RESEARCH ARTICLE   **Open Access**

Impact of students evaluation of teaching: a text analysis of the teachers qualities by gender

Kingsley Okoye[1], Arturo Arrona-Palacios[1], Claudia Camacho-Zuñiga[2], Nisrine Hammout[3], Emilia Luttmann Nakamura[4], Jose Escamilla[1] and Samira Hosseini[1,5]*

* Correspondence: samira.hosseini@tec.mx
[1]Writing Lab, TecLabs, Vicerrectoría de Investigación y Transferencia de Tecnología, Tecnologico de Monterrey, Monterrey, Mexico
[5]School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Mexico
Full list of author information is available at the end of the article

Abstract

Today, modern educational models are concerned with the development of the teacher-student experience and the potential opportunities it presents. User-centric analyses are useful both in terms of the socio-technical perspective on data usage within the educational domain and the positive impact that data-driven methods have. Moreover, the use of information and communication technologies (ICT) in education and process *innovation* has emerged due to the strategic perspectives and the process *monitoring* that have shown to be missing within the traditional education curricula. This study shows that there is an unprecedented increase in the amount of text-based data in different activities within the educational processes, which can be leveraged to provide useful strategic intelligence and improvement insights. Educators can apply the resultant methods and technologies, process innovations, and contextual-based information for ample support and monitoring of the teaching-learning processes and decision making. To this effect, this paper proposes an Educational Process and Data Mining (EPDM) model that leverages the perspectives or opinions of the students to provide useful information that can be used to enhance the end-to-end processes within the educational domain. Theoretically, this study applies the model to determine how the students evaluate their teachers by considering the gender of the teachers. We analyzed the underlying patterns and determined the emotional valence of the students based on their comments in the Students Evaluation of Teaching (SET). Thus, this work implements the proposed EPDM model using SET comments captured in a setting of higher education.

Keywords: Educational innovation, Teacher-student evaluation, Sentiment analysis, Higher education, Learning process, Teachers' competence, Gender perspective, Technology adoption

Introduction

Technology-focused management and utilization of information about the teacher-student learning experiences can enhance the level of impact of educational-process initiatives on students' satisfaction. Studies have shown that, frequently, higher education institutions (HEI) rely heavily on the outcomes of the Students Evaluation of Teaching

(SET), not only for improving the different learning activities and decision-making strategies about the organizations but also for determining academic performance and assessment of teachers (Badri, Abdulla, Kamali, & Dodeen, 2006; Bianchini, Lissoni, & Pezzoni, 2013; Boring, 2017). Indeed, using information from the learning activities to support decision-making (i.e., learning analytics) (Ferguson, 2012; Gedrimiene, Silvola, Pursiainen, Rusanen, & Muukkonen, 2019; Papamitsiou & Economides, 2014), is useful for the dissemination of student-generated data to create enhanced support for the learners (Slade & Prinsloo, 2013). Also, it is essential to capitalize on the performance indicators when addressing the different challenges and developments associated with educational technologies (Perrotta & Williamson, 2018). For example, Badri et al. (2006) notes that SET has become a factor in promotion, long-term contracts, merits, and award-related decisions, and contract renewals for teachers in most institutions. Perrotta and Williamson (2018) note that how the different educational choices and activities are collectively managed mainly depends on the data collection procedures, the methods applied for the sample analyses, and the interpretation and communication of the results or findings from the data.

This study assumes that there is a need for emerging innovative methods to accurately extract contextually-based educational information from the extensive collection of data recorded in SET and to transform it into actionable insights for decision-making. Interestingly, Lau, Lee, and Ho (2005) note that quite often, it is challenging to combine the qualitative text data with quantitative methods for analysis. For instance, existing research focuses on information extracted from questionnaires or surveys completed by students. These are designed by the researchers to obtain student opinions; they pay little attention to their actual viewpoints or emotions, as written in the SET comments. In this paper, text mining (e.g., sentiment analysis) is considered as a method that supports educational process innovation and information management. Technically, text mining is capable of analyzing large volumes of data that are recorded in various scholastic databases (Lau et al., 2005) to derive new and relevant information for driving business operations forward. Examples include ensuring the right decision-making strategies and monitoring potential deviations or bottlenecks. We note that by using the information correctly, educational organizations can define procedures and policies aimed to maintain a strong relationship with the stakeholders (Payne, 2006; Piedade & Santos, 2010). The models and methodologies discussed thus far have shown to be essential to promote the personalization of users' experiences and outcomes, e.g., learning processes, teachers' performance, and competencies (Pedró, Subosa, Rivas, & Valverde, 2019; Sánchez, Domínguez, Blanco, & Jaime, 2019; Yadav & Berges, 2019).

The rationale of the study

This study proposes the Educational Process and Data Mining (EPDM) model to demonstrate the capabilities of text mining and its related technologies within the educational domain. The main research questions are as follows:

- How can we analyze the textual data in SET to derive useful information from the opinion of the students to support the different educational processes and decision-making strategies? and
- How can we utilize the derived data to understand how the students evaluate their teachers and their qualities, considering the gender of the teachers?

Accordingly, the study develops a set of constructs and uses it to perform the investigations and analyses as follows:

- We identify the most frequent terms used by the student to describe the teachers and the number of occurrences within the comments that reference the teachers' gender and differences.
- We also perform sentiment analysis to determine the intensity of the students' comments towards the teachers by gender.
- We determine the qualities that the undergraduates consider essential in the evaluation of the teaching-learning process and the performance of the teachers.
- Finally, we establish the implications of the statistical significance and differences by the gender of the teachers and provide an empirical discussion of the results.

Based on the stated research questions and effort to provide answers to the identified objectives, this study makes the following contributions to knowledge:

1. It defines an educational process and data mining model (EPDM) that leverages the perspectives or opinions of the undergraduate students; stated in the Students' Evaluation of Teaching (SET). In turn, the method provides useful information to enhance the end-to-end processes within the educational domain.
2. It demonstrates the capabilities of text mining technologies and their application within the educational domain.
3. It describes a method utilized to define how students rate their teachers' qualities and performance based on gender differences.
4. It illustrates how data about students' evaluations of teaching can be analyzed to provide solutions to curricular challenges in a competitive and rapidly changing educational environment. This is paramount during a time when SET data are stored at an unprecedented rate within the different educational information systems or databases. Moreover, there is an ever-increasing need to improve and support the teacher-student learning processes.

The remainder of this paper is structured as follows: The *Background Information* section discusses the relevant state-of-the-art in this topic area, especially as it relates to analyzing SET outcomes by gender of the teachers. The *Methodology* section introduces the proposed method, the EDPM model and its main components, the participants' information, instruments, data analyses, and case study implementation. In the *Discussion* section, this work covers the outcomes and the impact of the EDPM method, and then concludes and points out the direction for future studies in the *Conclusion* section.

Background information
Gender bias in students' evaluation of teaching and performance
Currently, existing research has increasingly questioned the impact and validity of SET as an indicator of teachers' performance and assessment (Boex, 2000; Boring, 2017; Drowling, 2000; Rivera & Tilcsik, 2019). For example, Rivera and Tilcsik (2019) note that researchers have paid little attention to the architecture of assessment and how

biased evaluations play a primary role in sustaining gender inequalities. Boring (2017) observes that SET recently has not just been used to assess teachers' performance but has also led to other studies related to teaching dimensions or qualities that students value in their teachers. For instance, examining them in terms of gender stereotypes. Perhaps, most of the scientific research has focused on SET and its impact on educational systems due to an increase in the rate at which program data is collected over the academic semesters (Bianchini et al., 2013). Indeed, there is a need to explore the potential of the information from the SET datasets for decision-making purposes. Along these lines, we show that analyzing the textual data in SET allows us to identify the students' learning expectations, to translate the speech or text into meaningful insights about the teaching-learning processes, and to discover patterns and correlation of words. All of these, in turn, can provide the ability to make effective educational-process-related decisions (e.g., monitoring of the teaching-learning process, and content recommendation, etc).

Moreover, contemporary studies provide new evidence on gender bias in teaching evaluations (Boring, 2017; Heath et al., 2019; Islahi & Nasreen, 2013; Kafedžić, Maleč, & Nikšić, 2018; Laube, Massoni, Sprague, & Ferber, 2007; Mengel, Sauermann, & Zölitz, 2019; Rivera & Tilcsik, 2019; Sabbe & Aelterman, 2007). In those studies and the different contexts, we note several assumptions about differences in teaching qualities and performances of the male and female teachers. For example, Sabbe and Aelterman (2007) note that male and female teachers are often presumed to differ in teaching styles, capacities, and effects considering the students' perspectives and teaching as a career. A majority of the studies argue that females receive systematically lower teaching evaluations than their male counterparts (Boring, 2017; Mengel et al., 2019). Boring (2017) notes that male teachers are perceived by both male and female learners as being more knowledgeable and having more persuasive class leadership skills, even though the students appear to learn more from the females than from male teachers.

Interestingly, this observation aligns with the results and findings of our study, as discussed in detail in the Results and Discussion section. Furthermore, Kafedžić et al. (2018) measure competencies related to teaching skills for the different genders. Islahi and Nasreen (2013) note the need for both substantive and methodological focus on gender-specific responsibilities and requirements. Such methodologies would enable stakeholders like educators, researchers, and learners to benefit from a more effective instruction by both male and female teachers. Likewise, Rivera and Tilcsik (2019) highlight that minor technical performance and aspects of the teachers can have a high impact on teacher evaluations when analyzed by gender.

On the other hand, Laube et al. (2007) note that a gender-view of SET is an artifact in itself and further state that quantitative measures can only mask underlying gender bias. We note that most existing literature has focused on quantitative analysis. Few studies emphasize quantifiable linguistic methods to determine the differences between the male and female teachers through the analysis of free-text comments (Heath et al., 2019) such as the method introduced in this current study.

Conceptualizing text mining for teaching and learning assessment

Text mining methods, including sentiment analysis and opinion mining, are theoretical and technological accomplishments of the last decade that are closely related to natural

language processing (NLP) (Pandey & Pandey, 2019), otherwise referred to as *language technologies* (Wen, Yang, & Rosé, 2014). The benefits of such methods and analyses can be understood through the different views or patterns that are contained in the databases of the organizations in question (Binali, Wu, & Potdar, 2009; El-Halees, 2011; Jones, 2019; Tur, Marín, & Carpenter, 2017). Interestingly, Wen et al. (2014) note that text mining can be adopted to reveal relationships between the actual processes (the learning processes in social networks, government, businesses, for example) and the intended stakeholders. The reactions or intentions of the stakeholders are seen in teacher-student interactions, the mood of Twitter tweets, political opinions, and even stock market fluctuations, etc. (Bollen, Mao, & Zeng, 2011).

For example, the experiments conducted by Wen et al. (2014) showed that the forums and online discussion platforms for students in the massive open online courses (MOOCs) could provide ample opportunity to engage learners, even in social processes (Brinton et al., 2014; Crues et al., 2018). For this purpose, they proposed a survival modeling approach to analyze the impact of students' opinions over time in MOOCs. The model was envisioned as a predictive and monitoring vehicle for determining the effectiveness of certain language behaviors or human expressions on the probability of an event occurring. Similarly, our study shows that through automated modeling and processing of the SET datasets, text mining is capable of providing new and vital information about the patterns or expectations of the undergraduate students for improvement. Besides, there may not even be the necessity to rely on the results of the performance assessments or the post-course evaluations.

Educational data mining and state-of-the-art

There has been a lot of progress in the current literature about using data obtained from the educational settings to improve the intended processes (Wang & Zhu, 2019; Abu Zohair, 2019, Exter et al., 2019; Munro, 2018; Dommett, Gardner, & van Tilburg, 2019; Alizadeh, Mehran, Koguchi, & Takemura, 2019). It is noteworthy that existing research has demonstrated that HEIs are increasingly data-hungry, with the stakeholders (students, teachers, and administrators, for example) seeking data and study results for their own needs and use (Clark, 2015; Williamson, 2018). According to Williamson (2018), the *Data Future* program will pave the way for new modes of quantification and standardization. Indeed, programs that are allied to business intelligence systems, adaptive learning tools and environments, educational/learning analytics which include several "smart learning tools" and dashboards can be utilized by educators to support the different decision-making strategies and processes (Perrotta & Williamson, 2018). Moreover, nowadays, datasets are increasingly being captured about the teaching-learning process, promoting the advancement of educational data mining techniques. These, in turn, have led to what HEIs today could refer to as technology-focused quality education. For example, Daniel (2015) notes that digitalization of the educational processes and services has led to significant innovations and has paved the way for the new notion of *datafication*. Meanwhile, Williamson (2018) notes that HEIs have experienced transformations in the past as new forms of digital data were

being generated, analyzed, and used to support their decision-making processes. In other words, datafication plays a vital role in understanding and enabling development within the broad context of the higher education sectors.

As presented in this study, one of the main areas where datafication and its associated technologies (such as the text mining technique) are currently being applied and are gaining attention in recent years is in the educational innovation teaching models. The new technologies are useful during the implementation of modern educational models because they can leverage methodologies that develop disciplinary and transversal competencies. Examples include challenge-based learning, memorable university experience, and inspiring professors. The technologies provide flexibility as to how, when, and where learning occurs to achieve the goals of HEI (TEC, 2018).

This work shows that the EPDM model can help provide innovative practices within the HEI setting that range from the adoption of intelligent methodologies to the transformation of students' learning experiences by enabling the so-called *three-dimensional expressive-communication-relational skills*. The series of experiments and analyses presented in this paper is useful for improvement of the teaching practices that students value, primarily because it addresses gender stereotypes and teacher-student engagement (Crues et al., 2018; Sánchez et al., 2019). The model we present is a Competency-Based Education (CBE) model (UNESCO, 2015; Yadav & Berges, 2019), and it is adequate and effective in supporting the new educational process initiatives and practices (TEC, 2018).

Information technologies for quality education and innovation

Several institutions have had to integrate emerging educational technologies in their different activities to maintain or sustain the credibility and quality of their business models or operations in general (Kori, Pedaste, & Must, 2018). Therefore, recent research (Wang & Zhu, 2019; Abu Zohair, 2019, Exter et al., 2019; Munro, 2018; Dommett et al., 2019; Alizadeh et al., 2019; Kori et al., 2018) has focused on overcoming some of the challenges within the educational processes to provide opportunities for innovative teaching and learning methods (e.g., through educational data mining).

Typically, Wang and Zhu (2019) compared the learning performances of users in a MOOC-based, flipped classroom with those in a traditional class. They observed that adopting advanced network technologies for teaching can support the transformation and pave the way for in-depth construction of high-quality education (Crues et al., 2018). Also, Abu Zohair (2019) notes that accurate data analysis has not only become essential for improving the students' performance and experience (Crues et al., 2018; Kori et al., 2018; Weston, Dubow, & Kaminsky, 2019) but it also serves as a mechanism for elevating the university's ranking and reputation. The results of the study by Abu Zohair (2019) show that support vector machines or algorithms can be used to train and analyze the readily available datasets to produce acceptable learning-process-classification accuracy and reliability test rates (Liao et al., 2019).

Accordingly, Mayer-Schönberger and Cukier (2014) note that the datafication of learning, i.e., the integration of information technologies and education, has opened new innovative methods for HEIs and brought about modern educational models. Along these lines, Jones (2019) shows that by tracking, aggregating, and analyzing the

student profiles and behaviors recorded in the information systems or databases (Kori et al., 2018), the universities are adopting modern educational methods or strategies. The technologies employed include educational data mining, machine learning, and learning analytics technologies. Moreover, those new methods are useful in the evaluation of the quality of educational models and processes (Alizadeh et al., 2019; Dommett et al., 2019; Munro, 2018). For instance, Dommett et al. (2019) found that the teachers and students quite often have different perspectives about the teaching-learning process. They argue that information technologies could potentially serve as the bridge from observed behaviors to a useful educational model. Nonetheless, existing studies have shown that the students, educators, and administrators have benefited significantly through the use of information technologies to predict academic achievements. These studies include views of the students' perceptions of learning and the mediating roles of academic engagement, including digital readiness within the HEIs. (Exter, Caskurlu, & Fernandez, 2018; Kim, Hong, & Song, 2019; Kori et al., 2018; Sánchez et al., 2019).

Methodology

Sentiment analysis for educational process innovation and data mining

Early research has examined how best to apply the educational models or strategies to provide students with productive and positive learning experiences (Altrabsheh, 2016; Crues et al., 2018; Mackness, Mak, & Williams, 2010; Wen et al., 2014). On the one hand, Altrabsheh (2016) shows that there is a need to explore the best method to visualize and understand the performance and results of the teaching-learning processes. On the other hand, Wen et al. (2014) investigate what sentiment analysis can tell us about the students' opinions with regards to the learning experiences (Crues et al., 2018). Notably, Moshinskie (2001) states that it is essential to monitor the students' feelings or emotions because learners with a positive attitude have been seen as more confident and motivated to learn (Weston et al., 2019). Besides, considering the students as consumers of the educational systems and curricula is one of the main reasons to use digital learning strategies as a way to personalize learning experiences (Munro, 2018; Romero, Saucedo, Caliusco, & Gutiérrez, 2019).

Nowadays, text mining techniques are being applied within the educational sector, allowing researchers to analyze textual data such as formal documents, e-mails, chat messages, digital notes, and online discussions for teaching-learning improvement and decision-making purposes. For example, Kumakawa (2017) notes that to exploit the goldmine of educational data and its use to understand better how the students progress with learning, researchers should essentially employ data mining techniques.

Figure 1 shows the Educational Process and Data Mining model (EPDM) proposed in this paper. Likewise, Altrabsheh (2016) observes that most of the existing works within the educational domain have only focused on sentiment analysis in the broad spectrum of e-learning (Kechaou, Ammar, & Alimi, 2011; Ortigosa, Martín, & Carro, 2014; Song, Lin, & Yang, 2007; Troussas, Virvou, Espinosa, Llaguno, & Caro, 2013), with fewer studies being carried out on feedback or SET results (Munezero, Montero, Mozgovoy, & Sutinen, 2013). Interestingly, Ortigosa et al. (2014) studied the polarization of emotions in social networks in the context of e-learning: The positive classifications

Fig. 1 Educational Process Data Mining (EPDM) model
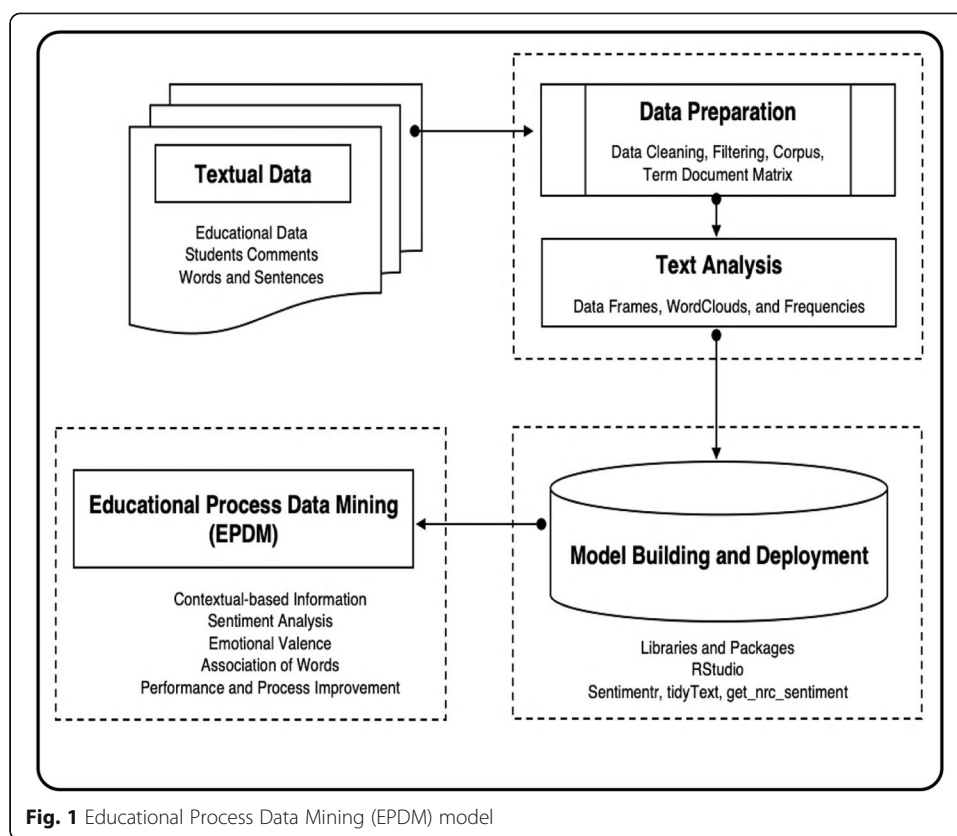
had good or optimistic feelings, whereas the negative classes were considered bad (aversive).

Thus far, one of the main contributions of the EPDM model is the text and sentiment analysis of comments given in SET. This study introduces the EPDM model for determining and analyzing the underlying sentiments (i.e., emotional valence) of the students based on their comments in SET, as demonstrated in the subsequent sections of this paper.

Educational process and data mining (EPDM) model

Figure 1 illustrates the architecture of the EPDM model that this study introduces for the text analysis of the educational data and case study implementation.

The EPDM model uses data from the educational domain and defines a method that allows for contextual-based information analysis, which is useful for improving the teaching-learning processes. The model is grounded on the following main components and design framework:

1. *Textual Data:* different sentences and words captured within the education domain. For this specific application, it consisted of the students' comments registered in SET. We extracted the feedback provided by the students at the end of their program when they rated the performance and teaching qualities of their instructors.

2. *Data Preparation:* involves filtering and cleaning of the available data to generate the library of words (corpus) used for the analysis. Thus, we cleaned out stop words, concatenated the same words which appeared in both singular and plural forms, and removed unnecessary and irrelevant words not required in the analysis. Then, we created a document matrix table for the analysis to follow.

3. *Model Building and Deployment:* performed using appropriate text-analysis tools, packages, and libraries to extract information that is used to understand the level of impact of the students' comments and emotions. For this study, we used RStudio (RStudio, 2018), an integrated development software that supports the sentimentr, pander, and tidyText packages, etc. which are used to handle textual data.

4. *Educational Process Innovation:* includes forming the contextual-based information for decision-making purposes and process-related monitoring and improvement. Thus, we extract, analyze, and visualize data that is used not only to answer the research questions developed in this paper but also to support the proposals and theoretical framework and results of this study.

Data sampling and participants

This study uses data from the Student Opinion Surveys to carry out the experimentations and practical implementation of the EPDM model. ECOA is a system designed for the collection of information about students' opinions about the course outcomes including the teaching and academic services offered throughout the various campuses of the host institution where this research took place. Specifically, we analyzed the data collected by ECOA from the undergraduates about their teachers' performances through a three-semester period (Semester Jan-May 2017, Semester Aug-Dec 2017, and Semester Jan-May 2018). The survey was carried out across six different campuses of the North Region of the host country, Mexico; covering around 14 Divisions or Schools, 78 Departments, and 1082 Courses. Therefore, a wide variety of students' views and opinions were represented. Although the ECOA contains some quantitative SET data, for this work, we focused on the comments provided by the students in response to the question, "Why would you recommend or not recommend the teacher, " which was also asked in the SET. To get practical, unbiased evaluations, ECOA does not record the names or the gender of the participants. Therefore, the dataset includes the gender of the teachers but not the gender of the participating students. From the ethical point of view, the students who provided the comments were informed about the purpose of the questions. However, they did not directly participate in the analyses performed in this study. The comments were provided by the students at the end of their respective courses when completing the ECOA survey.

Statistically, the dataset analyzed in this paper includes a total sample of $n = 93,915$ participants in the survey. However, after narrowing the data by filtering out the students who commented in the data, we noted a sample of $n_1 = 52,218$ comments for male teachers and $n_2 = 41,072$ for female teachers. However, this work uses the first 41, 072 comments for the male teachers and the $n_2 = 41,072$ for the females to conduct its analysis. This was done to distribute the data equally based on teachers' gender (variables) and for comparison purposes and validation of the results. In summary,

considering the gender of the teachers and the students who answered the survey, we used a total sample of $n_t = 82{,}144$ comments (i.e., $n_1 = 41{,}072$ for male teachers, and $n_2 = 41{,}072$ for female teachers) throughout the analysis and experiments presented in this paper.

Experimental setup

The text analysis method or EPDM model implementation was carried out to determine:

- The frequency of words by teachers' gender.
- The different word clusters and the correlations between the most frequent terms used by the students to describe the teachers.
- Emotional valence indicated by the comments and their impact, analyzed by gender.

Main tools and instrument

- Word Cloud and correlational analysis.
- Sentiment analysis.

Data analysis

To implement the different phases of the EPDM model as described in Fig. 1, we began the work by building a corpus (library of words) to allow for the subsequent modeling and implementation of the text-based analysis, features, functions, packages, and automation. It is also important to note that the comments given by the students were provided and analyzed in Spanish. However, this work reports the results in both Spanish and English to cover the objectives of this paper for an international audience.

Implementation and results

This study implements the EPDM model by first determining the frequency of each of the top twenty qualities that the students perceived as necessary in the evaluation of the teachers' performance across the data (ECOA). Technically, the sets of qualities were analyzed in terms of male and female teachers by determining the different clusters of words, frequencies, and correlation of words. The word frequencies were presented in highest to lowest order. The word correlations were determined by measuring the probability of the students using the associated terms in the data. The measurements fall between 0 and 1, representing 0% to 100%. The results are shown in Figs. 2 and 3 and Table 4 in appendix.

  ° Frequencies and word correlations are reported in Table 4 and Appendix.

  The EPDM model applied to the comments of the students allowed us to determine the top twenty qualities that the students considered necessary when evaluating the teachers' performance. In Fig. 2, we list them in two Word Clouds (English and Spanish), classified according to the gender of the teachers. In Fig. 3, we plot them in a bar chart. It is noteworthy that 19 of those terms are similar, and only one was different for the genders. Henceforth, "problems" only appear differently for males, while "patience"
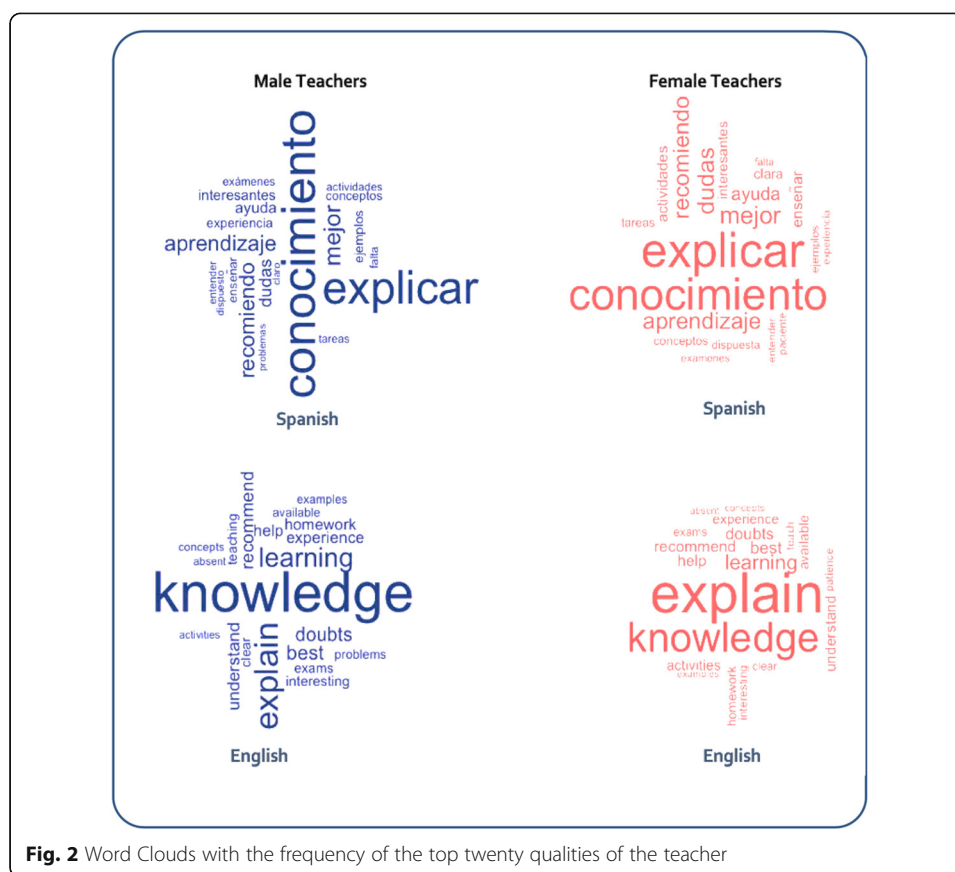
Fig. 2 Word Clouds with the frequency of the top twenty qualities of the teacher

only appears for females. Whereas, "problems" can be described as tools or mechanisms presented by the teachers to improve and evaluate the learning processes, "patience" is a term related to the personal abilities and performance of good teachers. Table 4 in appendix reports the word frequency in decreasing order, along with their correlations. The most mentioned terms are "conocimiento" (*knowledge*) and explicar (to *explain*); the performance of the teachers associated with these terms are dependent statistically by gender. "knowledge" is the most mentioned for male teachers, while "explain" is the most indicated for the females.

As previously mentioned, the total number of comments evaluated were 41,072 for each gender. Thus, Table 4 and the Appendix show the word "knowledge" appearing in 17.9% of the comments associated with male teachers and 15.4% of those with female teachers. The term "explain" appears in 16.8% of female teachers' comments and 17.0% of the male teachers. Interestingly, even though the word "experience" appears in both lists, for male teachers, it occupies the 9th place, and for female teachers, it holds the 20th place.

The outcome of the analysis of the comments shows that the students in the broad spectrum consider particular teachers' qualities as essential, independent of the SET scores. Although those qualities seem to be closely related in both genders, the frequency of use or the word correlations vary accordingly. In Fig. 2 and Table 4 in appendix, we note a plethora of essential words that were used by the students to

Fig. 3 The top 20 most frequently used words in students' comments about their teachers' performance

describe the qualities of the teachers. Table 4 in appendix also reports the correlation of words, which is the probability (or likelihood) of students using the associated words in their comments. Consistently, for teachers of both genders, the word "conocimiento" (*knowledge*) appears in phrases like "transmits knowledge," "wide knowledge," and "a lot of knowledge." The word "explicar" (infinitive of *explain*) appears in phrases expressing that the teacher makes something clear or more evident. Moreover, the first row in Table 4 in appendix shows that such phrases correlate significantly with the highest-rated word for male teachers, "conocimiento" (*knowledge*). This denotes that the students, on the one hand, consider the male teachers as more knowledgeable or intellectually deeper than their female counterparts. For female teachers, the words "clara" and "claras" (singular and plural feminine for the English word "clear," from "clarify") significantly correlated with the most common term associated with the females, namely "explicar" (to explain). This means that the students hold the teaching styles and abilities to pass across information; of the female teachers in higher regard than their male counterparts.

Furthermore, this study considers it essential to determine the weight (emotional valence) of the different combinations of words/comments given by the students. To do it, we determined the average sentiment of the individual student's comments, as shown in Table 1, and then established the foundation shape, represented as a graphical impact view of the different comments (see: Fig. 5). Typically, the EPDM allows us to analyze the sentiment scores, which refers to the sentiment of the individual comments by each student represented in each matrix in the corpus. Thus, the resultant values represent the average sentiment of the sentences of individual students, as shown in Tables 1 and 2, and Fig. 4.

Technically, we utilize the get_sentiment and get_nrc_sentiment function in R (RStudio, 2018) to determine the word count and the average sentiment scores with standard deviation. The sentiment scores are represented as interval values between − 2 and 2, denoting the level of impact for each of the comments, as shown in Table 1. For example, in Table 1, we show the first and last five comments in the datasets to expand on the method, whereby the comments with positive interval values represent a positive (goodness) sentiment, and comments with negative values mean an aversive feeling. Overall, as shown in Table 2 and Fig. 4, we found that male teachers received the most negative comments (sentiments) with a min value of − 1.12 in comparison to − 1.00 received by the females. However, the female teachers also received the most positive sentiments with a max value of 1.817 as opposed to the male counterparts who received an average of 1.570, respectively.

Henceforth, we assume that by visualizing the frequency or projection of the positive (attractiveness/goodness), neutral, and negative (aversiveness/badness) emotions of the students, we can identify the significance and/or the differences of impact among the male and female teachers. The visualization of those results is displayed in Fig. 5. There, we can identify positive, neutral, and negative emotions among the comments of the students, although it can sometimes be tough to discern any trend or behavior across the data. By utilizing the Syuzhet package with the simple plot function in R, we were able to visualize the data to soften (smooth out) and normalize it (i.e., between − 1 to 1, as shown in the upper and lower plots of Fig. 5. Interestingly, we observe that

Table 1 Word count and average sentiment scores for individual student comments

Element id	Male Teachers		Female Teachers	
	Word count	ave_sentm	Word Count	ave_sentm
1:	41	0.000	5	0.000
2:	2	0.000	4	0.000
3:	10	0.000	51	0.000
4:	15	0.000	6	− 0.102
5:	19	0.000	21	− 0.055
–	–		–	
41,068:	33	−0.044	22	0.000
41,069:	13	0.000	10	0.079
41,070:	92	−0.030	11	−0.075
41,071:	11	− 0.151	10	0.000
41,072:	14	0.000	1	0.000

Table 2 Summary of the sentiment scores across the data by gender

	Male Teachers			Female Teachers		
	Word count	sd	ave_sentm	Word count	sd	Ave_sentm
min	0	0.00	−1.12	0	0.00	−1.00
median	12	0.01	0.00	12	0.02	0.00
mean	17.69	0.05	−0.004	17.41	0.05	−0.005
max	744	1.17	1.570	386	1.19	1.817

Notes: element_id denotes the individual comments by the students
word_count accounts for the number of words in each comment
sd represents the standard deviation, and
ave_sentm is the score (valence) for the corresponding comments

most of the comments are centered evenly around the zero value, as noted earlier in Fig. 4.

Similarly, as seen in the lower plot in Fig. 5, the majority of the comments are clustered around the zero-point on the scale, which means that a large number of the comments provided by the students are considered neutral (see: Fig. 6). This observation was found for both genders of teachers. Nevertheless, considering the simplified macro shape (presented in the upper part of the graph) and the scatter plot (lower part of the graph), we found that the male teachers consistently received highly positive (+) comments, as opposed to the female teachers, who received these in an up and down manner. Although from the same graph (Fig. 5), it can be said that the teaching performance of females was slightly more favorable than those corresponding to males given the margin of comments, which are considered to be positive. In other words, as gathered in Table 2 and Fig. 5, the female teachers received the highest number of positive comments (sentiment) on average (max 1.817). In any case, there were some comments below the zero margin, which are considered as negative (−) for both genders.

To quantify the levels of emotional intensity for purposes of confirming and validating the results, we applied the EPDM model to establish the valence scores, otherwise referred to as the measures of the emotional valence. To do this, we applied the sentimentr miner (get_nrc_sentiment) algorithm in R to extract the different scores for each
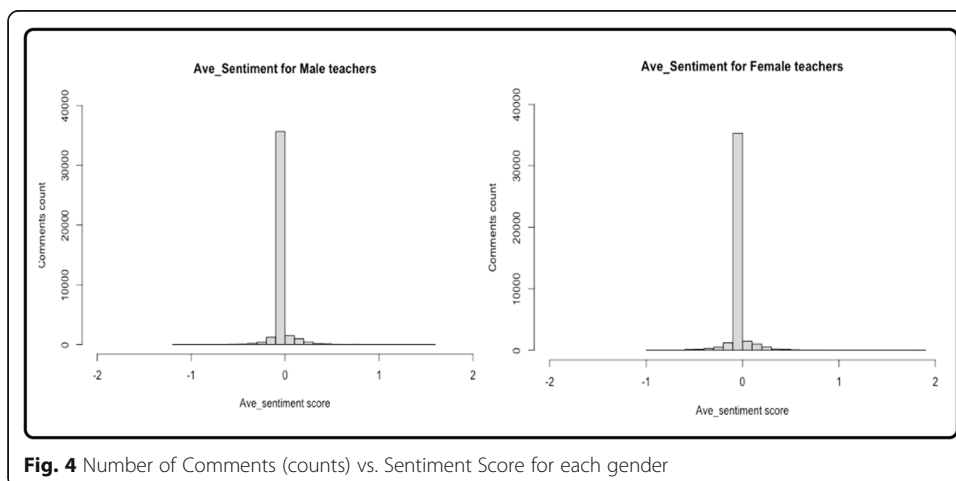

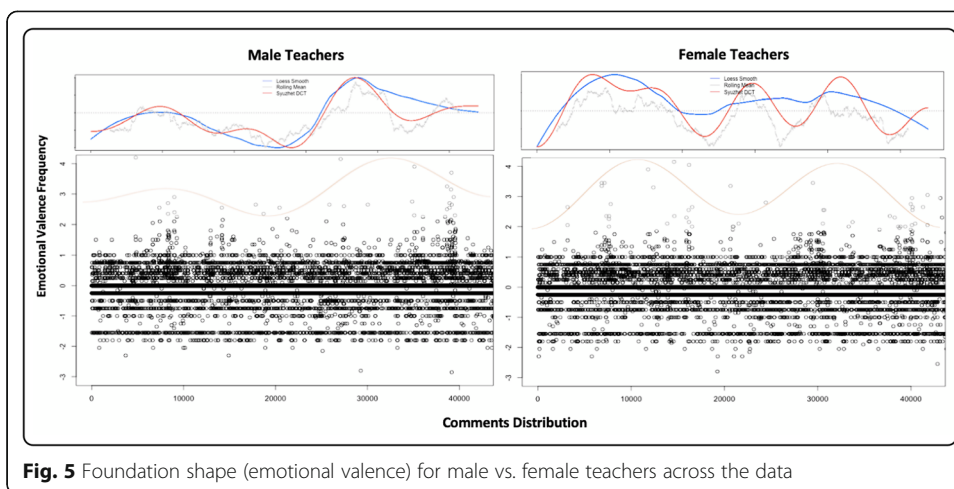**Fig. 4** Number of Comments (counts) vs. Sentiment Score for each gender

Fig. 5 Foundation shape (emotional valence) for male vs. female teachers across the data

comment about the male and female teachers. This is obtained from the intensity of the words used in the text by using positive (++), neutral (0) and negative (−−) values, as shown in Fig. 6.

Although the intensity element represents attractiveness (goodness), neutral, and aversiveness (badness) classifications of the comments, respectively, the scale also shows the intensity of emotions. As an example of the scores considering the different genders (i.e., male and female teachers), in Table 3, we show the emotional valence scores of the first 323 comments. The *Comments* column, [1] to [307], represents the id of the first individual comment within each row. The comments with positive values indicate an attractive valence (emotion), whereas the negative values indicate an aversive valence, and the zeros represent a neutral comment. Valence scores for male teachers rank between − 4 to 8, while those given to females rank between − 4 to 6. We summarize the different valence scores in Fig. 7, revealing that both genders received or showed similar distributions. In Fig. 8, we report some examples of specific comments and their associated emotional valence.
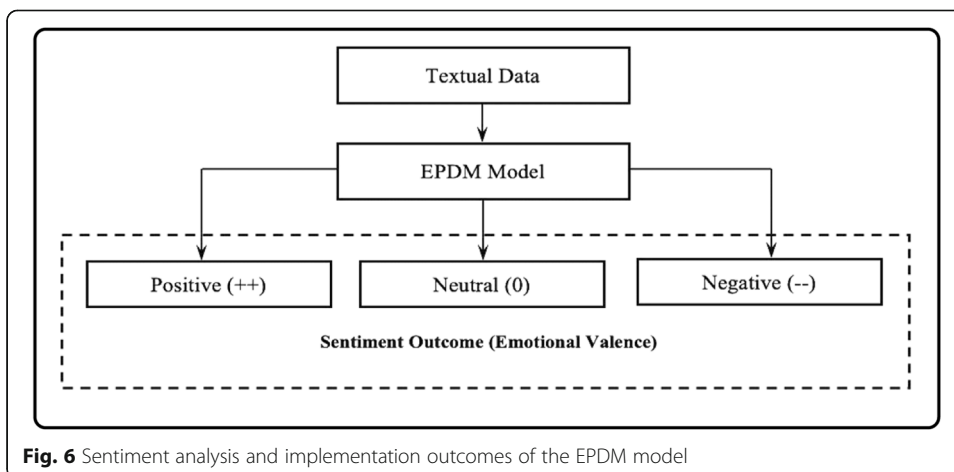

Fig. 6 Sentiment analysis and implementation outcomes of the EPDM model

Table 3 Fragment of the valence scores for the teachers by gender

Valence Score		
Comments	**Male Teachers**	**Female Teachers**
[1]	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -2 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
[18]	1 0 1 0 0 0 0 0 -1 0 -1 0 0 0 0 0	0 0 0 0 0 0 0 0 0 -1 0 0 0 1 0 0 0
[35]	0 0 0 1 0 0 0 0 0 0 0 0 -1 -1 0 0 0	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[52]	-1 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 -1	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1
[69]	1 0 0 -1 1 0 -1 -1 0 0 1 0 0 0 0 0 -1	0 0 0 0 0 0 0 0 0 -1 1 0 -1 0 0 0 0
[86]	0 0 0 -1 0 0 0 0 0 -1 0 -1 0 -1 0 0	0 0 0 0 0 0 0 -1 -2 0 -3 0 0 0 -1 0 0
[103]	0 0 0 1 0 0 0 0 0 0 0 0 0 -1 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[120]	0 0 0 0 -2 0 0 0 0 1 0 0 0 0 1 0 0	-2 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
[137]	0 0 0 -1 0 0 0 -1 0 -1 0 0 0 0 0 0	0 0 0 0 0 0 0 -2 0 0 0 0 0 -2 0 0 -2
[154]	0 0 0 -1 0 0 0 0 0 0 0 0 -1 1 0 0	-1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0
[171]	0 0 0 0 0 -1 0 0 0 0 0 0 -2 0 0 0	0 0 0 0 0 -3 0 0 1 0 1 0 0 1 -1 0 -1
[188]	0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0	0 0 0 0 1 0 0 -2 0 0 0 0 0 0 0 0 0
[205]	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	0 0 0 0 -1 -1 -2 0 0 0 0 0 0 0 0 0
[222]	-1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
[239]	0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 -2	0 0 0 0 0 0 0 -2 0 0 0 0 0 0 0 0 0
[256]	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1 1 0 0 -1 0 0 -1 0 0 0 0 0 0 0 0 0
[273]	0 0 0 0 -1 0 0 0 -1 0 0 0 0 0 0 1 0	0 0 0 0 0 0 0 -1 0 0 -1 -1 0 -1 0 0 0
[290]	0 0 0 0 0 0 1 0 0 0 0 -2 0 0 0 0 1	0 0 -2 0 0 0 0 0 0 0 0 0 0 0 0 0 0
[307]	0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 -1 0 0 -1 0 0
min	-4.00	-4.00
median	0.00	0.00
mean	-0.009174	-0.02743
max	8.00	6.00

In Fig. 9, we report the overall sentiment analysis of the students' comments about their teachers' performance, by gender of the teachers. We look at both genders and the different qualities that the students deem as crucial in the SET evaluation. There have been several definitions of emotional perception and analyses of these in the literature, especially highlighting its application in various domains (Kort, Reilly, & Picard, 2001; Litman & Forbes-Riley, 2004; Nosu & Kurokawa, 2006; Shen, Wang, & Shen, 2009; Tian, Zheng, & Zheng, 2010). Primarily, the EPDM model applies the sentiment classifications within the educational domain, as observed in Litman and
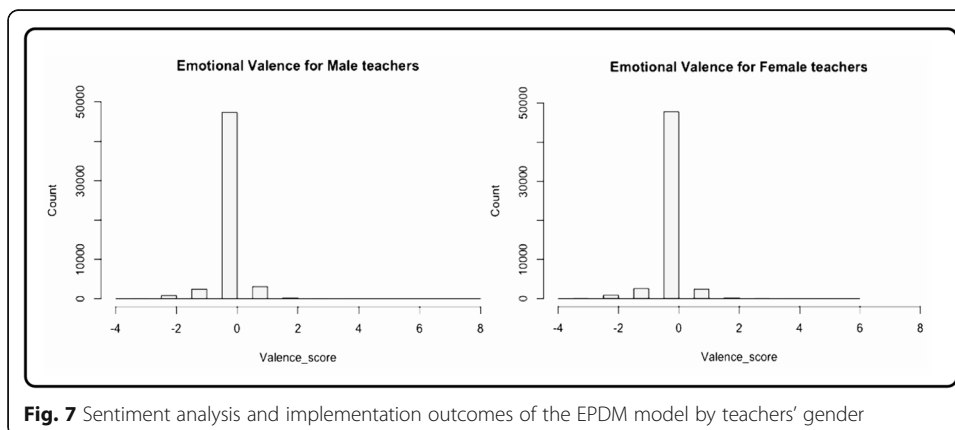

Fig. 7 Sentiment analysis and implementation outcomes of the EPDM model by teachers' gender
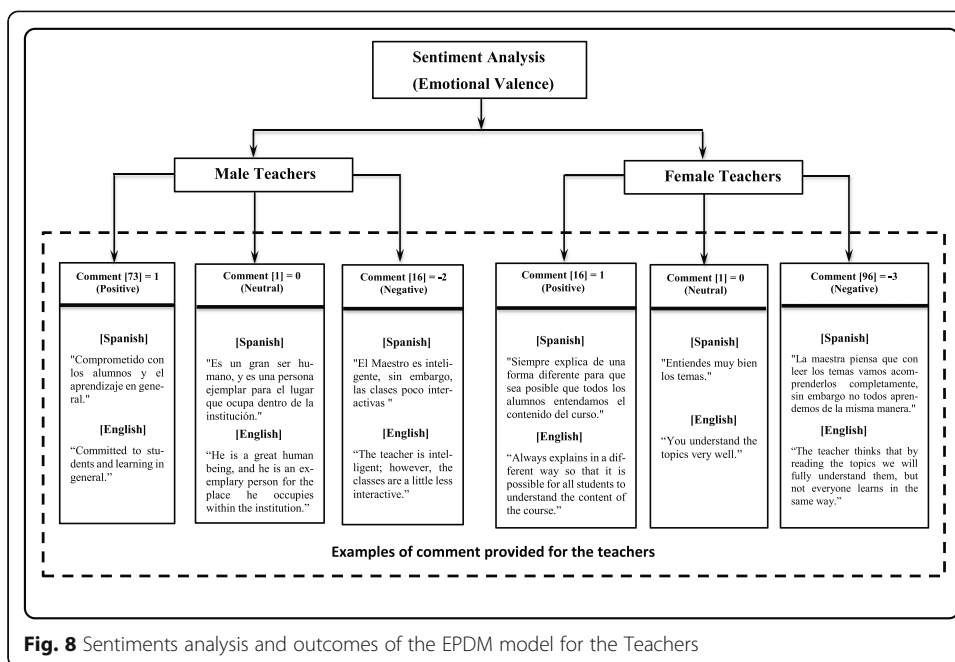
Fig. 8 Sentiments analysis and outcomes of the EPDM model for the Teachers

Forbes-Riley ([2004](#)). It is noteworthy that the feelings are slightly the same for both genders. While some elements of anxiety, resentment, dissatisfaction, and uncertainty appear in most of the optional comments or remarks, motivation, contentment, and amazement appear at a lower extent for both genders. The analysis reveals that the students choose whether or not to write comments about their teachers, but often these may reflect negative sentiments. Interestingly, another notable result is that for male teachers, the remarks or comments most frequently reflect confidence, which tends to appear less often for female teachers.

Discussion

This paper introduces the Educational Process and Data Mining model (EPDM) (Fig. [1](#)) as a method towards discovering and understanding the different sets of
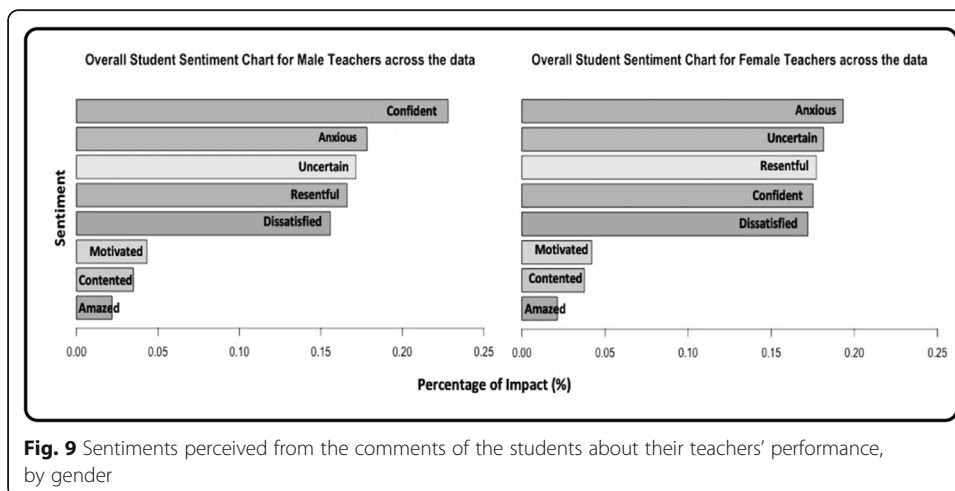

Fig. 9 Sentiments perceived from the comments of the students about their teachers' performance, by gender

qualities that are considered necessary for teaching-learning processes. The current research has shown that students perceive the qualities of the teachers to a significant level. Their comments include feedback on the teaching advice and interactions, intellectual challenge, guidance and supervision, and overall performance, for example. Moreover, this study also indicated that the perception of teachers' qualities could vary by the gender of the teachers. Interestingly, early studies has shown that the post-course or semester evaluations in SET, to some extent, are significantly biased against female teachers. Also, many of the educational institutions rely exclusively on SET for decision-making purposes regarding the teachers (Boring, Ottoboni, & Stark, 2016; Whitney, Hayter, & Marshall, 2019).

The findings on how the students evaluate their teachers' performance through the ECOA SET (see: the section on Data sampling and participants) revealed more comments for male teachers (52218) than for females (41072). Perhaps, this could be due to the higher number of male teachers across the institution. However, this work only used the first 41,072 comments for male teachers in its analysis for equality of feedback between the two genders and for conformability and validation of the results. We noted that 19 of the 20 words with the highest frequency were the same for the teachers, regardless of gender (see: Table 4 in appendix and Figs. 2 and 3).

Also noteworthy is the fact that the two most mentioned words were "knowledge" and "explain," the first one being on top for males while the second one led for females. Although Table 4 in appendix does not show whether the comments were positive or negative, this reflects gender bias. The students valued female teachers on their teaching methodology or the way they were able to explain clearly the concepts/materials that they teach; whereas, they valued male teachers for the knowledge they could demonstrate.

Although the frequency of the words appearing in Table 4 in appendix varies by gender, we note that the only difference between male and female when considering the top twenty words were "problem" and "patience." While the term "problem" is used to refer to males, "patience" is used to refer to females, respectively. Consequently, in Table 4 in appendix, we link the two words (problem and patience) with the correlated words to deduce how the terms relate to the different genders independently. Thus, while words such as problems (problemas), resolve (resolver), and imagination (imaginación) were correlated for the male teachers, the term "patience" did not correlate with any other in the data. Nonetheless, these observations also aligned with or supported the fact that the students viewed the male teachers as more knowledgeable and the female teachers as having a better ability to explain concepts and topics very well. Perhaps, this was also the reason why the students used the attribute "patience" to describe the female teachers.

Some examples of specific comments provided by the students related to male teachers were "Smart teacher; however, there is little interaction in classes," (*El Maestro es inteligente, sin embargo, las clases poco interactivas*) and "He has general knowledge of all topics." (*Tiene conocimiento general sobre todos los temas*). Comments related to female teachers included "You understand the topics very well " (*Entiendes muy bien los temas*) and "She always explains differently to make it possible for all

students to understand the course content. She is always patient and kind." (*Siempre explica de una forma diferente para que sea posible que todos los alumnos entendamos el contenido del curso. Siempre es paciente y amable*). We confirm that the noted affirmations and statements aligned with the results of this study and settle that the outcome and results of the EPDM method described in this paper are strongly evident.

In this study, it is noteworthy that nine of the questions in the ECOA SET survey were mandatory and multiple-choice, while the comments from which the data was extracted were optional. We believe there is a need for educators to emphasize and explore more the impact of the students' opinions on the teacher-student dynamics in learning and on the educational processes/management at large. Moreover, Slade and Prinsloo (2013) note that for the stakeholders to achieve an effective datafied-education ethically, the students should partake in defining the context, conditions, and purposes for which the data collected from them are used.

The sentiment chart (Fig. 9) demonstrates that the majority of the available comments show adverse emotions to some extent. It could be that the students who write negative comments do so because they have bad evaluations from their teachers. The figure (Fig. 9) also shows that the assessment of the teachers correlates strongly with feelings beyond satisfaction or dissatisfaction. Perhaps, the two affirmations mentioned previously could suggest that SET is not an objective evaluation for teachers' performance. Therefore, a more profound and cross-related analysis with descriptive statistics should be done, such as the text mining method and analysis described in this paper.

We found that even though the students tended to confide in the male teachers, there significantly existed also some elements of anxiousness, uncertainty, resentfulness, and dissatisfaction. For the female professors, the overall sentiment ratings showed a mixture of both positive and negative feelings, which perhaps confirms the results of the analysis of emotional valence, as represented in Fig. 5, where the emotional slope moves up and down intermittently. In any case, the students appear to be motivated, contented, and amazed by the teaching qualities of the professors or tutors. Moreover, it is essential to mention that most of the comments provided by the students are considered neutral with little or no sentiment (emotions) attached (see Figs. 4, 5, and 7).

In the wider spectrum of scientific research, several works have studied the impact of the state-of-the-art methodologies on the teacher-student learning process (Engen, 2019; Gallego-Arrufat, Torres-Hernández, & Pessoa, 2019; Gordillo, López-Pernas, & Barra, 2019; Silva, Usart, & Lázaro-Cantabrana, 2019). Particularly, Silva et al. (2019) note that a larger scale of the early indicators and success factors have been as a result of using the information or digital technologies to improve teaching and professional development (Kori et al., 2018). Whereas, Engen (2019) argues that there is also a need for a greater understanding of new technologies and their effective use by stakeholders such as teachers and students. They (Engen, 2019) argue that the different cultural and social aspects of the users of digital technologies for transforming modern educational institutions and models should be considered. Interestingly, the work of (Silva et al., 2019) comments that there has been much progress demonstrated in organizational management, teaching spaces, planning, and technology to support the student learning

process. However, evidence from the existing works has shown that male students achieved a higher level of competencies compared with female students (Silva et al., 2019). Perhaps, future research questions might address whether such differences in skills also exist for male and female teachers (Sánchez et al., 2019; Exter et al., 2019; Yadav & Berges, 2019; Crues et al., 2018).

Nonetheless, the results of this paper show that by analyzing the real comments (opinions) of the students, we were able not only to provide a better way to understand their feelings and what they expect from the teachers but also what would help inform the decision-making processes and strategies of higher education institutions. Not only does the EPDM method show a high level of aptitude to offer solutions for the different educational frameworks and practices, but also the model can provide new and better ways to monitor and improve the educational processes.

There are two main drivers for the method introduced in this research. First, data about student evaluations of teaching are captured and stored at an unprecedented rate within the educational information systems/databases. Second, there is now more than ever an increasing need to improve and support the teacher-student learning processes because of the competitive and rapidly changing educational environment and curricula. Thus, we developed the EPDM model as a theoretical bridge for the aforementioned gaps or challenges. Moreover, the EPDM design framework can easily be applied by the owners of educational processes, innovators, and educational advisory boards to understand the organizational structures and policies they need for informed teaching-process decisions and the provision of valuable support to the stakeholders. This also includes improvement and monitoring of the various activities that underly the educational processes and learning in general.

The concept of *datafication* used and explained in this study cannot be fully emphasized or practically applied without acknowledging the ethical considerations that surround the resultant methods. Therefore, we turned our attention to some of the ethical implications, especially the sociotechnical perspective on data usage in education. Whereas Slade and Prinsloo (2013) highlight the need to harvest data under provisions that ensure trust among the different institutions and students, we took steps to ensure that the contexts in which the readily available datasets were analysed and the suggested predictions were made are within the social construction and moral objectives of technical expertise (Perrotta & Williamson, 2018; Prinsloo, 2017; Slade & Prinsloo, 2013). We bore in mind the need for suitable methods of data-driven segmentation and diversification (Perrotta & Williamson, 2018) and the ethical framework and practices that require the HEIs to offer context-appropriate solutions or strategies that increase the quality and effectiveness of teaching and learning processes (Slade & Prinsloo, 2013). Whereas Prinsloo (2017) notes that the socio-technical imaginary of HEIs and algorithmic decision-making methods (such as the EPDM model developed in this paper) offer huge potential; we must also acknowledge the risks and ethical concerns that are attached to such data-driven methods. Besides, it does no harm to map out the opportunities and threats in applying new algorithmic or technology-focused decision-making practices within the higher educational settings, especially when

adequate procedures such as sensing, processing, acting, and learning are put in place (Prinsloo, 2017).

Indeed, this study has identified and discussed the different problems with current tools and methods that support the SET evaluations, especially as it concerns the genders of the teachers. This work considers how to resolve the identified challenges, especially regarding new technologies and data-driven approaches such as text mining (natural language processing) in the education domain. To this effect, this paper proposes the EPDM model, which shows to be effective in the analysis of the textual data collected in SET for improved educational process management and technology-based decision-making. However, while we believe that the method is suitable for a contextual-based analysis of the different opinions and perspectives of the students in understanding the teacher-student processes, there might also exist a number of limitations or threats to its validity. For instance, although this paper has introduced the framework for the sets of descriptive and quantifiable text analyses that provide ample understanding of the student comments found in the datasets, there could be potentially many ways to approach this, and there may be bigger areas of considerations not yet addressed. This is due to the fact that the text analysis methods (e.g. sentiment analysis) represent a new area within the educational domain, and there are not many tools or methods that support this approach currently in the literature. Therefore, the study assumes that this work is an incentive and methodological road map for more robust and intensive research to come, particularly within the broad and overlapping field of educational-process mining and innovation.

Conclusion

This paper shows that the textual data gathered in students' evaluation of teaching (SET) can provide insightful and strategic decision-making information with regards to the teacher-student learning processes. To this end, the paper proposed an Educational Process Data Mining model (EPDM) that leverages the opinions or perspectives of the students to provide useful information to understand the relations or correlation of words and sentiments of the students towards their teachers by considering the gender differences. Fundamentally, the work shows that the state-of-the-art idea of text mining for educational process innovation can be employed to provide a more robust analysis of the students' comments or viewpoints, and consequently, adopted or used by the educational process owners or advisories. Practically, this study applies the EPDM model on a case study of students' comments given in a SET for the private university, in order to demonstrate the application of the method in real-time. Theoretically, the study presumes that it must become the responsibility of educators to employ and apply the EPDM model in their current teacher-students' evaluations and experiences in order to ensure a robust and effective educational process innovation and curricula. Future works can adopt the proposed model and text mining approach described in this paper to analyze datasets about any given process domain, or yet, modify the proposed model to include further types of processes or data analysis that may have not been already introduced in this paper.

Appendix

Table 4 Top qualities, frequencies, and correlation of words for the male and female teachers

Male Teachers

Frequency of Words		Correlation of Words			
Spanish word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Conocimiento (knowledge)	7344 (17.9%)	Transmitir (transmit) 0.18	Amplio (wide range) 0.17	Muchísimo (a lot) 0.1	–
Explicar (explain)	6963 (17.0%)	Claras (clear, f., plural) 0.19	Claro (clear) 0.12	–	–
Mejor (Best)	3101 (7.6%)	Profesores (teachers) 0.22	Maestros (male teachers) 0.21	–	–
Aprendizaje (learning)	2796 (6.8%)	N/A 0	–	–	–
Recomiendo (recommend)	2752 (6.7%)	N/A –	–	–	–
Dudas (doubts)	2018 (4.9%)	Resuelve (resolves) 0.37	Aclarar (to clarify) 0.19	Responde (responds) 0.14	Dispuesto (available) 0.13
Ayuda (help)	1653 (4.0%)	Dispuesto (available) 0.25	–	–	–

Female Teachers

Frequency of Words		Correlation of Words			
Spanish Word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Explicar (explain)	6899 (16.8%)	claras (clear, f., plural) 0.15	clara (clear, f.) 0.11	–	–
Conocimiento (knowledge)	6342 (15.4%)	transmitir (transmit) 0.2	amplio (wide-range) 0.15	muchísimo (a lot) 0.1	–
Mejor (best)	3088 (7.5%)	profesoras (female professors) 0.14	–	–	–
Dudas (doubts)	2991 (7.3%)	resuelve (resolves) 0.29	aclarar (to clarify) 0.15	responde (responds) 0.13	disponible (available) 0.13
Aprendizaje (learning)	2829 (6.9%)	Dependen (depend) 0.11	Exploratoria (exploratory) 0.11	Resolvían (resolved) 0.11	Externamente (external) 0.1
Recomiendo (recommend)	2798 (6.8%)	N/A 0	–	–	–
Ayuda (help)	2019 (4.9%)	dispuesta (available) 0.21	–	–	–

Table 4 Top qualities, frequencies, and correlation of words for the male and female teachers *(Continued)*

Male Teachers

Frequency of Words		Correlation of Words			
Spanish word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Interesantes (interesting)	1478 (3.6%)	N/A — 0	–	–	–
Experiencia (experience)	1244 (3.0%)	amplia (wide-range) 0.14	laboral (work) 0.13	profesional (professional) 0.12	Campo (field) 0.11
Ejemplos (examples)	1213 (2.9%)	Reales (real, plural) 0.18	Bueno (good) 0.16	Espeficar (speculate) 0.14	Claros (clear) 0.13
Conceptos (concepts)	1152 (2.8%)	Claros (clear, m., plural) 0.19	Aprendidos (learned, m., plural) 0.12	Entender (to understand) 0.1	–
Enseñar (teach)	1094 (2.7%)	N/A — 0	–	–	–
Exámenes (exams)	969 (2.4%)	Admite (admits) 0.22	parcial (mid-term) 0.2	compañero (peer) 0.17	Tareas (homework) 0.11
Falta (absent or lack)	965 (2.3%)	N/A — 0	–	–	–
Actividades (activities)	940 (2.3%)	N/A	–	–	–

Female Teachers

Frequency of Words		Correlation of Words			
Spanish Word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Enseñar (teach)	1539 (3.8%)	N/A — 0	–	–	–
Actividades (activities)	1474 (3.6%)	ayudan (help) 0.12	–	–	–
Interesantes (interesting)	1330 (3.2%)	N/A — 0	–	–	–
Clara (clear)	1279 (3.1%)	dispuesta (available) 0.15	–	–	–
Tareas (homework, tasks)	1045 (2.5%)	Encarga (assignment) 0.16	múltiple (multiple) 0.12	Exámenes (exams) 0.11	Proyectos (Projects) 0.1
Conceptos (concepts)	1012 (2.4%)	Claros (clear, m. plural) 0.17	Entender (to understand) 0.12	Exploratoria (exploratory) 0.11	Narrativo (narrative) 0.11
Ejemplos (examples)	934 (2.3%)	Bueno (good) 0.18	reales (real) 0.12	Claros (clear, m. pl) 0.12	–
Dispuesta (available)	897 (2.2%)	Apoyar (to support) 0.14	ayudate (help you) 0.12	Dudas (doubts) 0.12	resolver (resolve) –

Table 4 Top qualities, frequencies, and correlation of words for the male and female teachers *(Continued)*

Male Teachers

Frequency of Words		Correlation of Words			
Spanish word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Tareas (homework)	926 (2.2%)	— — (0)	N/A — (0)	—	—
Entender (to understand)	851 (2.1%)	fácil (easy) 0.11	conceptos (concept) 0.1	Darse (notice) 0.1	—
Dispuesto (available)	764 (1.9%)	ayuda (help) 0.25	dudas (doubts) 0.21	resolver (resove) 0.13	apoyar (support) 0.1
Claro (clear)	749 (1.8%)	Queda (remains) 0.2	explicar (to explain) 0.13	Forzosamente (necessarily) 0.12	Rigidez (rigid) 0.1
Problemas (problems)	674 (1.6%)	resolver (to resolve) 0.2	Imaginación (imagination) 0.14	exámenes (exams) 0.12	compañero (peer) 0.11

Female Teachers

Frequency of Words		Correlation of Words			
Spanish Word (English Meaning)	Freq (n = 41,072)	Associated words with Average probability (Spanish/English)			
Entender (understand)	813 (2.0%)	— 0.31	Conceptos (concepts) 0.13	— 0.12	— 0.11
Paciente (patient)	812 (2.0%)	N/A 0.12	—	—	—
		— 0	—	—	—
Exámenes (exams)	781 (1.9%)	parcial (mid-term) 0.19	Confundién (confusing) 0.19	dificultaron (made difficult) 0.12	Señala (points out) 0.12
Falta (absent)	722 (1.8%)	manipulada (manipulated) 0.11	organizacion (organization) 0.11	—	—
Experiencia (experience)	715 (1.7%)	Impartan (impact) 0.11	Probando (testing) 0.11	campo (field) 0.1	—

Author details
[1]Writing Lab, TecLabs, Vicerrectoría de Investigación y Transferencia de Tecnología, Tecnologico de Monterrey, Monterrey, Mexico. [2]School of Engineering and Sciences, Tecnologico de Monterrey, Toluca, Mexico. [3]Emines- School of Industrial Management, Mohammed VI Polytechnic University, Benguerir, Morocco. [4]Institutional Effectiveness Department, North Region ECOA, Tecnologico de Monterrey, Monterrey, Mexico. [5]School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Mexico.

References
Abu Zohair, L. M. (2019). Prediction of Student's performance by modelling small dataset size. *International Journal of Educational Technology in Higher Education*, *16*, 27springer. https://doi.org/10.1186/s41239-019-0160-3.

Alizadeh, M., Mehran, P., Koguchi, I., & Takemura, H. (2019). Evaluating a blended course for Japanese learners of English: Why quality matters. *International Journal of Educational Technology in Higher Education*, *16*, 6springer. https://doi.org/10.1186/s41239-019-0137-2.

Altrabsheh, N. (2016). *Sentiment analysis on students' real-time feedback* PhD Thesis, University of Portsmouth, United Kingdom.

Badri, M., Abdulla, M., Kamali, M., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management*, *20*(1), 43–59. https://doi.org/10.1108/09513540610639585.

Bianchini, S., Lissoni, F., & Pezzoni, M. (2013). Instructor characteristics and students' evaluation of teaching effectiveness: Evidence from an Italian engineering school. *European Journal of Engineering Education*, *38*(1), 38–57. https://doi.org/10.1080/03043797.2012.742868.

Binali, H. H., Wu, C., & Potdar, V. (2009). A new significant area: Emotion detection in e-learning using opinion mining techniques. In *3rd IEEE international conference on digital ecosystems and technologies*, DEST'09, (pp. 259–264). https://doi.org/10.1109/DEST.2009.5276726.

Boex, L. (2000). Attributes of effective economics instructors: An analysis of student evaluations. *The Journal of Economic Education*, 211–227. https://doi.org/10.1080/00220480009596780.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007.

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006.

Boring, A., Ottoboni, K., & Stark, P. B. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*. https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1.

Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., & Wong, F. M. F. (2014). Learning about social learning in moocs: From statistical analysis to generative model. *IEEE Transactions on Learning Technologies*, *7*(4), 346–359. https://doi.org/10.1109/TLT.2014.2337900.

Clark, P. (2015). The Green Paper needs big data. Times Higher Education, retrieved from https://www.timeshighereducation.com/blog/green-paper-needs-big-data

Crues, R. W., Henricks, G. M., Perry, M., Bhat, S., Anderson, C. J., Shaik, N., & Angrave, L. (2018). How do gender, learning goals, and forum participation predict persistence in a computer science MOOC? *ACM Transactions on Computing Education*, *18*(4) article 18, 14. https://doi.org/10.1145/3152892.

Daniel, B. K. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904–920. https://doi.org/10.1111/bjet.12230.

Dommett, E. J., Gardner, B., & van Tilburg, W. (2019). Staff and student views of lecture capture:A qualitative study. *International Journal of Educational Technology in Higher Education*, *16*, 23springer. https://doi.org/10.1186/s41239-019-0153-2.

Drowling, W. (2000) Why we should abolish teaching evaluations. The Daily Targum, Dec 2010.

El-Halees, A. (2011). Mining opinions in user-generated contents to improve course evaluation. *Software Engineering and Computer Systems*Springer, 107–115. https://doi.org/10.1007/978-3-642-22191-0_9.

Engen, B. K. (2019). Understanding social and cultural aspects of teachers' digital competencies. Comprendiendo los aspectos culturales y sociales de las competencias digitales docentes. *Comunicar, 61*, XXVII. https://doi.org/10.3916/C61-2019-01.

Exter, M., Caskurlu, S., & Fernandez, T. (2018). Comparing computing professionals' perceptions of importance of skills and knowledge on the job and coverage in undergraduate experiences. *ACM Transactions on Computing Education, 18*(4) article 21, 29. https://doi.org/10.1145/3218430.

Exter, M.E., Gray, C.M. & Fernandez, T.M. (2019). Conceptions of design by transdisciplinary educators: disciplinary background and pedagogical engagement. International Journal of Technology and Design Education. Springer. https://doi.org/10.1007/s10798-019-09520-w.

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning., 4*(5/6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816.

Gallego-Arrufat, M., Torres-Hernández, N., & Pessoa, T. (2019). Competence of future teachers in the digital security area. Competencia de futuros docentes en el área de seguridad digital. *Comunicar, 61*, XXVII. https://doi.org/10.3916/C61-2019-05.

Gedrimiene, E., Silvola, A., Pursiainen, J., Rusanen, J., & Muukkonen, H. (2019). Learning analytics in education: Literature review and case examples from vocational education. *Scandinavian Journal of Educational Research*, 1–15. https://doi.org/10.1080/00313831.2019.1649718.

Gordillo, A., López-Pernas, S., & Barra, E. (2019). Effectiveness of MOOCs for teachers in safe ICT use training. Efectividad de los MOOC Para docentes en el uso seguro de las TIC. *Comunicar, 61*, XXVII. https://doi.org/10.3916/C61-2019-09.

Heath, J. K., Weissman, G. E., Clancy, C. B., Shou, H., Farrar, J. T., & Dine, C. J. (2019). Assessment of gender-based linguistic differences in physician trainee evaluations of medical faculty using automated text mining. *JAMA Network Open, 2*(5), e193520. https://doi.org/10.1001/jamanetworkopen.2019.3520.

Islahi, F., & Nasreen, N. (2013). Who make effective teachers, men or women? An Indian perspective. *Universal Journal of Educational Research, 1*(4), 285–293. https://doi.org/10.13189/ujer.2013.010402.

Jones, K. M. L. (2019). Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education, 16*, 24springer. https://doi.org/10.1186/s41239-019-0155-0.

Kafedžić, E., Maleč, D., & Nikšić, E. (2018). Differences between male and female secondary school students in assessing their physical and health education teachers ' competences. *Sport Science, 11*(Suppl 1), 64–70.

Kechaou, Z., Ammar, M. B., & Alimi, A. M. (2011). Improving e-learning with sentiment analysis of users' opinions. *Global Engineering Education Conference, 6*, 1032–1038. https://doi.org/10.1109/EDUCON.2011.5773275.

Kim, J. H., Hong, A. J., & Song, H. (2019). The roles of academic engagement and digital readiness in students' achievements in university e-learning environments. *International Journal of Educational Technology in Higher Education, 16*, 21springer. https://doi.org/10.1186/s41239-019-0152-3.

Kori, K., Pedaste, M., & Must, O. (2018). The academic, social, and professional integration profiles of information technology students. *ACM Transactions on Computing Education, 18*(4) article 20, 19. https://doi.org/10.1145/3183343.

Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogybuilding a learning companion. Int. Conference on Advanced Learning Technologies (ICALT), IEEE Computer Society, Aug 2001, Madison, WI, USA. p. 0043. https://doi.org/10.1109/ICALT.2001.943850.

Kumakawa, T. (2017). A text mining examination of university students' learning program posters. *Open Access Library Journal, 4*, e3639. https://doi.org/10.4236/oalib.1103639.

Lau, K., Lee, K., & Ho, Y. (2005). Text Mining for the Hotel Industry. *Cornell Hotel and Restaurant Administration Quarterly, 46*(3), 344–362. https://doi.org/10.1177/2F0010880405275966.

Laube, H., Massoni, K., Sprague, J., & Ferber, A. (2007). The impact of gender on the evaluation of teaching: What we know and what we can do. *NWSA Journal, 19*(3), 87–104 Retrieved March 17, 2020, from www.jstor.org/stable/40071230.

Liao, S. N., Zingaro, D., Thai, K., Alvarado, C., Griswold, W. G., & Porter, L. (2019). A robust machine learning technique to predict low-performing students. *ACM Transactions on Computing Education, 19*(3) article 18, 19. https://doi.org/10.1145/3277569.

Litman, D. J., & Forbes-Riley, K. (2004). Predicting student emotions in computer human tutoring dialogues. In *42nd annual meeting on ass. For computational linguistics*, (pp. 351–358). https://doi.org/10.3115/1218955.1219000.

Mackness, J., Mak, S., & Williams, R. (2010). *The ideals and reality of participating in a mooc*Networked Learning Conference, University of Lancaster, (pp. 266–275).

Mayer-Schönberger, V., & Cukier, K. (2014). *Learning with big data: The future of education*. New York: Houghton Mifflin Harcourt.

Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association, 17*(2), 535–566. https://doi.org/10.1093/jeea/jvx057.

Moshinskie, J. (2001). How to keep e-learners from e-scaping. *Performance Improvement, 40*(6), 30–37. https://doi.org/10.1002/pfi.4140400607.

Munezero, M., Montero, C. S., Mozgovoy, M., & Sutinen, E. (2013). Exploiting sentiment analysis to track emotions in students' learning diaries. In *13th ACM international conference on computing education research*, (pp. 145–152). https://doi.org/10.1145/2526968.2526984.

Munro, M. (2018). The complicity of digital technologies in the marketisation of UK higher education: Exploring the implications of a critical discourse analysis of thirteen national digital teaching and learning strategies. *International Journal of Educational Technology in Higher Education, 15*, 11springer. https://doi.org/10.1186/s41239-018-0093-2.

Nosu, K., & Kurokawa, T. (2006). A multi-modal emotion-diagnosis system to support e-learning. *Innovative Computing, Information and Control, 2*, 274–278. https://doi.org/10.1109/ICICIC.2006.203.

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in facebook and its application to e-learning. *Computers in Human Behavior, 31*, 527–541. https://doi.org/10.1016/j.chb.2013.05.024.

Pandey, S., & Pandey, S. K. (2019). Applying natural language processing capabilities in computerized textual analysis to measure organizational culture. *Organizational Research Methods, 22*(3), 765–797. https://doi.org/10.1177/1094428117745648.

Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systemic literature review of empirical evidence. *Journal of Educational Technology & Society*, *17*(4), 49–64.

Payne, A. (2006). Handbook of CRM. Achieving excellence in customer management. Butterworth-Heinemann, imprint Elsevier, Oxford, UK.

Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. *Education Sector*, UNESCO, ED-2019/WS/8, 46.

Perrotta, C., & Williamson, B. (2018). The social life of learning analytics: Cluster analysis and the 'performance' of algorithmic education. *Learning, Media and Technology*, *43*(1), 3–16. https://doi.org/10.1080/17439884.2016.1182927.

Piedade, M. B., & Santos, M. Y. (2010). Business intelligence in higher education: Enhancing the teaching-learning process with a SRM system. In *5th Iberian conference on information systems and technologies*, (pp. 1–5).

Prinsloo, P. (2017). Fleeing from Frankenstein's monster and meeting Kafka on the way: Algorithmic decision-making in higher education. *E-Learning and Digital Media*, *14*(3), 138–163. https://doi.org/10.1177/2042753017731355.

Rivera, L. A., & Tilcsik, A. (2019). Scaling down inequality: Rating scales, gender bias, and the architecture of evaluation. *American Sociological Review*, *84*(2), 248–274. https://doi.org/10.1177/0003122419833601.

Romero, L., Saucedo, C., Caliusco, M. L., & Gutiérrez, M. (2019). Supporting self-regulated learning and personalization using ePortfolios: A semantic approach based on learning paths. *International Journal of Educational Technology in Higher Education*, *16*, 16, springer. https://doi.org/10.1186/s41239-019-0146-1.

RStudio (2018). *Open source and enterprise-ready professional software for R* Available at: https://www.rstudio.com/ last accessed September 2019.

Sabbe, E., & Aelterman, A. (2007). Gender in teaching: A literature review. *Teachers and Teaching*, *13*(5), 521–538. https://doi.org/10.1080/13540600701561729.

Sánchez, A., Domínguez, C., Blanco, J. M., & Jaime, A. (2019). Incorporating computing professionals' know-how: Differences between assessment by students, academics, and professional experts. *ACM Transactions on Computing Education*, *19*(3) article 26, 18. https://doi.org/10.1145/3309157.

Shen, L., Wang, M., & Shen, R. (2009). Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment. *Educational Technology & Society*, *12*(2), 176–189.

Silva, J., Usart, M., & Lázaro-Cantabrana, J. (2019). Teacher's digital competence among final year pedagogy students in Chile and Uruguay. Competencia digital docente en estudiantes de último año de Pedagogía de Chile y Uruguay. *Comunicar*, *61*, XXVII. https://doi.org/10.3916/C61-2019-03.

Slade, S., & Prinsloo, P. (2013). Learning analytics, ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366.

Song, D., Lin, H., & Yang, Z. (2007). Opinion mining in e-learning system. In *Network and parallel computing workshops, 2007*, (pp. 788–792. IFIP Int. Conference, IEEE). https://doi.org/10.1109/NPC.2007.51.

TEC (2018). *TEC21 Modelo Educativo* Retrieved August 2019, from Tecnologico de Monterrey (ITESM): http://modelotec21.itesm.mx/files/folletomodelotec21.pdf.

Tian, F., Zheng, Q., & Zheng, D. (2010). Mining patterns of e-learner emotion communication in turn level of chinese interactive texts: Experiments and findings. In *Computer supported cooperative work in design*, (pp. 664–670). IEEE. https://doi.org/10.1109/CSCWD.2010.5471892.

Troussas, C., Virvou, M., Espinosa, K. J., Llaguno, K., & Caro, J. (2013). Sentiment analysis of facebook statuses using naive bayes classifier for language learning. *Information, Intelligence, Systems and Applications*, *4*, 1–6. https://doi.org/10.1109/IISA.2013.6623713.

Tur, G., Marín, V. I., & Carpenter, J. (2017). Using twitter in higher education in Spain and the USA. *Comunicar*, *25*(51), 19–27. https://doi.org/10.3916/C51-2017-02.

UNESCO (2015). *Competency based education. Tecnologico de Monterrey, educational innovation observatory*. Mexico: Learning Portal - Planning education for improved learning outcome.

Wang, K., & Zhu, C. (2019). MOOC-based flipped learning in higher education: Students' participation, experience and learning performance. *International Journal of Educational Technology in Higher Education*, *16*, 33springer. https://doi.org/10.1186/s41239-019-0163-0.

Wen, M., Yang, D., & Rosé, C. P. (2014). Sentiment analysis in MOOC discussion forums: What does it tell us? In *7th international conference on educational data mining* EDM 2014.

Weston, T. J., Dubow, W. M., & Kaminsky, A. (2019). Predicting Women's persistence in computer science- and technology-related majors from high school to college. *ACM Transactions on Computing Education*, *20*(1) article 1, 16. https://doi.org/10.1145/3343195.

Whitney, B., Hayter, J., & Marshall, E. C. (2019). Gender bias and temporal effects in standard evaluations of teaching. *AEA Papers and Proceedings*, *109*, 261–265. https://doi.org/10.1257/pandp.20191104.

Williamson, B. (2018). The hidden architecture of higher education: Building a big data infrastructure for the 'smarter university'. *International Journal of Educational Technology in Higher Education*, *15*(1), 12. https://doi.org/10.1186/s41239-018-0094-1.

Yadav, A., & Berges, M. (2019). Computer science pedagogical content knowledge: Characterizing teacher performance. *ACM Transactions on Computing Education*, *19*(3) article 29, 24. https://doi.org/10.1145/3303770.

Publisher's Note