as faculty members go from one phase to another, whereas a negative value of the correlation coefficient indicates a generally decreasing parameter value.

Considering the feature "general email communication", the value of correlation coefficient in Table I for the parameters gives an indication of pattern change. The parameter – number of email communication – clearly shows that the value of correlation coefficient for those faculty members who left the business school changes from 0.370 to −0.126. This denotes that the number of emails sent by faculty members goes up as they move from Phase 1 to Phase 2, and as faculty members transit from Phase 2 to Phase 3, the number of email communication goes down. This indicates that the number of email communication decreases as faculty members move closer to the decision of departure from the business school. Hence, the parameter "number of email communication" becomes a

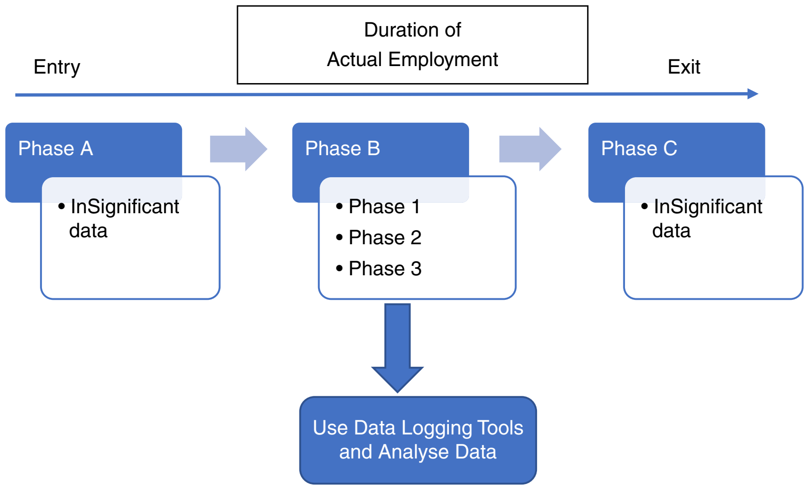


Figure 2.
The phases of employment of an employee in an organisation

Feature	Specific parameters	Faculty members who left the business school		Faculty members who did not leave the business school	
		Transition from Phase 1 to Phase 2	Transition from Phase 2 to Phase 3	Transition from Phase 1 to Phase 2	Transition from Phase 2 to Phase 3
General email communication	Number of emails	0.370	−0.126	0.430	0.446
	Number of email attachments	0.355	0.312	0.362	0.361
Specific email communication	Number of email forwards	0.150	0.525	0.171	0.172
	Number of recipients marked in "To" address	0.179	−0.076	0.165	0.162
	Average recipients marked Cc	0.309	0.012	0.212	0.207
	Number of replies	0.512	−0.316	0.539	0.598
After office hour communication	Number of emails sent after office hours	0.338	−0.085	0.312	0.349
Communication inside the organisation	Number of emails to colleagues inside the organisation	0.155	−0.063	0.156	0.143
	Number of copies to colleagues inside the organisation	0.194	0.022	0.176	0.182
Sentiment	Number of emails with positive sentiments	0.157	0.161	0.221	0.228
	Number of emails with negative sentiments	0.218	−0.012	0.218	0.215
Expressive communication	Number of "?" symbol in email	0.146	−0.055	0.154	0.159
	Number of "!" symbol in email	0.204	−0.053	0.231	0.245
	Number of superlatives used in emails	−0.007	−0.061	−0.058	−0.078
	Number of emails with complex words	0.144	−0.055	0.143	0.149

Table I. Feature set comparison of correlation coefficient during the transition of phases

predictor for attrition. This is also justified, as the value of correlation coefficient for the same parameter for faculty members who did not quit the business school has increased from 0.430 to 0.446. Email attachments as a parameter did not qualify to be a predictor, as all correlation coefficient values were positive. Hence, *H1* is rejected.

Specific email communication has the number of email forwards, the number of recipients marked in "To" address, average recipients marked copy and the number of replies as the parameters. It can be seen from Table I that there is no significant change in the values of correlation coefficient for all parameters for those faculty members who did not leave the business school.

The correlation coefficient values significantly change from transition Phase 1 to Phase 2 and transition Phase 2 to Phase 3 for those faculty members who had left the business school for the parameters – number of email forwards, number of recipients marked in "To" address, number of replies and average recipients marked copy of the mail. The correlation coefficient values show that faculty members who had quit the business school tend to forward the emails (0.150 to 0.525), reply very less to the emails (0.512 to −0.316) and send fewer emails to specific individuals (0.179 to −0.076). The average number of people kept in copy of the email reduced (0.309 to –0.012) but had not slipped to a negative. Hence, those faculty members who had decided to leave initiated fewer email conversations and opted to forward emails to colleagues. This also indicates that the number of email forwards, the number of replies and also the number of emails sent to the specific number of people become predictors for attrition. Therefore, *H2* is rejected.

The parameter number of emails sent after office hours of the feature "after office hour communication" is certainly a predictor of attrition. The correlation coefficient values for those faculty members who had not quit the business school are more than 0.3 (0.312 and 0.349) at both the transition stages, whereas the value has moved from 0.338 to −0.085 for those faculty members who had left. The parameter can be interpreted as those who had left the business school did not take work home after completing their office and hence did not send official emails after office hours. This helps us to conclude that *H3* is rejected.

Similarly, the parameter – number of email communications and the parameter – number of emails to colleagues inside the organisation of the feature "communication inside the organisation" also help in predicting attrition. While the correlation coefficient values were hovering around 0.1 for the faculty members who did not leave, and it changed from 0.155 to −0.063 for the parameter – number of emails to colleagues inside the organisation. Although the value of correlation coefficient reduces from 0.194 to 0.022, it is not negative and hence it is not a predictor for attrition. Comparing the value with the number of emails sent, it can also be found that those faculty members who decided to leave were involved more in external communication, i.e., communicating more with those outside the organisation and less in internal communications, i.e., with those within the organisation. So, *H4* is rejected.

Majority of the parameters of the feature "sentiment" and "expressive communication" qualified to be the predictors of attrition. The statistical data revealed that the positive sentiments marginally increased (correlation coefficient changed from 0.157 to 0.161) and the negative sentiments decreased (correlation coefficient changed from 0.218 to −0.012) for the faculty members who left the business school but the values were around 0.2 for those who did not leave. Literature indicates that positive sentiments and emotions are linked to events that help the fulfilment of an individual's objectives, whereas negative sentiments and emotions are linked with those events that hinder the fulfilment of an individual's objectives (Barclay *et al.*, 2005). If this postulate is applied to an employee leaving an organisation or a faculty member leaving a business school, his or her notice period would be a hindrance for fulfilling their objective of taking a new job offer. Hence, they are expected to express negative sentiments. In contrary, the current findings show that the negative sentiments are reduced. The correlation coefficient value shows that the negative sentiment goes down as faculty members leave the organisation and this is in contrary to the existing review of literature. Hence, *H5* is rejected.

Also, those faculty members who left the business school were less expressive in their communication. This can be seen with the values of correlation coefficient for the parameters – number of "?" symbol in email (0.146 to −0.055), number of superlatives used in emails (−0.007 to −0.061) and number of emails with complex words (0.144 to −0.055). Therefore, *H6* is also rejected.

Based on the correlation values, Table II gives the list parameters which help in predicting attrition and which have been derived after analysing email communication. The exhibited behaviour has also been mentioned. The parameters and the correlation values give a behavioural pattern for the group of faculty members who left the business school *vis-à-vis* those who did not leave. The results also have relevance to the TTM. The change in correlation coefficient values indicates that the individuals are going through the phases of contemplation, preparation and action. Hence, this study shows that the change in behaviours is involuntarily and inadvertently exhibited by the individuals in their email communication as they go through the change process. Hence, *H7* is rejected.

The analysis carried out using "R" sentiment analysis showed that there were less negative sentiments expressed by the people who left the B school. "R" programming and sentiment R package were further used to understand and predict the intention to leave the business school. The emotional valence graph was generated and analysed to find if there was significant difference in the pattern of emotional valence of faculty members who left the business schools when compared to those who did not leave.

The emotional valence graph, which portrays the positive sentiment of individuals, was generated for all the phases of Phase B. It was found that the emotional valence of all faculty members who left the business school followed a downward swing pattern. The values of emotional valence differed but they all exhibited a trend of down swing.

Figure 3 shows the emotional valence of some faculty members who left the business school, and Figure 4 shows the emotional valence of some faculty members who did not leave the business school. It can be clearly seen from Figure 4 that those who did not leave the business school had the emotional valences swing but the end trail was always positive. This was not the case for faculty member who left the business school, which is shown in Figure 3, where the emotional valence end trail was downward.

This also indicates that emotional valence, which is the sentiment exhibited by the faculty members in their official email, can be a predictor for attrition. A constant downward swing of emotional valence clearly indicates the chance for attrition to be high. Hence, the triggering point or the intention to leave is positively correlated to a continuous downward swing of the emotional valence (positive sentiment).

5. Conclusion

The analysis of email repository has helped in identifying specific aspects related to email usage and behaviour that can help in predicting the attrition of faculty members in a business school. Though the results are preliminary, the analysis establishes a data-driven predictive correlation model between faculty members behaviour on email and the real-world event of they leaving the business school.

5.1 Contributions of the study

This research study has immensely contributed in identifying specific features of email that can help in predicting attrition. The study has also contributed to the TTM by presenting that the change in behaviours is also involuntarily and inadvertently exhibited by the individuals in their email communication as they go through the change process, while they decide to change their jobs. Informative features in email and sentiment trends have been discovered, which can also help in predicting the attrition of faculty members in a business school. The analysis has also revealed that faculty members become less communicative and expressive before they decide to leave the organisation and has also contributed to the

Parameter	Interpretation
Number of emails	Fewer number of email communication
Number of email forwards	Higher number of email forwards
Number of recipients marked in "To" address	Lesser number of specific emails sent
Number of replies	Lesser replies to emails
Number of emails sent after office hours	Fewer emails after office hours
Number of emails to colleagues inside the organisation	Lesser number of email communication with colleagues
Number of emails with negative sentiments	Less negative sentiments expressed
Number of "?" symbol in email	Less expressive
Number of "!" symbol in email	
Number of emails with complex words	

Table II. Parameters from email communication that predict attrition

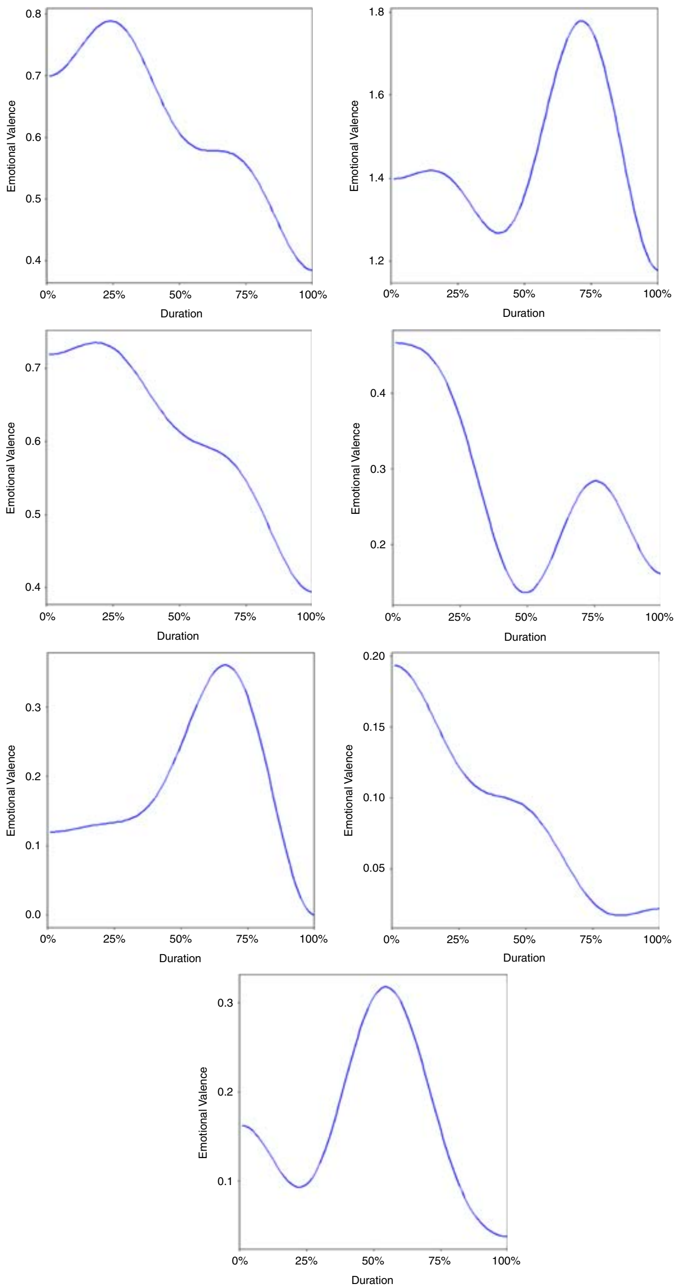


Figure 3.
Emotional valence of some faculty members who left the business school

existing body of knowledge, on the change in sentiment, as exhibited in the email communication before they quit the workplace. Hence, this research study has shown how the predictive analytics on email communication can be used to predict incipient departure of a faculty member with a reasonable accuracy.

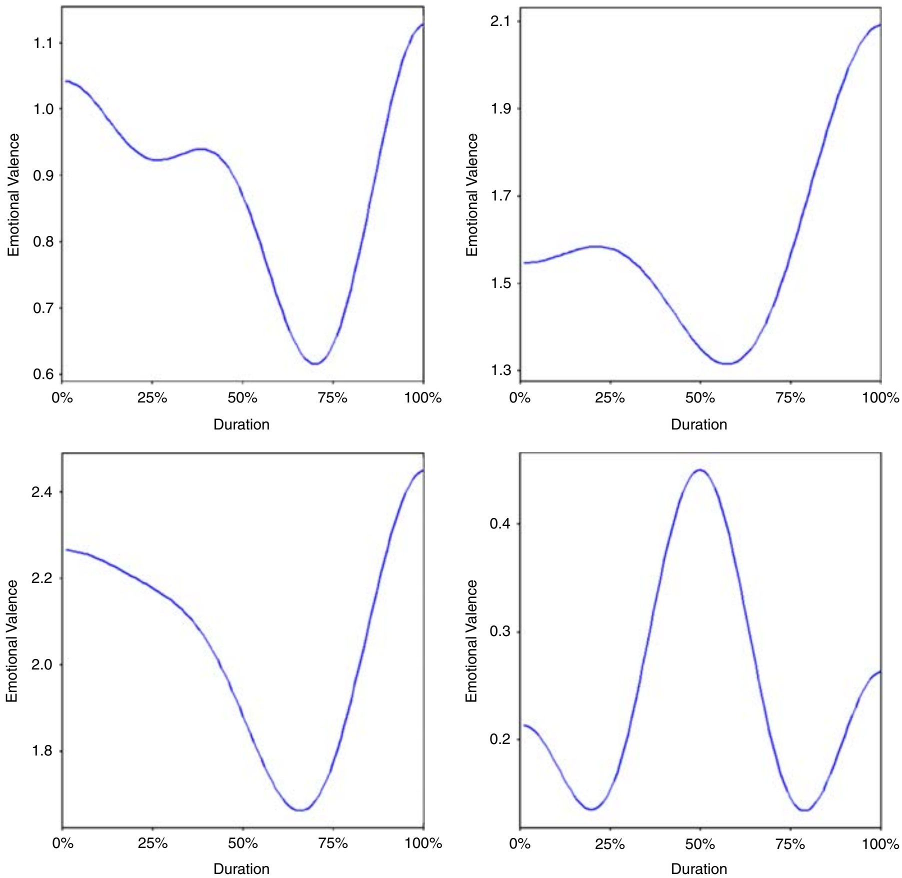


Figure 4.
Emotional valence of some faculty members who did not leave the business school

5.2 Limitations and future research

Our study has considered the email repository of one business school alone. If cost, privacy and complexities involved in procuring and handling the data can be addressed, the data set could include the email repository of several business schools. Also, the reasons for exhibiting the sentiments and their root cause have not been studied. The link between the designation of the faculty member, his/her roles and responsibility and their influence on the email behaviour and attrition has not been studied. Further study can consider eliminating the above limitations and also consider data from social media platforms along with the email data sets, which can give better insight in predicting the attrition of faculty members from business school.

References

Al-Ali, A.A., Singh, S.K., Al Nahyan, M. and Sohal, A.S. (2017), "Change management through leadership: the mediating role of organizational culture", *International Journal of Organizational Analysis*, Vol. 25 No. 4, pp. 723-739.

Alhosani, A.A., Singh, S.K. and Al Nahyan, M. (2017), "Role of school leadership and climate on student achievement", *International Journal of Educational Management*, Vol. 31 No. 6, pp. 843-851.

Al Mehrzi, N. and Singh, S.K. (2016), "Competing through employee engagement: a proposed framework", *International Journal of Productivity and Performance Management*, Vol. 65 No. 6, pp. 831-843.

Barclay, L.J., Skarlicki, D.P. and Pugh, S.D. (2005), "Exploring the role of emotions in injustice perceptions and retaliation", *Journal of Applied Psychology*, Vol. 90 No. 4, pp. 629-643.

Berliner, D.C. (2004), "Describing the behavior and documenting the accomplishments of expert teachers", *Bulletin of Science, Technology, and Society*, Vol. 24 No. 3, pp. 200-212.

Boe, E.E., Bobbitt, S.A. and Cook, L.H. (1997), "Whither didst thou go? Retention, reassignment, migration, and attrition of special and general education teachers from a national perspective", *The Journal of Special Education*, Vol. 30 No. 4, pp. 371-389.

Borman, G.D. and Dowling, N.M. (2008), "Teacher attrition: a meta-analytic and narrative review of the research", *Review of Educational Research*, Vol. 78 No. 3, pp. 367-409.

Boyd, D., Grossman, P., Lankford, H., Loeb, S. and Wyckoff, J. (2007), "Who leaves? Teacher attrition and student achievement", available at: www.teacherpolicyresearch.org (accessed 22 May 2008).

Chauhan, K. and Pillai, A. (2013), "Role of content strategy in social media brand communities: a case of higher education institutes in India", *Journal of Product & Brand Management*, Vol. 22 No. 1, pp. 40-51, available at: https://doi.org/10.1108/10610421311298687

Chen, S., Yang, C., Shiau, J. and Wang, H. (2006), "The development of an employee satisfaction model for higher education", *The TQM Magazine*, Vol. 18 No. 5, pp. 484-500.

Cousins, K. and Robey, D. (2015), "Managing work-life boundaries with mobile technologies: an interpretive study of mobile work practices", *Information Technology & People*, Vol. 28 No. 1, pp. 34-71, available at: https://doi.org/10.1108/ITP-08-2013-0155

Coons, J. and Chen, S.-L.S. (2014), "Social network analysis for Facebook: locating cliques and visualizing sociability", in Johns, M.D., Chen, S.-L.S. and Terlip, L.A. (Eds), *Symbolic Interaction and New Social Media (Studies in Symbolic Interaction)*, Vol. 43, Emerald Publishing Limited, pp. 43-61.

Coussement, K. and Den Poel, D.V. (2008), "Churn prediction in subscription services: an application of support vector machines while comparing two parameter-selection techniques", *Expert Systems with Applications*, Vol. 34 No. 1, pp. 313-327, available at: www.sciencedirect.com/science/article/pii/S0957417406002806

Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A.A. and Joshi, A. (2008), "Social ties and their relevance to churn in mobile telecom networks", *EDBT*, pp. 668-677.

Edwards, J.A., Van Laar, D., Easton, S. and Kinman, G. (2009), "The work-related quality of life scale for higher education employees", *Quality in Higher Education*, Vol. 15 No. 3, pp. 207-219.

Efraty, D., Sirgy, M.J. and Claiborne, C. (1991), "The effects of personal alienation on organizational identification: a quality-of-work-life model", *Journal of Business and Psychology*, Vol. 6 No. 1, pp. 57-78.

Ehrenberg, R.G. and Brewer, D.J. (1994), "Do school and teacher characteristics matter? Evidence from high school and beyond", *Economics of Education Review*, Vol. 13 No. 1, pp. 1-17.

Ferguson, R.F. and Ladd, H.F. (1996), "How and why money matters: an analysis of Alabama schools", in Ladd, H.F. (Ed.), *Holding Schools Accountable: Performance-Based Reform in Education*, Brookings Institution, Washington, DC, pp. 265-298.

Ferreira, J.B., Vellasco, M., Pacheco, M.A. and Barbosa, C.H. (2004), "Data mining techniques on the evaluation of wireless churn", *European Symposium on Artificial Neural Networks*, *Bruges, 28–30 April*, pp. 483-488.

Fields, D., Dingman, M.E., Roman, P.M. and Blum, T.C. (2011), "Exploring predictors of alternative job changes", *Journal of Occupational and Organizational Psychology*, Vol. 78 No. 1, pp. 63-82.

Fuller, G. (2006), "Anti-stress scheme boosts health and morale of city of London police", *Personnel Today*, Vol. 3 No. 7, pp. 207-219.

Goldhaber, D.D., Gross, B. and Player, D. (2007), "Are public schools really losing the 'best'? Assessing the career transitions of teachers and their implications for the quality of the teacher workforce", paper presented at the American Education Finance Association, Baltimore, MD.

Guarino, C., Santibañez, L. and Daley, G. (2006), "Teacher recruitment and retention: a review of the recent empirical literature", *Review of Educational Research*, Vol. 76 No. 2, pp. 173-208.

Gupta, S., Kumar, S., Singh, S.K., Foropon, C. and Chandra, C. (2018), "Role of cloud ERP on the performance of an organization: contingent resource based view perspective", *EDBT '08 Proceedings of the 11th International Conference on Extending Database Technology: Advances in Database Technology*, pp. 668-677.

Hallward, M.C. and Armstrong, C. (2016), "Tweeting resistance: the evolution of engagement frameworks", in Hancock, L.E. (Ed.), *Narratives of Identity in Social Movements, Conflicts and Change (Research in Social Movements, Conflicts and Change)*, Vol. 40, Emerald Publishing Limited, pp. 227-261.

Hanushek, E.A., Kain, J.F. and Rivkin, S.G. (2004), "Why public schools lose teachers", *Journal of Human Resources*, Vol. 39 No. 2, pp. 326-354.

Hanushek, E.A., Kain, J.F., O'Brien, D.M. and Rivkin, S.G. (2005), "The market for teacher quality", Working Paper No. 11154, NBER Working Paper Series, National Bureau of Economic Research, Cambridge.

Harhara, A., Singh, S.K. and Hussain, M. (2015), "Correlates of employee turnover intentions in oil & gas industry in the UAE", *International Journal of Organizational Analysis*, Vol. 23 No. 3, pp. 493-504.

Harris, D.N. and Adams, S. (2007), "Understanding the level and causes of teacher turnover: a comparison with other profession", *Economics of Education Review*, Vol. 26 No. 3, pp. 325-337.

Haustein, S., Bowman, T.D., Holmberg, K., Peters, I. and Larivière, V. (2014), "Astrophysicists on Twitter: an in-depth analysis of tweeting and scientific publication behavior", *Aslib Journal of Information Management*, Vol. 66 No. 3, pp. 279-296, available at: https://doi.org/10.1108/AJIM-09-2013-0081

Helvie-Mason, L. (2011), "Facebook, 'friending,' and faculty–student communication", in Wankel, C. (Ed.), *Teaching Arts and Science with the New Social Media (Cutting-Edge Technologies in Higher Education)*, Vol. 3, Emerald Publishing Limited, pp. 61-87.

Herrera, O. and Znati, T. (2007), "Modeling churn in p2p networks", *Annual Simulation Symposium, IEEE Computer Society*, pp. 33-40.

Horn, C.E.V. and Fichtner, A.R. (2003), "An evaluation of state-subsidised, firm-based training: the workforce development partnership programme", *International Journal of Manpower*, Vol. 24 No. 1, pp. 97-110.

Hua, G. and Haughton, D. (2009), "Virtual worlds adoption: a research framework and empirical study", *Online Information Review*, Vol. 33 No. 5, pp. 889-900, available at: https://doi.org/10.1108/14684520911001891

Huang, B., Buckley, B. and Kechadi, T.M. (2010), "Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications", *Expert System Applications*, Vol. 37 No. 5, pp. 3638-3646, available at: http://dx.doi.org/10.1016/j.eswa.2009.10.027

Hung, S.Y., Yen, D.C. and Wang, H.Y. (2006), "Applying data mining to telecom churn management", *Expert Systems with Applications*, Vol. 31 No. 3, pp. 515-524, available at: www.sciencedirect.com/science/article/pii/S0957417405002654

Ingersoll, R.M. (2003), *Is There Really a Teacher Shortage?*, Center for the Study of Teaching and Policy, Seattle, WA.

Jay, L. (2005), "Linking social network analysis with the analytic hierarchy process for knowledge mapping in organizations", *Journal of Knowledge Management*, Vol. 9 No. 1, pp. 76-86, available at: https://doi.org/10.1108/13673270510582974

Jiang, K., Liu, D., McKay, P.F., Lee, T.W. and Mitchell, T.R. (2012), "When and how is job embeddedness predictive of turnover? A meta-analytic investigation", *Journal of Applied Psychology*, Vol. 97 No. 5, pp. 1077-1096.

Johnson, S.M., Berg, J.H. and Donaldson, M.L. (2005), *Who Stays in Teaching and Why: A Review of the Literature on Teacher Retention*, Harvard Graduate School of Education, Cambridge, MA, available at: http://assets.aarp.org/www.aarp.org_/articles/NRTA/Harvard_report.pdf (accessed 1 July 2005).

Khobzi, H. and Teimourpour, B. (2014), "How significant are users' opinions in social media?", *International Journal of Accounting & Information Management*, Vol. 22 No. 4, pp. 254-272, available at: https://doi.org/10.1108/IJAIM-02-2014-0010

Kirby, S. and Grissmer, D.W. (1993), *Teacher Attrition: Theories, Evidence, and Suggested Policy Options*, RAND, Santa Monica, CA.

Lankford, M., Loeb, S. and Wyckoff, J. (2002), "Teacher sorting and the plight of urban schools: a descriptive analysis", *Educational Evaluation and Policy Analysis*, Vol. 24 No. 1, pp. 37-62.

Luekens, M.T., Lyter, D.M., Fox, E.E. and Chandler, K. (2004), *Teacher Attrition and Mobility: Results from the Teacher Follow-Up Survey, 2000-2001*, National Center for Education Statistics, Washington, DC, available at: http://nces.ed.gov/pubs2004/2004301.pdf (accessed 1 July 2005).

Macdonald, D. (1999), "Teacher attrition: a review of literature", *Teaching and Teacher Education*, Vol. 15 No. 8, pp. 835-848.

Masand, B., Datta, P., Mani, D. and Li, B. (1999), "Champ: a prototype for automated cellular churn prediction", *Proceedings of the 10th WSEAS International Conference on APPLIED MATHEMATICS, Published by World Scientific and Engineering Academy and Society (WSEAS), Wisconsin, 2006*, pp. 523-528.

Mason, D. (2010), "The University of Google: education in the (Post) information age", *The Electronic Library*, Vol. 28 No. 2, pp. 348-350, available at: https://doi.org/10.1108/02640471011033729

Moecke, C.T. and Volkamer, M. (2013), "Usable secure email communications: criteria and evaluation of existing approaches", *Information Management & Computer Security*, Vol. 21 No. 1, pp. 41-52, available at: https://doi.org/10.1108/09685221311314419

Mozer, M.C., Wolniewicz, R., Grimes, D.B., Johnson, E. and Kaushansky, H. (2000), "Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry", *Transaction Neural Networks*, Vol. 11 No. 3, pp. 690-696, available at: http://dx.doi.org/10.1109/72.846740

Murnane, R.J. and Olsen, R.J. (1989), "The effects of salaries and opportunity costs on length of stay in teaching: evidence from Michigan", *Review of Economics and Statistics*, Vol. 71 No. 2, pp. 347-352.

Murnane, R.J., Singer, J.D. and Willett, J.B. (1989), "The influences of salaries and 'opportunity costs' on teachers' career choices: evidence from North Carolina", *Harvard Educational Review*, Vol. 59 No. 3, pp. 325-346.

Murnane, R.J., Singer, J.D., Willet, J.B., Kemple, J. and Olsen, R.J. (1991), *Who Will Teach? Policies That Matter*, Harvard University Press, Cambridge, MA.

Nie, G., Wang, G., Zhang, P., Tian, Y. and Shi, Y. (2009), "Finding the hidden pattern of credit card holder's churn: a case of china", *Proceedings of the 9th International Conference on Computational Science, Springer-Verlag, Berlin and Heidelberg*, pp. 561-569, available at: http://dx.doi.org/10.1007/978-3-642-01973-963

Oshagbemi, T. (2003), "Personal correlates of job satisfaction: empirical evidence from UK universities", *International Journal of Social Economics*, Vol. 30 No. 12, pp. 1210-1232, available at: https://doi.org/10.1108/03068290310500634

Ozturkcan, S., Kasap, N., Cevik, M. and Zaman, T. (2017), "An analysis of the Gezi Park social movement tweets", *Aslib Journal of Information Management*, Vol. 69 No. 4, pp. 426-440, available at: https://doi.org/10.1108/AJIM-03-2017-0064

Papis, J. (2006), "Understanding the workforce: the key to success in a youth hostel in Scotland", *International Journal of Contemporary Hospitality Management*, Vol. 18 No. 7, pp. 593-600.

Podgursky, M., Monroe, R. and Watson, D. (2004), "The academic quality of public school teachers: an analysis of entry and exit behaviors", *Economics of Education Review*, Vol. 23 No. 5, pp. 507-518.

Prochaska, J.O. and Di Clemente, C.C. (1982), "Transtheoretical therapy: toward a more integrative model of change", *Psychotherapy: Theory, Research and Practice*, Vol. 19 No. 3, pp. 276-288.

Prochaska, J.O., DiClemente, C.C. and Norcross, J.C. (1992), "In search of how people change: applications to the addictive behaviors", *American Psychologist*, Vol. 47 No. 6, pp. 1102-1114.

Prochaska, J.O., Redding, C.A. and Evers, K. (2002), "The transtheoretical model and stages of change", in Glanz, K., Rimer, B.K. and Lewis, F.M. (Eds), *Health Behavior and Health Education: Theory, Research, and Practice*, 3rd ed., Jossey-Bass, San Francisco, CA, pp. 226-231.

Prochaska, J.O., Butterworth, S., Redding, C.A., Burden, V., Perrin, N., Lea, M., Flaherty, R.M. and Prochaska, J.M. (2008), "Initial efficacy of MI, TTM tailoring, and HRI's in multiple behaviors for employee health promotion", *Preventive Medicine*, Vol. 46 No. 4, pp. 226-231.

Raman, R. and Pamod, D. (2015), "A study on user perception and awareness related to online privacy during online shopping", *Journal of Theoretical and Applied Information Technology*, Vol. 77 No. 3, pp. 382-389.

Raman, R., Vijayakumar Bharathi, S., Sesha, V. and Joseph, S. (2013), "Use of structural equation modeling to empirically study the turnover intentions of information technology professionals in Pune City", *Indian Journal of Science and Technology*, Vol. 6 No. 12, pp. 5612-5624.

Ramanathan, J. and Purani, K. (2014), "Brand extension evaluation: real world and virtual world", *Journal of Product & Brand Management*, Vol. 23 No. 7, pp. 504-515, available at: https://doi.org/10.1108/JPBM-04-2014-0559

Razaq, A., Yameen, M., Sabir, S., Iqbal, J.J., Ali, K.S. and Khan, M.A. (2011), "Impact of CSR, quality of work life and organizational structure on employee's performance in Pakistan", *Journal of Social and Development Sciences*, Vol. 2 No. 2, pp. 66-72.

Rice, J.K. (2003), *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*, Economic Policy Institute, Washington, DC.

Rivkin, S., Hanushek, E.A. and Kain, J. (2005), "Teachers, schools, and academic achievement", *Econometrica*, Vol. 73 No. 2, pp. 417-458.

Ruhul, A.M. and Nafeez, A.A. (2003), "Benchmarking learning outcomes of undergraduate business education", *Benchmarking: An International Journal*, Vol. 10 No. 6, pp. 538-558, available at: https://doi.org/10.1108/14635770310505175

Sahney, S., Banwet, D. and Karunes, S. (2008), "An integrated framework of indices for quality management in education: a faculty perspective", *The TQM Journal*, Vol. 20 No. 5, pp. 502-519.

Sanders, W.L. and Horn, S.P. (1994), "The Tennessee Value-Added Assessment System (TVAAS): mixed-model methodology in educational assessment", *Journal of Personnel Evaluation in Education*, Vol. 8 No. 3, pp. 299-311.

Seng, J.-L. and Yang, H.-F. (2017), "The association between stock price volatility and financial news – a sentiment analysis approach", *Kybernetes*, Vol. 46 No. 8, pp. 1341-1365, available at: https://doi.org/10.1108/K-11-2016-0307

Sirin, A.F. (2009), "Analysis of relationship between job satisfaction and attitude", *Journal of Theory and Practice in Education*, Vol. 5 No. 1, pp. 85-104.

Stinebrickner, T.R. (2001), "A dynamic model of teacher labor supply", *Journal of Labor Economics*, Vol. 19 No. 1, pp. 196-230.

Su, N., Reynolds, D. and Sun, B. (2015), "How to make your Facebook posts attractive: a case study of a leading budget hotel brand fan page", *International Journal of Contemporary Hospitality Management*, Vol. 27 No. 8, pp. 1772-1790, available at: https://doi.org/10.1108/IJCHM-06-2014-0302

Summers, A.A. and Wolfe, B.L. (1977), "Do schools make a difference?", *American Economic Review*, Vol. 67 No. 4, pp. 639-652.

Tsai, C.F. and Lu, Y.H. (2009), "Customer churn prediction by hybrid neural networks", *Expert System Application*, Vol. 36 No. 10, pp. 12547-12553, available at: http://dx.doi.org/10.1016/j.eswa.2009.05.032

Tschannen-Moran, M. and Woolfolk Hoy, A. (2001), "Teacher efficacy: capturing an elusive construct", *Teaching and Teacher Education*, Vol. 17 No. 7, pp. 783-805.

Wayne, A.J. and Youngs, P. (2003), "Teacher characteristics and student achievement gains: a review", *Review of Educational Research*, Vol. 73 No. 1, pp. 89-122.

Wilson, S. and Floden, R. (2003), *Creating Effective Teachers: Concise Answers for Hard Questions*, AACTE Publication, New York, NY.

Wilson, S., Floden, R. and Ferrini-Mundy, J. (2001), *Teacher Preparation Research: Current Knowledge, Gaps, and Recommendations*, Center for the Study of Teaching and Policy, University of Washington, Seattle, WA.

Worrall, L. and Cooper, C.L. (2006), "The quality of working life: managers' health and well-being", Executive Report, Chartered Management Institute, London.

Xue, W. and Deng, H. (2012), "Unstructured queries based on mobile user context", *International Journal of Pervasive Computing and Communications*, Vol. 8 No. 4, pp. 368-394, available at: https://doi.org/10.1108/17427371211283038

Zhang, X., Yu, Y., Li, H. and Lin, Z. (2016), "Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews", *Online Information Review*, Vol. 40 No. 1, pp. 119-145, available at: https://doi.org/10.1108/OIR-04-2015-0101

Zhao, Y., Li, B., Li, X., Liu, W. and Ren, S. (2005), "Customer churn prediction using improved one-class support vector machine", *Proceedings of the First International Conference on Advanced Data Mining and Applications, Springer-Verlag, Berlin and Heidelberg*, pp. 300-306, available at: https//dx.doi.org/10.1007/11527503 36

Further reading

Pinho, J.C. and Pinheiro, M.L. (2015), "Social network analysis and the internationalization of SMEs: towards a different methodological approach", *European Business Review*, Vol. 27 No. 6, pp. 554-572, available at: https://doi.org/10.1108/EBR-10-2014-0075

Corresponding author

Ramakrishnan Raman can be contacted at: raman06@yahoo.com