



Towards teaching analytics: a contextual model for analysis of students' evaluation of teaching through text mining and machine learning classification

Kingsley Okoye, et al. [full author details at the end of the article]

Received: 12 July 2021 / Accepted: 19 September 2021 / Published online: 11 October 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Recent trends in *educational technology* have led to emergence of methods such as teaching analytics (TA) in understanding and management of the teaching–learning processes. Didactically, *teaching analytics* is one of the promising and emerging methods within the Education domain that have proved to be useful, towards scholastic ways to make use of substantial pieces of evidence drawn from educational data to improve the teaching–learning processes and quality of performance. For this purpose, this study proposed an educational process and data mining plus machine learning (EPDM + ML) model applied to contextually analyze the teachers' performances and recommendations based on data derived from students' evaluation of teaching (SET). The EPDM + ML model was designed and implemented based on amalgamation of the Text mining and Machine learning technologies that builds on the descriptive decision theory, which studies the rationality behind decisions the learners are disposed to make based on the textual data quantification and statistical analysis. To this effect, the study determines pedagogical factors that influences the students' recommendations for their teachers, what role the sentiment and emotions expressed by the students in the SET play in the way they evaluate the teachers by taking into account the gender of the teachers. This includes how to automatically predict what a student's recommendation for the teachers may be based on information about the students' gender, average sentiment, and emotional valence they have shown in the SET. Practically, we applied the Text mining technique to extract the different sentiments and emotions (intensities of the comments) expressed by the students in the SET, and then utilized the quantified data (average sentiment and emotional valence) to conduct an analysis of covariance and Kruskal Wallis Test to determine the influential factors, as well as, how the students' recommendation for the teachers differ by considering the gender constructs, respectively. While a large proportion of the comments that we analyzed ($n=85,378$) was classified to be neutral and predominantly interpreted to be positive in nature considering the sentiments (76.4%), and emotional valence (88.2%) expressed by the students. The results of our analysis shows that for the students' comments which contain some kind of positive or negative sentiment (23.6%) and emotional valence (11.8%); that females students recommended the teachers taking into account the sentiments

($p=.000$). While the males appear to be slightly borderline in terms of emotions ($p=.056$) and sentiment ($p=.077$). Also, the EPDM+ML model showed to be a good predictor and efficient method in determining what the students' recommendation scores for the teachers would be, going by the high and acceptable values of the precision (1.00), recall (1.00), specificity (1.00), accuracy (1.00), F1-score (1.00) and zero error-rate (0.00) which we validated using the k -fold cross-validation method, with 63.6% of optimal k -values observed. In theory, we note that not only does the proposed method (EPDM+ML) proves to be useful towards effective analysis of SET and its implications within the educational domain. But can be utilized to determine prominent factors that influences the students' evaluation and recommendation of the teachers, as well as helps provide solutions to the ever-increasingly need to advance and support the teaching–learning processes and/or students' learning experiences in a rapidly changing educational environment or ecosystem.

Keywords Teaching analytics · Text mining · Machine learning · Educational innovation · Higher education · Performance assessment

1 Introduction

The application of *teaching analytical* (TA) methods for analysis of the different datasets collected about the educational process and performances, e.g., the students learning experiences and outcome, can enhance the level of impact of several educational initiatives and resultant technologies, both in terms of students' satisfaction and at institutional levels. Studies have shown that many educators depend on the outcomes of the teachers and/or students evaluation of teaching (SET), to explore the effectiveness of the various strategies used to create, manage, or improve the various educational activities that underlie the higher educational institutions' (HEIs) programs and curricula (Badri et al., 2006; Bianchini et al., 2013; Boring, 2017; Holmes et al., 2019; Tondeur et al., 2020). Also, there have been speculations or discourse on the strengths, weaknesses, opportunities, and threats to using information collected about the teaching–learning activities to support the educational processes (Ferguson, 2012; Gedrimiene et al., 2019; Mangaroska & Giannakos, 2019; Papamitsiou & Economides, 2014; Renz & Hilbig, 2020). This ranges from the proliferation of SET-generated data to support and improve the several learning processes (Larrabee S nderlund et al., 2019; Prinsloo et al., 2012; Slade & Prinsloo, 2013), to theoretical approaches that aim to combine the teaching practices with design-based research, and innovative methods that support educators with didactic/technological tools to improve the quality of teaching (Herodotou et al., 2019a, b; Ndukwe & Daniel, 2020).

Nowadays, the need to effectively use information (data) derived from the students' evaluation of teaching or the teacher's performance in addressing the different challenges with educational technologies and data mining have not been overemphasized (Altrabsheh et al., 2014; Badri et al., 2006; Crues et al., 2018; Perrotta & Williamson, 2018; Romero & Ventura, 2020). This, perhaps, has

particularly become important, when the outcomes of the developed methods can be used to adapt, monitor, predict, and recommend best practices for the teaching and learning process within the higher education domains (Abu Zohair, 2019; Aldowah et al., 2019; De Quincey et al., 2019; Nganji, 2018). To note, while Badri et al (2006) stated that SET has become a factor in the renewal and promotion of long-term contracts, merits, and award-related decisions for the teachers in several HEIs. Perrotta and Williamson (2018), on the other hand, notes that how the various educational activities/processes are managed essentially depends on the methods which are applied for the (educational) data collection, including its measurement, analyses, and interpretation within the educational context.

For all intents and purposes, this study believes that there is not only the need for teaching analytical methods that aim to efficiently extract contextual-based insights from the several collections of educational datasets and information recorded in SET. But also, there is the necessity to transform the derived information into actionable plans or practices that can help improve the teaching–learning processes across the different HEIs. In this study, the *text mining* and *machine learning* technique was used to analyze the students' evaluation of teaching and recommendation of the teachers based on the opinions or perceptions (e.g., sentiment analysis, and emotional valence) of the students, as well as determines what the students' recommendation may be based on the identified sentiments and emotional valence scores. Whereas, the Text mining method is held to support the educational processes and/or information management, due to its capability to analyze and derive (new) relevant information from the textual datasets which are recorded in the various databases of the HEIs (Kumakawa, 2017; Lau et al., 2005; Tseng et al., 2018). The Machine learning classification model or technique is applied to make predictions based on the identified students' features or input datasets (Abu Zohair, 2019; Ghosh et al., 2020; Ofli et al., 2016; Viji et al., 2020; Wong & Yeh, 2019). Moreover, the study also believes that by efficiently using the extracted information or results of the two methods (Text mining and Machine learning) correctly, the higher educational institutions or educationalists can define adequate teaching analytical methods or innovative technologies that can not only be used to support the effectiveness of the teaching–learning processes in general, but can also be utilized to maintain a strong relationship amongst the stakeholders (e.g., HEIs, teachers, students) (Ndukwe & Daniel, 2020; Payne, 2006; Piedade & Santos, 2010; Renz & Hilbig, 2020; Tseng et al., 2018). Therefore, in theory, the resultant model (EPDM + ML) and methodological approach used in this study shows to be important towards the promotion and advancement of the stakeholders' experiences and/or teaching–learning performances/outcomes at large (Bowdre, 2020; De Fortuny et al., 2013; Dollinger & Lodge, 2018; Er et al., 2019; Gomes & Ma, 2020; Hilliger et al., 2020; Pedró et al., 2019; Sánchez et al., 2019; Tóth & Surman, 2019; Yadav & Berges, 2019).

1.1 The rationale of the study

The use information or insights drawn from analyzing educational datasets such as SET, can help the Educators in the development of innovative teaching practices

to foster the learning processes. Moreover, achievement of strategic or idiosyncratically analysis of the (educational) datasets, otherwise allied to the notion of “datafied-Education” has shown to be one of the most pertinent challenges that faces the effective delivery of the teaching–learning processes, both in the literature and in practice (Cerratto Pargman & McGrath, 2021; Hilliger et al., 2020; LALA, 2020; Mahmoud et al., 2020; Martens et al., 2020; Ndukwe & Daniel, 2020; Pettersson, 2020; Slade & Prinsloo, 2013). For example, while Cerratto Pargman and McGrath (2021) noted the main gaps with education in the literature and practice to be that the educational data-driven practices are highly context sensitive, not synonymous with evidence-based practices, and are not sustainable per se. The authors (Cerratto Pargman & McGrath, 2021), on the other hand, notes that with the growth in digitalization or datafication of Education, that availability of significant amounts of data, i.e., educational “big” data, have created opportunities for the use of technologies such as AI and/or machine learning techniques to gain valuable insight into how students learn in higher education, as described in this study. To this end, this study strongly believes that useful dissemination of “student-generated data” can be used not only to understand, support, and provide an increased learning process or performance metrics for the stakeholders (teachers and students) particularly in addressing the social-technical or pedagogical challenges related to the teaching–learning process (Dimitriadis et al., 2021; Mahmoud et al., 2020; Raffaghelli et al., 2020; UNESCO, 2014, 2021). But it also demonstrates the importance of the use of data through a predictive analytical approach to inform or shift the current culture of academic advising or teaching paradigms, from one of compliance to one that focuses on students’ learning success and recommendations based on their educational experience (Bowdre, 2020).

Along these lines, this study shows that there is a need for innovative methods or approaches, such as the EPDM+ML model proposed in this paper, for extraction of *educational-based* information from the unprecedented datasets recorded and stored about the students’ evaluation/recommendation of the teaching–learning performances, to help transliterate them into actionable plans for education in general. In our analysis, we extended the educational process and data mining (EPDM) model proposed in Okoye et al., (2020) to show how the amalgamation of the Text mining and Machine learning techniques which we grounded on the descriptive decision theory (Baucells & Katsikopoulos, 2011; Chandler, 2017), can be used to analyze the (educational) data (SET) towards improvement of the end-to-end teachers-students learning process and interactions within the higher educational setting. To this end, we proposed an educational process and data mining + machine learning (EPDM+ML) model for fostering the teaching analytics and performance evaluation.

The main research questions studied in this paper are as follows:

- How can we analyze the educational datasets (SET) within the higher education context to understand the pedagogical and social-technical factors that influences the students’ recommendation for the teachers?
- How do we exploit the resultant information or pieces of evidence (e.g., sentiment and emotional valence) extracted from the SET data to determine

whether the students evaluate and recommend their teachers by taking into account the gender of the teachers or construct?

- What role does the average sentiment and emotional valence expressed by the students play in the way they evaluate the teachers?
- How can we predict what a student's recommendation for the teachers may be based on the gender of the students, average sentiment, and emotional valence scores?

To answer the identified research questions, the study designed a number of constructs it used to conduct the analysis and investigations:

- In the text mining approach, we applied the EPDM model to determine the different sentiment and emotions expressed by the students in the comments for the teachers by considering the gender differences. To this effect, we performed a sentiment and emotional valence analysis (i.e., polarity or textual data quantification) to determine the intensities of the comments provided by the students and their impact on the teachers' evaluation.
- Also, we determined the main factors or differences in the way the students evaluated and/or recommended the teachers based on the quantified data (average sentiment and emotional valence) by holding out the students' gender as one of the potential heightening factors.
- For the machine learning method, we built a textual data classification model to determine to what extent the approach (EPDM + ML) is capable of predicting what a students' recommendation or evaluation (scores) for the teachers would be by taking into account the students' gender, average sentiment, and emotional valence.
- Finally, we provided an empirical discussion of the implications of the study findings based on the significant factors, gender differences or perspectives, and machine learning classifications outcomes.

Consequently, the study makes the following contributions to knowledge, based on the analyzed constructs and effort to provide answers to the research questions:

- (1) It shows the capability of the Text mining approach and Machine learning technologies towards effective (educational) data analysis, and understanding of its main implications/application within the education domain.
- (2) It defines a text mining and machine learning model (EPDM + ML) that makes use of the comments (textual data) provided by the students in SET to determine the impact of the different emotions and sentiments expressed by the students in connection to their recommendation of the teachers based on the gender differences.
- (3) It develops a machine learning classification model that was trained to predict to a significant level (high accuracy), the ratings or scores a student is most likely to give to their teachers by considering the students' gender, sentiments, and emotional valence constructs.

- (4) It demonstrates how data about students' evaluations of teaching (SET) can be utilized to provide solutions to the ever-increasingly need to social-technically and pedagogically advance/support the teaching–learning processes and/or students' experiences in a rapidly changing educational market or ecosystem.
- (5) It introduces a teaching analytical method that shows to be effective towards the understanding/improvement of the end-to-end teaching–learning processes in education, as well as, how compounding factors such as the gender differences affects the way students rate or evaluate the teachers' performances and/or learning outcomes.

2 Background information

2.1 Text mining and machine learning techniques for educational process evaluation

The fundamental feature of the Text mining (Altrabsheh et al., 2014; Binali et al., 2009; El-Halees, 2011; Pandey & Pandey, 2019; Wen et al., 2014) and Machine learning (Abu Alfeilat et al., 2019; Abu Zohair, 2019; Dey et al., 2016; Ghosh et al., 2020; Ofli et al., 2016; Viji et al., 2020) techniques, both in what can be defined in theoretical and technological paradigms; is that both methods can be used to understand the several patterns or relationships that exist in the (educational) datasets stored in the information systems or databases of the several organizations' processes (Jones, 2019; Tur et al., 2017; van der Aalst, 2016). With Text mining, we note that the supported methods can be applied to determine the connections between the real-time processes and the intended users or stakeholders (Wen et al., 2014). Whereas, with Machine learning methods, we note that the technique can be used to predict (e.g., through automatic classification) (De Fortuny et al., 2013; Ofli et al., 2016) the relationships that exist between the process instances in connection to the processes in question. Indeed, the congruence of the aforementioned features of both methods (Text mining and Machine learning) can be harnessed to understand and foster the teaching–learning processes for the said stakeholders (HEIs, teachers, and students) based on the descriptive decision theory or concept (Baucells & Katsikopoulos, 2011; Chandler, 2017) as demonstrated in this paper. Typically, whilst the sentiments/emotions or experiences of the users (e.g., learners) can be determined from data collected about the teachers-students interactions through the text mining technique, the machine learning technique or classification models, on the other hand, can be used to predict what the learners' recommendation or assessment about the teachers-students interactions could be based on the expressed/extracted sentiments or emotional valence (Abu Zohair, 2019; Bollen et al., 2011; Dey et al., 2016; Ghosh et al., 2020).

As an example, Wen et al (2014) mined collective sentiments in a massive open online courses' (MOOCs) forum post, to monitor the students' thoughts towards the several offered courses, and the result shows that sentiment analysis (text mining) can be used to provide an effective method to engage the learners even in social settings (Bollen et al., 2011; Brinton et al., 2014; Crues et al., 2018). The outcome

of the exploratory study comprised of a survival model (Wen et al., 2014) that was built as a predictive/monitoring tool for determining the efficacy of certain human-expressions or language-behaviours (e.g., the impact of students' opinions in the MOOCs environment) based on the probability of certain events happening. Also, taking into account the connectedness between the text mining technique (e.g., sentiment analysis) and machine learning or classification models (Ofli et al., 2016), the study of Dey et al. (2016) notes that the sentiments which are often found in the comments or feedbacks (e.g., SET) can be categorized by polarity (i.e., positive, neutral, or negative Kalaivani, 2013; Litman & Forbes-Riley, 2004; Okoye et al., 2020), and then utilized to provide valuable pointers or indicators in connection to the various reasons or purposes for which the datasets are analyzed (e.g., the advances in teaching analytical methods and/or students' evaluation of teaching described in this study). Besides, the authors (Dey et al., 2016) also used a statistical method that supports the K-nearest neighbour (KNN) (Abu Alfeilat et al., 2019; Ghosh et al., 2020; Viji et al., 2020) and Naïve Bayes' (Zhou et al., 2020) supervised machine learning algorithms to capture the different words/sentence polarities and elements of the subjective styles or patterns.

Along these lines, this study shows that through extraction of the different polarities or intensities of the sentiments and emotional valence expressed by the students in the SET data, that we are able to ascertain and provide new and vital information in relation to the top emotions the undergraduates (students) show when completing the SET instrument and/or rating of the teachers' performance, and then provide strategies for further improvements through the machine learning and statistical analysis procedures. Consequently, we applied the k-nearest neighbour (KNN) classifier/algorithm (Abu Alfeilat et al., 2019; Ghosh et al., 2020; Viji et al., 2020) to model and make predictions about the students' recommendation scores for the teachers based on the quantified (polarized) sentiments and emotional valence.

2.2 Teaching analytics towards educational management and process analysis

Teaching analytics (TA) is a term used to describe emerging methods and technologies that are used to support the educational processes to ensure an increased teaching or scholastic practices/pedagogies within the education domain (Ndukwe & Daniel, 2020; Romero & Ventura, 2020; Wise & Jung, 2019). Regardless of the context in which the TA methods or tools are applied, the main goal of the TA-supported methods is to facilitate some form of knowledge extraction (sense-making analytics) from the readily available educational datasets and then foster a response (actionable analytics) based on the derived information (Wise & Jung, 2019). According to Wise and Jung (2019), embedding the use of educational technologies or learning analytical tools into teaching practices to help inform instructional-based decisions may represent to be a cumbersome and/or time taking process. For example, the authors (Wise & Jung, 2019) noted that the gap between finding/discovering some of the interesting patterns or knowledge from the educational data/processes, to taking actionable responses or decisions is one of the most critical considerations to take into account when aiming to bridge the pedagogical support or complexities

within the higher educational processes. On the other hand, TA is also allied to the concept of “business intelligence” (BI) (Ndukwe & Daniel, 2020; van der Aalst, 2016) in education, as the resultant insights drawn from the methods can be used to not only improve the teaching practices in the various HEIs’ settings, but also utilized to provide meaningful or valuable actions to drive the business operations forward, based the pieces of evidence or insights drawn from analyzing the datasets (e.g., educational data) collected about the processes in question. Moreover, whereas Romero and Ventura (2020) notes the educational data mining (EDM) methods to be closely related to the overlapping terms such as the Teaching Analytics, Big Data in Education, Academic Analytics, Datafied or Data-Driven Education, Institutional Analytics, Educational Data Science, and Data-Driven Decision-Making (Cech et al., 2018; Cerratto Pargman & McGrath, 2021; De Fortuny et al., 2013) in Education within the current works of literature. Ndukwe and Daniel (2020) notes that analysis of educational datasets collected about interactions of the teachers in relation to the students’ learning processes (i.e., teacher-centric learning design) is a promising way of increasing knowledge about the teaching processes, and how it can be effectively sustained. Therefore, improving the effectiveness and efficiency of the HEIs and the teaching practices, including other benefits such as students’ success and active learning engagement, curriculum enhancement and development, etc. (Ndukwe & Daniel, 2020).

2.3 The state-of-the-art in educational data mining and process innovation

Educational data mining (EDM) and the supported methods which are mainly used to enhance the different organizational processes through insights drawn from educational data, have experienced significant growth and attention over the years both in theory and in practice (Abu Zohair, 2019; Alizadeh et al., 2019; Bogarín et al., 2018; Bowdre, 2020; Cerratto Pargman & McGrath, 2021; Dommert et al., 2019; Exter et al., 2018, 2019; Romero & Ventura, 2020; Sánchez-Mena et al., 2019; Wang & Zhu, 2019). A majority of existing studies demonstrates that the educational institutions are gradually becoming “data-hungry”, thus, are increasingly seeking data and the results which are obtained from the applied methods for their own use or educational purposes (Clark, 2015; Romero & Ventura, 2020; Williamson, 2018). For example, Williamson (2018) noted the “Data Future” program as one of the many educational initiatives that will foster new ways of standardizing and quantifying accumulated data and information for education purposes. Also, the several “smart learning tools” and “learning analytical platforms” (Aldowah et al., 2019; Jones, 2019; Larrabee Sønderslund et al., 2019; Perrotta & Williamson, 2018), otherwise allied to the notion of “business intelligence” (Ndukwe & Daniel, 2020; van der Aalst, 2016) are now being used by the different institutions to harness the various educational-process-related decisions and strategies (De Fortuny et al., 2013).

Indeed, in the modern-day educational settings or twenty first Century education; data are now being collected and stored about the teaching and/or learning processes at an unprecedented rate, either for use in understanding the different activities or sub-processes that make up the educational processes, or for use in advancing the

technical capabilities of the educational technologies and impact in the diaspora. Perhaps, such advancements, both in development and use of the educational technologies, have resulted in what the different organizations or educators now call "technology-based education" or yet "educational innovation". As an example, the research by Daniel (2015) notes that *digitalization* of educational processes have led to a significant amount of innovation, and has inadvertently spanned the conception of *datafication* or *datafied-Education* (Cerratto Pargman & McGrath, 2021; Perrotta & Williamson, 2018; Prinsloo, 2017; Prinsloo & Slade, 2017; Slade & Prinsloo, 2013; Williamson, 2018). Nonetheless, the several educational institutions in question have experienced systematic transformations in the various scholastic endeavors, thanks to the new trends in educational data mining and technologies which are purportedly used to foster the way in which educational data are being generated, collected, and analyzed to support the educational processes.

Perhaps, most HEIs have had to amalgamate the digital (educational) technologies in their different activities in order to ensure the quality or reliability of the underlying business models and operations (Kori et al., 2018; Lawrenz et al., 2019; Medne et al., 2020; Mourad, 2017). Moreover, studies within the EDM field have also focused on addressing some of the challenges in using those methods and technologies to provide innovative opportunities for teaching and learning in the different contexts (Abu Zohair, 2019; Alizadeh et al., 2019; Bogarín et al., 2018; Dommett et al., 2019; Exter et al., 2018, 2019; Munro, 2018; Romero & Ventura, 2020; Wang & Zhu, 2019). For instance, Wang and Zhu (2019) observed that the use of digital technologies in teaching is capable of supporting high-quality and transformed educational process by comparing the performances/outcomes of the learners in a MOOC-based, flipped classroom, and traditional class settings, respectively. Whereas, the study by Abu Zohair (2019) that used the machine learning techniques to analyze educational datasets, shows that adequate (accurate) prediction of students' data or analysis may have not only been crucial in improving the students' performance/experiences (Benkwitz et al., 2019; Crues et al., 2018; Kori et al., 2018; Weston et al., 2019), but also represents as a useful tool towards the promotion of the various university's ranking or status (Medne et al., 2020; Mourad, 2017; Tóth & Surman, 2019).

2.4 The future of educational data mining and technologies in HEIs

This study notes the implications of the educational data mining (EDM) in terms of quality of education, innovation, and the future of education (UNESCO, 2014, 2015, 2021). For example, the research by Mayer-Schönberger and Cukier (2014) noted that the acclimatization of *digital technologies* and *education* has paved the way for ground-breaking innovative systems used to drive the higher educational institutions forward. Symptomatically, the resultant models which incorporates modern-day modes of teaching and learning such as Challenge-based learning, memorable university experience, inspiring professors, and flexibility as to how, when, and where learning takes place have become the contemporary goals of the HEIs (TEC, 2018, 2020). Moreover, Dommett et al. (2019) opined that digital technologies for Education

would ostensibly stand as the practical bridge between the several activities that underlies the educational processes and the usefulness of the said educational models. To note, Jones (2019) shows that Educators are consumably adopting the modern educational models to track, aggregate, and analyze the students learning behaviors or profiles that are logged in the databases of the several HEIs (Kori et al., 2018; Medne et al., 2020). Apparently, the most adopted technologies include the educational data mining (Bogarín et al., 2018; Romero & Ventura, 2013, 2020), machine learning (Abu Alfeilat et al., 2019; Abu Zohair, 2019; Dey et al., 2016; Ghosh et al., 2020; Herodotou et al., 2019a, b; Litman & Forbes-Riley, 2004; Muldner et al., 2011; Viji et al., 2020), and learning analytics (Ferguson, 2012; Ferguson & Clow, 2016; Larrabee Sönderlund et al., 2019; Mangaroska & Giannakos, 2019; Ndukwe & Daniel, 2020; Noroozi et al., 2019; Papamitsiou & Economides, 2014, 2019). Interestingly, all of the aforementioned methods and studies have one common goal; which are directed towards achieving an effective and quality of Education, models, and learning outcomes (UNESCO, 2015, 2021).

Likewise, the work in this study believes that the developed educational models and technologies can provide innovative methods and practices within the higher education contexts. Ranging from the development and application of intelligent methodologies that are aimed to transform the students' learning experiences, to empowering the teaching analytical methods and processes, and what could be called the three-dimensional "expressive-communication-relational" pedagogic skills implemented within the higher education settings. To these effects, the study proposed the educational process and data mining plus machine learning model (EPDM+ML) as an extension to the EPDM model proposed in Okoye et al. (2020) that proves to be useful in improving the teaching process and practices in HEIs, as well as, qualities and experiences that the students value in their instructors, e.g., by addressing the gender preconceptions in terms of teachers-students engagement, and the methods' adequacy in supporting the educational process initiatives in the diaspora.

3 Methodology

3.1 Educational process and data mining technologies

Prior studies have looked into how best to apply Educational technologies and models to effectively facilitate the teachers-students learning processes, practices, and experiences (Altrabsheh, 2016; Barton & Dexter, 2020; Bowdre, 2020; Cerratto Pargman & McGrath, 2021; Crues et al., 2018; Mackness et al., 2010; Ndukwe & Daniel, 2020; Renz & Hilbig, 2020; Tondeur et al., 2020; Wen et al., 2014). Ndukwe and Daniel (2020) notes that TA can help Educators (e.g., teachers) to improve the teaching pedagogies and learning outcomes through the provision of tools or platforms that allows them to use data to reflect on teaching. Indeed, the utmost goal of TA-supported methods should be on how to extract or derive meaningful information from the several educational datasets that are stored at an unprecedented rate within the educational databases or information systems; that normally would

not be observable by the ordinary eyes but with ample application of state-of-the-art models and/or methods to help uncover hidden patterns/relationships or knowledge from the readily available datasets (Ndukwe & Daniel, 2020). For example, whereas Tondeur et al (2020) proposed a technological pedagogical content knowledge (TPACK) method to reflect on the role of technology in Education. Ndukwe and Daniel (2020) introduced what could be called a theoretic road map that is directed towards guiding the researchers or educationalists in improving the quality of teaching and/or learning processes by engaging with the educational datasets. The outcome of their approach was a teaching outcome model (TOM) that showed to be useful not only for understanding the state-of-the-art methods or educational technologies related to teaching analytics and its implications for the future of education, prospects or mechanism in higher educational settings, but also in understanding the connection between the conceptual frameworks of teaching analytics (TA) (Ndukwe & Daniel, 2020), Learning Analytics (LA) (Ferguson, 2012; Herodotou, et al., 2019a, b; Jones, 2019; Papamitsiou & Economides, 2014; Renz & Hilbig, 2020; Romero & Ventura, 2020), and learning design (LD) (Holmes et al., 2019; Mangaroska & Giannakos, 2019) in general. Accordingly, this study proposes the Educational Process and Data Mining plus Machine Learning model (EPDM+ML) that is built on conceptual frameworks of the TA, LA, and LD that is based on descriptive decision theory (Baucells & Katsikopoulos, 2011; Chandler, 2017) to provide a data-focused or analytical method that shows to be useful not only towards understanding of the teachers-students learning processes/outcomes to help inform and improve the quality of teaching pedagogies, but also the need for creating teaching-data-literacy or contextual-based analysis to uncover and address the different teaching-learning challenges that can be found within the higher education settings. The method (EPDM+ML) is described in detail in the next section of this paper.

3.2 Educational process and data mining plus machine learning model (EPDM + ML)

The study presents in Fig. 1, the main architecture or building blocks of the EPDM+ML model which it applied for implementation of the method described in this paper.

As shown in Fig. 1, the EPDM+ML model design and implementation is described in two phases. First, we applied the main functions of the EPDM model defined in Okoye et al. (2020) to analyze the comments (textual data) provided by the student in the SET to help deduce and quantify the average sentiments and emotional valence scores of the individual comments. The process also comprised of the cleaning and filtering of the dataset to allow for the text mining and model deployment to follow. The text analysis or textual data quantification is then performed using appropriate text mining tools, packages, and libraries to extract the values that were utilized to comprehend the level of impact of the students' comments or intensity of the sentiment/emotions they expressed across the datasets. Technically, we used the R statistics tool (Rstudio, 2020), an integrated development environment that supports the text mining algorithms and methods such as sentimentr, sentiment Analysis, pander, etc. to analyze the data. Thus, we extracted the sentiment scores

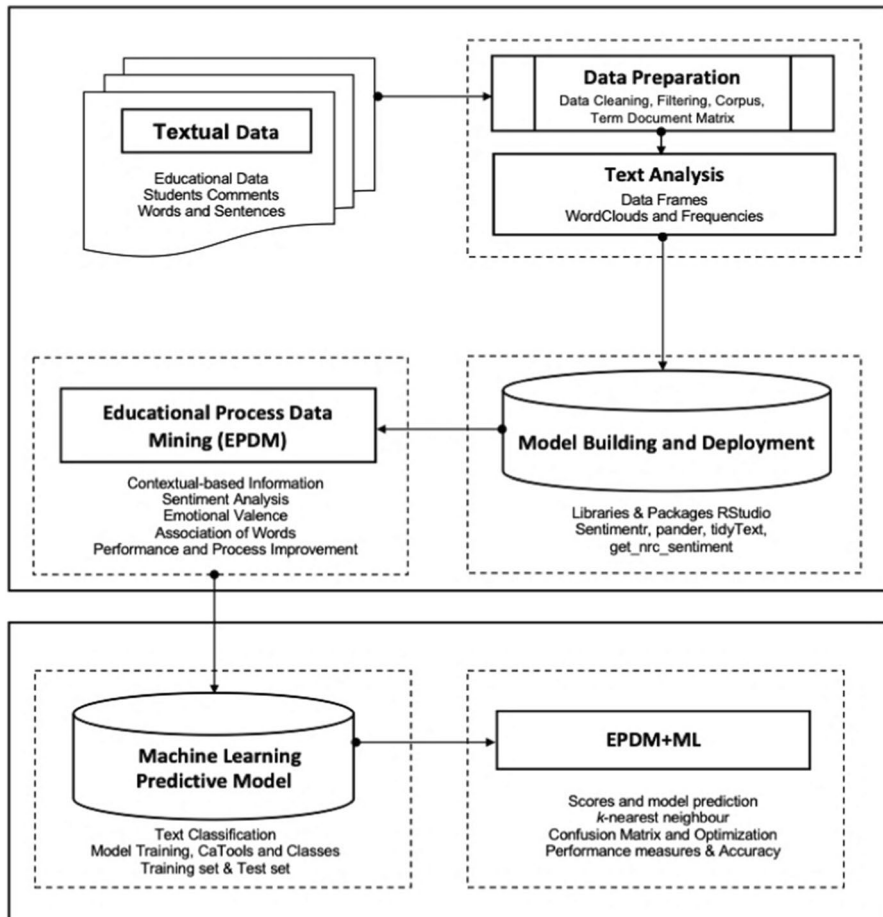


Fig. 1 Educational process and data mining plus machine learning (EPDM + ML) model

and emotional valence of the different comments, and consequently, utilized the quantified results (i.e., *ave_sentiment* and *emotional_valence*) to compare its relatedness to the quantitative measures or data (*recommendation_by_students*) as contained in the provided SET instrument, by considering the influence and significant differences in the way the students recommended the teachers by gender.

For the second phase of the model (Fig. 1), we developed a machine learning classification model that predicts what a students' recommendation for the teachers may be based on the students' gender, *ave_sentiment*, and *emotional_valence* scores. To do this, we trained and analyzed the extracted information (i.e., *recommendation_by_students*, *student_gender*, *ave_sentiment*, and *emotional_valence*) using the *k*-nearest neighbour machine learning algorithm (Abu Alfeilat et al., 2019; Ghosh et al., 2020; Viji et al., 2020; Wong & Yeh, 2019) supported by libraries such as CaTools and Class Library in R (Rstudio, 2020). Also, a cross-validation or

performance evaluation measure using the confusion matrix and performance metrics (van der Aalst, 2016) was used to determine the accuracy, error-rate, specificity, precision, recall, and F1-score of the model being able to classify the predicted scores.

3.3 Data sampling and SET instrument

The study makes use of data we collected from the Student Opinion Survey (ECO, 2013) within the higher education context to conduct the series of experiments and implementation of the EPDM+ML model described in this paper. ECO is a (SET) system designed for the collection of information about students' opinions with respect to the outcome of the different offered courses and teaching programs across the various campuses of the host institution, where this research was conducted. We analyzed the SET data collected through the (ECO) survey from undergraduates about their teachers' performances for the academic year of 2019. The survey instrument was applied across the 26 campuses of the institution spread across the entire national regions of the host country (Mexico); covering around 14 Divisions/Schools, 78 Departments, and 1082 Courses. Therefore, we assume that a wide range of the students' opinions and/or recommendations about the courses and teachers' performance were represented in the collected data. For the purpose of this study, we analyzed both the textual data (comments) provided by the students in response to the question, "Why would you recommend or not recommend the teacher", the gender of the teachers and students, and the quantitative data they provided in response to the question "Please rate your recommendation for the teacher?" which was measured on interval scale of 0 to 10 where 0 means the lowest rating and 10 means the highest rating. It is also important to mention that to get an unbiased and ethical evaluation/analysis of the available data; the names of the students who completed the survey were withheld from the data for anonymity purposes, even though, their gender distribution or demographic information were disclosed. The questionnaires were completed by the students at the end of their respective programs or courses. Also, considering the privacy and ethical point of view; the students who provided the comments were informed about the purpose of the applied questionnaires, and were not directly involved in the analysis performed in this study.

Considering the validity and reliability of the SET data; we note that the ECO instrument is an institutional survey administered and maintained by the host university, and has been used for several years by the institution for the purpose of evaluation/assessment of the teachers' performance based on answers or comments provided by the students. The instrument has been used and validated in previous studies (Hernández, 2013; Montemayor-Gallegos, 2002; Salinas & Martínez, 2018). While, the comments given by the students were a free choice open-ended question, the recommendation of the teachers was a close-ended interval scale question between 1 and 10. The estimated minimum sample size for the study was 40 participants which we considered to be the scientifically acceptable sample size ($n > 30$ or

40) (Roscoe, 1975) for conducting the different experimentations and analysis in this study when compared to the large enough sample size ($n=85,378$) we have used.

Describing the data sample and size, the dataset we used for the study includes a total sample of $n=85,378$ responses we analyzed after cleaning and filtering out the incomplete datasets and students who did not comment in the data. We noted a sample size of $n_1=45,294$ for the male students, and $n_2=40,084$ for female students which we utilized throughout the series of experiments and analyses in this paper. Also, for the training and testing of the proposed machine learning model, we have randomly selected a total sample of $n_3=1000$ ($n=700$ used as training set, and $n=300$ for test set) to evaluate the classification process or performance metrics of the model.

3.4 Experimental setup

The EPDM + ML model implementation and analysis was carried out to:

- Determine the average sentiment and emotional valence of the individual comments given by the students in the SET.
- Determine the marginal mean differences and/or influence that the average sentiment and emotional valence scores has on the students' recommendation of the teachers, and how the results differ by considering the gender differences.
- Determine the extent or capability of the machine learning classification model being able to predict what the students' recommendation for the teachers would be by considering the students' gender, sentiment, and emotional valence displayed in the data or comments.

4 Main methods used for the data analysis

- Text mining or Sentiment analysis: Used to extract the intensities (polarity) of the comments provided by the students in the SET through polarization or textual data quantification.
- Analysis of covariance (ANCOVA) and Kruskal Wallis Test: Used to determine the marginal means of effects that the extracted sentiments and emotional valence has on the students' recommendation for the teachers, and how the results differ by considering the gender construct.
- *K*-nearest neighbor (KNN): Classification algorithm or predictive model used to predict what the students' recommendation for the teachers may be by considering the students' gender, sentiment, and emotional valence scores.

4.1 Data analysis, implementation, and results

For the data analysis and implementation of the various phases of the EPDM + ML model as defined in Fig. 1; the study used the text mining method (EPDM) to extract the different values or scores representing the polarity (intensities) of the average

sentiment and emotional valence expressed by the students in the comment (textual data) using the relevant packages, functions, automation, and quantification methods in R. It is noteworthy to mention that the text (comments) provided by the students were mainly in Spanish and was analyzed in its original form. However, we report the results and outcomes of the experiments in English to cover the wider spectrum of international audience/readers and educational objectives of this study.

4.2 Text mining method and analysis

The study implemented the first phase of the EPDM + ML model by determining the intensities of the students' comments towards the teachers as contained in the ECOA SET by using the EPDM method, a Text Mining technique previously proposed in Okoye et al (2020). Typically, as shown in Tables 1, 2 and 3, the outcome of

Table 1 Word count and average sentiment scores for the students' comments (SET) broken down by gender

| Gender | Element id | Word count | ave_sentm |
|--|------------|------------|-----------|
| Male students ($n = 45,294$) | 1: | 32 | 0.000 |
| | 2: | 2 | 0.000 |
| | 3: | 2 | 0.000 |
| | 4: | 12 | 0.216 |
| | 5: | 3 | 0.000 |
| | ... | ... | ... |
| | 45290: | 3 | 0.000 |
| | 45291: | 4 | 0.000 |
| | 45292: | 24 | 0.000 |
| | 45293: | 3 | 0.000 |
| | 45294: | 3 | 0.000 |
| | | | |
| | Element id | Word count | ave_sentm |
| Female students ($n = 40,084$) | 1: | 12 | 0.000 |
| | 2: | 6 | 0.000 |
| | 3: | 7 | 0.000 |
| | 4: | 19 | 0.000 |
| | 5: | 13 | 0.000 |
| | ... | ... | ... |
| | 40080: | 29 | 0.000 |
| | 40081: | 13 | 0.000 |
| | 40082: | 21 | − 0.054 |
| | 40083: | 4 | 0.000 |
| | 40084: | 3 | 0.000 |

Note: values represent the first and last five comments in the SET data

Table 2 Summary of the sentiment scores across the SET data considering the gender of the students, broken down by word count, standard deviation, and the sentiment score

| | Male students | | | Female students | | |
|--------|---------------|------|-----------|-----------------|------|-----------|
| | Word count | sd | ave_sentm | Word count | sd | Ave_sentm |
| Min | 0 | 0.00 | − 0.716 | 0 | 0.00 | − 0.707 |
| Median | 10 | 0.00 | 0.00 | 12 | 0.00 | 0.00 |
| Mean | 13.8 | 0.04 | − 0.006 | 16.2 | 0.04 | − 0.005 |
| Max | 407 | 1.14 | 1.395 | 419 | 0.65 | 1.223 |

element_id denotes the individual comments by the students; word_count represents the number of words in each comment; sd represents standard deviation of the sentiment scores in the comments; ave_sentm is the average sentiment score for the individual comments

Table 3 Fragment of the emotional valence scores expressed by the students towards the teachers analyzed by gender

| Emotional valence scores | | |
|--------------------------|-------------------------------------|-------------------------------------|
| Comments | (Male students > > > teachers) | (Female students > > > teachers) |
| [1] | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 |
| [16] | 0 0 − 1 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 0 0 0 0 − 1 0 0 0 0 0 |
| [31] | 0 0 1 0 0 0 0 − 1 − 1 0 − 1 0 0 0 0 | 0 0 0 0 0 0 0 0 0 0 − 3 − 1 0 0 0 0 |
| [46] | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 0 0 0 0 0 − 1 0 0 0 0 0 0 1 0 |
| [61] | 0 − 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 0 − 1 0 0 0 0 0 0 0 0 0 0 0 0 0 |
| [76] | 4 0 0 0 0 0 0 0 1 0 0 0 2 0 0 | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 |
| [91] | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 0 0 − 1 0 0 0 0 0 − 2 0 0 0 0 − 1 |
| [106] | 0 0 0 0 − 1 0 0 0 0 0 0 0 0 1 − 1 | 0 0 − 1 0 1 0 − 1 0 0 0 0 0 0 0 0 |
| [---] | --- | --- |
| Min | − 4 | − 4 |
| Median | 0.000 | 0.000 |
| Mean | − 0.012 | − 0.021 |
| Max | 6 | 13 |

the method (EPDM), is a quantified or polarized values that are used to denote the intensities of the different comments provided by the students by using the positive (+), neutral (0), and negative (-) connotations to represent the scores (Litman & Forbes-Riley, 2004; Okoye et al., 2020); whereby the values with positive (sentiment and/or emotional valence) (+) scores represent an attractive sentiment/emotion, whilst the negative (-) scores signify an aversive sentiment/emotion. The zeros represent sentiment/emotions that are classified as neutral (0), and thus, have no emotions or sentiment attached to them. It is also important to mention that by sentiment scores; we refer to the average or impact of the different individual

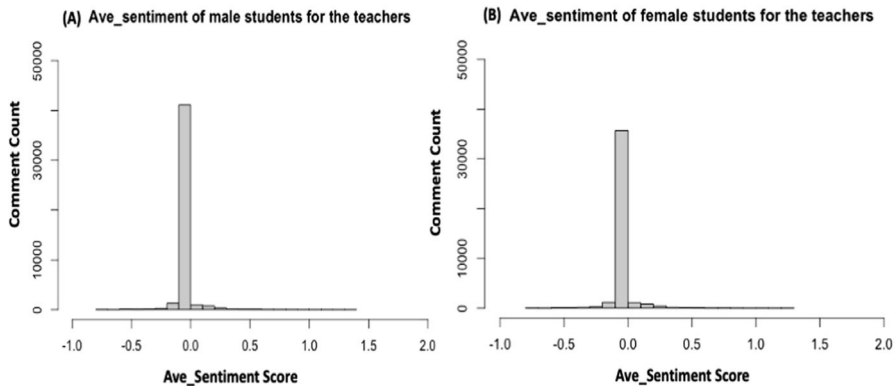


Fig. 2 Comments (counts) vs. Ave_Sentiment score broken down by students gender

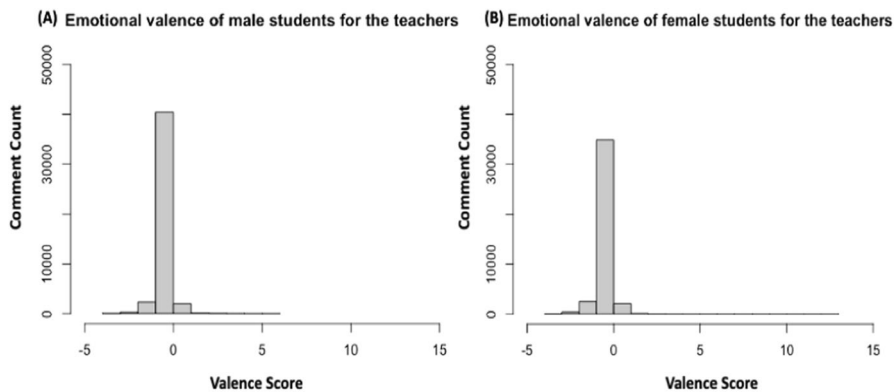


Fig. 3 Emotional valence scores for the SET comments broken down by students gender

comments provided by the students. Whereas the emotional valence is obtained by summing up the scores of the words which the model has identified as a term that can be used to express an emotion in the texts (comments).

As shown in Table 1; the sentiment analysis we did to extract the scores for the different comments were analyzed based on the students' gender distribution. The results are as shown in Table 1 and summarized in Table 2 and Fig. 2.

As shown in Tables 1 and 2, and Fig. 3; we made use of the `sentiment_by`, `get_sentiment`, and `get_nrc_sentiment` functions in R (Rstudio, 2020) to establish the different word counts and the average_sentiment scores for each comment in the data including the standard deviations. The sentiment scores were represented as interval values between -1 and 2 (Fig. 2), denoting the levels of intensity or impact of

each individual comments that was analyzed by the model. For example, in Table 1, we made use of the first and last five comments in the analyzed data to explain the results of the method; whereby the comments that came out with a positive (+) interval value signifies a positive (good) sentiment, whereas comments with a negative (−) value represent an aversive (bad) one. In total, we found that the male students expressed the utmost negative sentiments (comment) with a minimum (min) value of −0.716 when compared to the min (−0.707) expressed by the females (Table 2). Nevertheless, the male students also expressed the most positive sentiments with a maximum (max) value of 1.395 in contrast to the female counterparts who expressed a max average of 1.223, vice and versa.

Furthermore, to quantify the levels of emotional valence of the comments provided by the students by gender, we applied the `get_nrc_sentiment` function which is supported by the `pander` method or algorithm in R to extract the different scores for each comment provided by the students based on the gender differences. Technically, the `get_nrc_sentiment` functions by obtaining and quantitatively labeling the intensities of the words which can be used to express an emotion in the texts by using the positive (++), neutral (0) and negative (−) values (Litman & Forbes-Riley, 2004; Okoye et al., 2020) to represent each relevant word it finds in each case. The results of the method is as shown in Table 3 and Fig. 3.

In Table 3, we showed the emotional valence scores of the first 120 comments. The *Comments* column, [1] to [106], represents the id of the first individual comment in each case within each row. The comments with positive (valence) scores represent an attractive emotion, whilst the negative scores signify an aversive valence. The zeros represent comment which are classified as neutral, and thus, with no emotions attached or words which can be used to express emotions were not found. We note, as shown in Table 3 and Fig. 3, that the valence scores for the male students ranged between −4 and 6, whereas those found for the females ranged between −4 to 13. As example, we show the specific comments for the *min* and *max* emotional valence that we found for the different genders including an example of the neutral comments, as follows:

Male student comment [15043] Positive (Max) Valence Score: 6 > “Excellent professor. He [teacher] not only shows his mastery but also his love for literature in class. Very fun class, challenging due to the time management needed to complete the assignments, but fun and relaxing either way.”

Female student comment [7758] Positive (Max) Valence Score:13 > “I’d recommend her [teacher] because her classes were full of “hands-on” activities, examples and interesting information, and she explain very clearly theoretical information, and didn’t asked us to memorize but to learn. I liked that her quizzes were short but continuous, and just about relevant information. Seriously, I discovered the huge passion I have for art, and she inspire me to get deeper and learn by myself more about the background of art and artists; this changed my perception of art, from just appreciating it to analyzing and feeling it. As well, I felt the confidence to approach

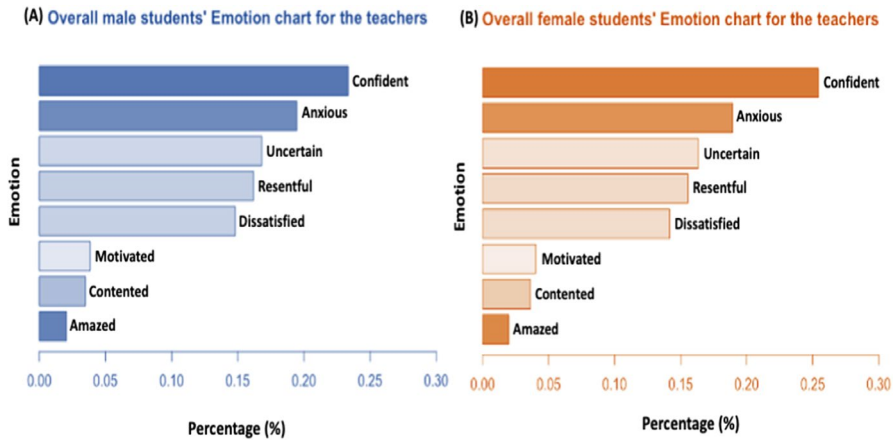


Fig. 4 Overall Emotions expressed by the students for the teachers broken down by gender of the students

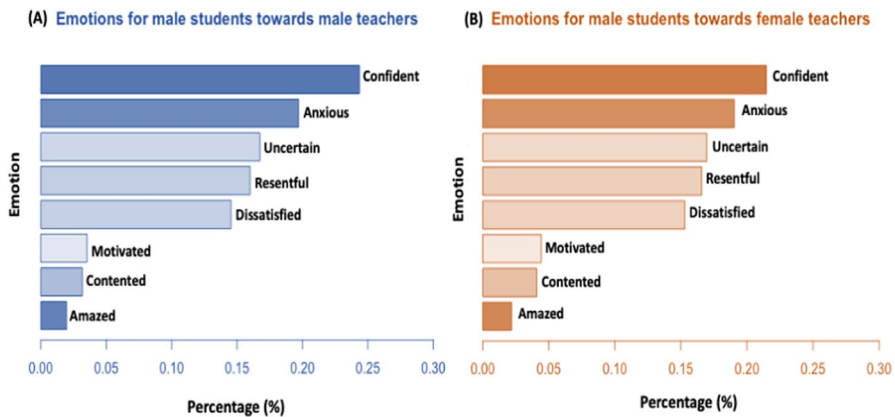


Fig. 5 Emotions expressed by the male students for the male vs female teachers

and ask for help when I needed; once I received a word of support from her, that helped me a lot”.

Male student comment [43944] Negative (Min) Valence Score: $-4 >$ “He is a teacher who has a great knowledge of the subject, however I think that he does not manage to transmit that knowledge to the students at all and that makes one lose interest in the subject and only take the subject simply by passing it even if one does not understand the subject. everything seen in class”.

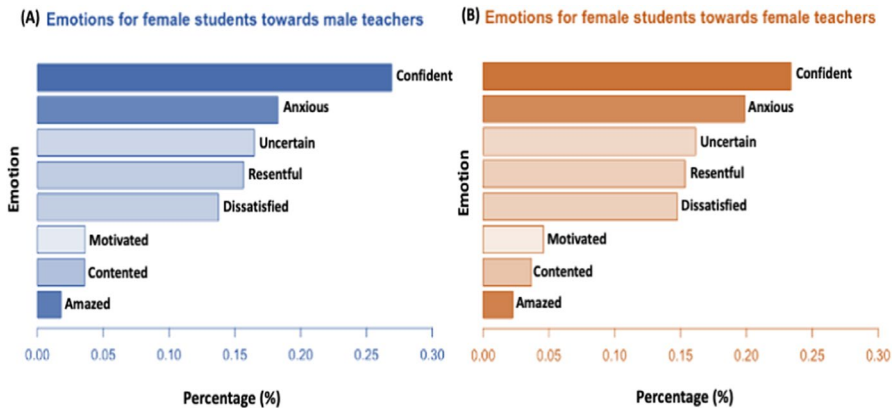


Fig. 6 Emotions expressed by the female students for the male vs female teachers

Table 4 Summary of the Emotional Valence expressed by the students towards the teachers broken down by gender

| EV | Male teachers | | | | Female teachers | | | |
|-----------------|---------------|--------|---------|-----|-----------------|--------|---------|-----|
| | Min | Median | Mean | Max | Min | Median | Mean | Max |
| Male students | − 4 | 0.00 | − 0.008 | 6 | − 3 | 0.00 | − 0.019 | 6 |
| Female students | − 3 | 0.00 | − 0.017 | 10 | − 4 | 0.00 | − 0.027 | 13 |

EV emotional valence, *min* minimum, *max* maximum

Female student comment [38731] *Negative* (Min) Valence Score: $-4 >$ “Excellent teacher, however sadly the course has lost my interest since it is not a dynamic class, or with additional material to that offered by the book”.

Male student comment [1] (Neutral) Valence Score: $0 \rightarrow$ “Because he knows a lot about the subject, they are things that will serve us in the future, in addition, his way of teaching is very specific which makes the class interesting”.

Female student comment [2] (Neutral) Valence Score: $0 \rightarrow$ “She explains very well, has a lot of patience”.

In Figs. 4, 5 and 6 and Table 4, we report the overall emotions of the students about the teachers’ broken down by gender of the teachers and students. We looked at both genders (teachers and students), and the emotions that the students deem as crucial in the SET evaluations. Principally, we applied the sentiment/emotions classifications within the educational domain as noted in Litman and Forbes-Riley, (2004) and Okoye et al. (2020) in the results.

Table 5 Analysis of Covariance for the Recommendation of professors by the students

Analysis of Covariance (ANCOVA)

DV = Recommendation of teachers, IV = Students_Gender (Ave_sentiment*Emotional_valence), $n = 85,378$

| | Mean square | F | p-value (Sig.) | Partial Eta Sq |
|--|-------------|--------|----------------|----------------|
| Student_gender | 14.876 | 3.908 | .048** | .000 |
| Ave_sentiment | 210.602 | 55.324 | .000** | .001 |
| Emotional_valence | 2.983 | 0.784 | .376 | .000 |
| Student_gender*Ave_sentiment*Emotional_valence | 27.133 | 7.127 | .001** | .000 |
| Marginal mean effect (ANCOVA) | 14.892 | 3.912 | .048** | .000 |

The factors marked with asterisk (*) were the found elements or considerations that came out statistically significant in the ANCOVA test

DV dependent variable, IV independent variable

Significance levels $p \leq .05$

Table 6 Kruskal–Wallis test for recommendation of the teachers broken down by students' genderKruskal–Wallis test (Teachers' Recommendation vs Students gender) IV = Ave_sentiment*Emotional_valence, $n = 85,378$

| SET Data | Factor | Mean | Std. Dev | H (X^2) | p-value (Sig.) |
|----------------------------------|-------------------|---------|----------|-------------|----------------|
| Male students ($n = 45,294$) | Ave_sentiment | – 0.006 | 0.074 | 15.238 | .124 |
| | Emotional_valence | – 0.010 | 0.394 | 14.190 | .164 |
| Female students ($n = 40,084$) | Ave_sentiment | – 0.005 | 0.076 | 35.389 | .000* |
| | Emotional_valence | – 0.020 | 0.441 | 5.774 | .834 |

IV independent variable

Significance levels $p \leq .05$

5 Analysis of covariance: gender differences and effects in recommendation

Having established the polarity (intensities or textual data quantification) of the different comments provided by the students in the SET; the study turned its attention to determine the effect that the students' gender has on the recommendation of the teachers based on the linearity of the independent pairwise comparisons (i.e., the estimated marginal means) by controlling the *ave_sentiment* and *emotional_valence* expressed by the students using an Analysis of Covariance test (ANCOVA) (Alao et al., 2019; Wang et al., 2019). In other words, we determined the effect that the students' gender have on the recommendation of the professors whilst also controlling (taking into account) the impact or influence of the covariates (students' average_sentiment and emotional_valence) on the

Table 7 Kruskal Wallis test for recommendation of teachers analyzed by students' and teachers' gender

Kruskal Wallis test for Recommendation of the teachers (DV), analyzed by students' gender vs gender of the teachers

IV = Ave_sentiment*Emotional_valence, $n = 85,378$

| | | Factor | Mean | Std. Dev | H (X^2) | p-value |
|-------------------------------------|-------------------------------------|-------------------|---------|----------|-------------|---------|
| Male students ($n = 45,294$) | Male teachers ($n = 27,855$) | Ave_sentiment | – 0.006 | 0.069 | 8.644 | .566 |
| | | Emotional_valence | –0.010 | 0.402 | 17.953 | .056 |
| | Female teachers ($n = 17,439$) | Ave_sentiment | – 0.007 | 0.080 | 16.889 | .077 |
| | | Emotional_valence | – 0.020 | 0.381 | 9.661 | .471 |
| Female students ($n = 40,084$) | Male teachers ($n = 22,015$) | Ave_sentiment | – 0.004 | 0.070 | 23.425 | .009* |
| | | Emotional_valence | – 0.020 | 0.446 | 8.594 | .571 |
| | Female teachers ($n = 18,069$) | Ave_sentiment | – 0.007 | 0.083 | 29.717 | .001* |
| | | Emotional_valence | – 0.030 | 0.435 | 11.151 | .346 |

DV dependent variable, IV independent variable

Significance levels $p \leq .05$

resultant outcomes. We believed that the covariates might explain some of the differences in the marginal means in terms of the recommendations given to the teachers by the students considering the students' gender. For instance, whether the varying negative/positive sentiments and emotional valence as described in the earlier section (text mining) may lead to a higher or lower recommendations by the students taking into account the gender differences of the students. Henceforth, we analyzed the contributing effect that the students' average sentiment and emotional valence may have on the outcome of the recommendations using the ANCOVA method, and then, conducted a Kruskal Wallis test (Elliott & Hynan, 2011; Frey, 2018) to determine where the significance differences may lie by considering the different genders. To do this, we assume that there is homogeneity between the covariates and the students' gender. The results of the ANCOVA and Kruskal–Wallis tests are as reported in Tables 5, 6, and 7.

As gathered in Table 5, the marginal mean of effect or result of the ANCOVA test shows that there is a significant difference or effect between the students' gender and the recommendation of the teachers ($p = .048$). This means that the students' gender plays a part in the recommendation scores given to the teachers, and also varies by gender. Also, when we took into consideration the influence (covariance) that the ave_sentiment and emotional_valence of the students played in the recommendation of the teachers; we found that whereas the ave_sentiment ($p = .000$) expressed by the students contributed to the test outcome, the students emotional_valence ($p = .376$) do not influence their recommendation of the teachers. Nevertheless, when analyzing the effect of the controlled independent variables (Ave_sentiment*Emotional_valence) while taking into account the influence of the uncontrolled independent variable

(Student_Gender) i.e., Student_Gender*Ave_sentiment*Emotional_valence (see: Table 5), we found that the overall mean effect of all the combined factors were significant ($p=.001$).

6 Kruskal–Wallis test for differences in the students' vs teachers' gender

Having found that there is a significant difference in the students' gender and the average sentiment of the students when they recommend the teachers; we deemed it necessary to conduct a Kruskal–Wallis test to help determine where the significant differences may lie between the genders. The result of the method is as reported in Table 6.

As reported in Table 6, we found that the significant differences in terms of the recommendation of the teachers when analyzed by the students' gender, alongside the impact of the ave_sentiment and emotional_valence of the students when doing so; is observed for the female students ($p=.000$). This means that the female students significantly take into account the sentiment when recommending the teachers, whereas their male counterparts do not ($p=.124$). In any case, both genders (male $p=.164$, female $p=.834$) do not consider their expressed emotions when recommending the teachers.

Finally, we checked to see if the significant results as presented in Table 6 differ by considering the gender of the teachers. The result of the method is presented in Table 7.

In Table 7, we found that the marginal means effect of the average sentiment for female students, as explained in Table 6, was significant for both genders of the teachers (male teachers $p=.009$, female teachers $p=.001$). Moreover, another interesting finding is the fact that the male students appeared to be borderline in terms of the sentiment and emotions when recommending the teachers; with the male students attaching emotional_valence of $p=.056$ for the male teachers, and ave_sentiment of $p=.077$ for the female teachers, respectively.

7 The EPDM + ML machine learning classification/predictive model

To implement the second component of the EPDM+ML approach; we developed a machine learning classification model that predicts what a students' recommendation for the teachers would be based on the students' gender, average sentiment, and emotional valence parameters. As defined in the provided steps or procedures in Algorithm 1, we trained the model with the students' gender (Gd), recommendation scores given by the students in the SET (Rec), and the extracted average sentiments (ave_sentm), and emotional valence scores (EV) (see: Tables 1 and 2), using the k -nearest neighbour (KNN) algorithm (Abu Alfeilat et al., 2019; Ghosh et al., 2020; Viji et al., 2020; Wong & Yeh, 2019) in R statistics (Rstudio, 2020).

Algorithm 1: EPDM+ML textual data classification and predictive model

```

1: Input:   ED, educational data (SET)
           ML, machine learning dataset
           Rec, students' recommendations
           Gd, gender of students
           df.TR, dataframe for training set
           df.TS, dataframe for test set

2: Output: Ave_sentm, average sentiment of the individual students' comments
           EV, emotional valence score of the individual students' comments
           KNN, confusionMatrix, precision, recall, specificity, accuracy, error-rate, F1-score

3: Procedure: EPDM+ML classifier: predictive model and textual data classification

4: Begin

5:   For all extracted dataset, ED

6:     Apply EPDM functions using sentiment_by, get_sentiments, get_nrc_sentiment

7:     Extract average sentiment, Ave_sentm, and emotional valence, EV ← from ED

8:     while no more constructs or considerations (A) is left do

9:       Create ML, by concatenating c(Rec, Gd, Ave_sentm, EV) ← from ED and line 7

10:    Normalize object ML, ← function(x) = return((x - min(x)) / (max(x) - min(x)))

11:    Set seed, df.TR and df.TS, as dataframes ← from ML

12:    Analyze df.TR and df.TS to produce the EPDM+ML predicted scores and classifier results

13:    If df.TR and df.TS interpretation ← Null then

14:      obtain the relevant A from ED and loop to line 9

15:    Else If df.TR and df.TS interpretation ← 1 then

16:      run the knn() method and analyze the outputs, KNN

17: Return: outputs, KNN

18: End If statements

19: End while

20: End For

21: End

```

As gathered in the algorithm (Algorithm 1), the EPDM + ML model which can be applied to analyze any given educational data, especially as it concerns the students-generated dataset (SET), functions as follows; First, the captured dataset (ED) is imported into the integrated development environment, and then the EPDM method/

Table 8 Confusion matrix (performance metrics) for optimal k -value, $knn=31$

| Actual scores | Predicted scores | | | | | | | | | | |
|---------------|------------------|---|---|---|---|----|---|----|----|----|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 0 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 2 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 9 | 2 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 15 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 28 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 165 |

Accuracy = 1.00 (100%), 95% CI = (0.98, 1), $p < .00$

functions (Line 6) (Okoye et al., 2020) are applied to the (textual) data, in this case, the comments provided by the students, to extract the average sentiment (`ave_sentm`) and emotional valence (EV) scores (Line 7). The next steps defined in Lines 9 to 11, involves the creation of the training (`df.TR`) and test (`df.TS`) dataset which we used for the model predictions and classification process by concatenating the considered variables which are, consequently, stored as an object we called ML, i.e., `c(Rec, Gd, Ave_sentm, EV)` (see: Line 9). In Lines 12 to 16, the defined dataframe or objects (`df.TR` and `df.TS`) containing the values of the *training* and *test* sets, respectively, was analyzed by using the k -nearest neighbour method, `knn()`, and in turn, the outputs (ConfusionMatrix, precision, recall, specificity, accuracy, error-rate, F1-score) are returned in Line 17. It is noteworthy to mention that the (ML) input dataset ($n=1000$) was randomly selected from the data sample which perhaps contained the student genders and the recommendation variables; with the training set (`df.TR`) consisting of $n=700$ cases (i.e., 70% of the input dataset) and the remainder utilized as test set (`df.TS`) ($n=300$). For the experimentation, we set the value of k (k -value) to be $k=\sqrt{n}$ (Cover, 1968) where $n=1000$. The results of the predictions or classifications by the model/outcomes are as represented in the Confusion matrix (Ariza-López et al., 2019; van der Aalst, 2016) (see: Table 8), and Tables 9 and 10, respectively. In Table 9, we reported the details of the Optimal value of k (i.e., `knn31`) based on the closest $k=\sqrt{n}$ in order to illustrate how the model predicts each of the recommendation scores by the students in each run test of the model. As shown in the confusion matrix or performance metrics table (Table 8) and Table 9, the recommendation scores by the students was denoted from 0 to 10, with 0 being the lowest and 10 being the highest rating. Furthermore, in Table 10, we conducted a k -fold cross-validation method (Dehghani et al., 2019; Wong & Yeh, 2019; Xiong et al., 2020) to determine the performance of each of the input run test by the `knn` model. To do this, we determined the closest k -values for the model by establishing the square root of the input dataset (i.e., $k=\sqrt{n}$, where $n=1000$),

Table 9 Optimal KNN model (knn.31) classifier and statistics for the different classes of recommendation (scores) by the students

knn.31 model classifier statistics and performance metrics for the optimal $k = \text{sqrt}(n)$, $n = 1000$
 $p = .000$ (2.2e-16), Cohen's Kappa Coefficient: 1.00, CI: 95% (0.98, 1)

| Performance criteria | Score0 | Score1 | Score2 | Score3 | Score4 | Score5 | Score6 | Score7 | Score8 | Score9 | Score10 | Ave_met (knn.31) |
|------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|---------------------|
| Recall (sensitivity) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Specificity | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Precision (pos. pred. value) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Negative pred. value | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Prevalence of score | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 | 0.05 | 0.03 | 0.05 | 0.08 | 0.09 | 0.55 | 0.08 |
| Detection rate | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 | 0.05 | 0.03 | 0.05 | 0.08 | 0.09 | 0.55 | 0.08 |
| Detection prevalence | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 | 0.05 | 0.03 | 0.05 | 0.08 | 0.09 | 0.55 | 0.08 |
| Balanced accuracy | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Values are represented in percentage (%) where 0.00 equals 0% and 1.00 equals 100%

Table 10 K-fold cross-validation of the KNN models used for predicting the recommendation scores by the students

K-fold Cross validation method and results of the KNN models classifications and performance metrics

Where k -values are represented using the closest $k = \text{sqrt}(n)$, i.e., k_2, \dots, k_{36} , $n = 1000$

| Performance criteria | knn.26 | knn.27 | knn.28 | knn.29 | knn.30 | knn.31 | knn.32 | knn.33 | knn.34 | knn.35 | knn.36 |
|----------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|--------------|--------------|--------------|
| Recall (sensitivity) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 0.94 | 0.88 | 0.88 |
| Specificity | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Precision (positive pred. value) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.95 | 1.00 | 1.00 | 1.00 |
| Negative pred. value | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 | 0.99 |
| Prevalence of score | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 |
| Detection rate | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 |
| Detection prevalence | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 |
| Balanced accuracy | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 0.97 | 0.94 | 0.94 |
| p -value: $\leq .05$ | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| Cohen's Kappa (expected at 0.9) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 0.95 | 0.91 | 0.91 |
| Conf. Int (CI): 95% | (0.98, 1) | (0.98, 1) | (0.98, 1) | (0.98, 1) | (0.98, 1) | (0.98, 1) | (0.98, 1) | (0.96, 0.99) | (0.94, 0.98) | (0.91, 0.96) | (0.91, 0.96) |

Values are represented in percentage (%) where 0.00 equals 0% and 1.00 equals 100%

and then executed in each run test for the predictions, the values of the closet square root of n for k , i.e., k_{26} to k_{36} as shown in Table 10.

Accordingly, we utilized the results of the classification process as shown in the confusion matrix (Table 8), and Tables 9 and 10 to calculate the precision, recall, specificity, accuracy, F1-score, and error-rate by the model, respectively, whereby the cross-validation or performance metrics are defined as follows (van der Aalst, 2016):

- TP—number of *true positives*; representing the instances of the scores that were correctly classified as positive
- TN—number of *true negatives*; representing instances of the scores that were correctly classified as negative
- FP—number of *false positives*; representing instances of the scores that are predicted to be positive but should have been classified as negative
- FN—number of *false negatives*; representing instances of the scores that were predicted to be negative but should have been classified as positive.

8 Performance metrics: precision, recall, specificity, accuracy, error-rate, and F1-score

The closest k -values for the ML input dataset ($n=1000$) was used for calculating the performance of the outputed (scores) using a k -fold cross validation method (Dehghani et al., 2019; Wong & Yeh, 2019; Xiong et al., 2020). This was done in order to assess and to validate the performance of the knn model (see: Table 9). The outcome for each of the run test (knn26 to knn36) for the model is as represented in Table 10. In turn, we found that majority of the k -values (i.e., knn26 to knn32) utilized in each run test of the model happened to be high representing a total of 63.6% (7 out of 11 models) (see: Table 10) of the executed parameters. Moreover, the remainder of the run tests (36.4%) also presented a high accuracy and acceptable levels of performance measures (see: Table 10).

As gathered in Tables 8, 9 and 10, the *precision* of the knn model, otherwise known as the positive predicted values, i.e., $(TP)/(TP+FP)$ determines what proportion of the predicted scores actually tallies with the actual scores as contained in the dataset. As an example, we note in Table 9 for the closest Optimal k -value knn31 ($k=\sqrt{n}$) also reported in Table 10, that the precision value which was 1.00 indicates that all the predicted vs actual scores did match. In other words, given the high precision value (1.00), we can say that there is an empirically evidence of 100% chance that a specific students' recommendation score that has been predicted by the model were essentially the correct scores.

On the other hand, *recall* (sensitivity) which describes what proportion of the students' recommendations (scores) by considering the students' gender, ave_sentiment, and emotional_valence was correctly identified by the model. For example, as reported in Table 9 for the knn36 optimal value for the model, the recall, which is also known as true positive rate, i.e., $(TP)/(TP+FN)$ equals to 1.00 (100%) which

means that the model did not necessarily miss any of the considered variables or parameters when calculating the recommendation scores.

Accordingly, when considering the *specificity* which represents the true negative rate, i.e., $(TN)/(TN + FP)$ by calculating, for instance, the proportion of the students' gender that were actually classified as either male or female. Henceforth, given the specificity score of 1.00 (100%) which is again very high, we assume or accept that those classifications by the model were correct.

Therefore, from the sensitivity (recall) result (1.00), we know that if the model predicts what a students' recommendation score for the teachers would be based on the students' gender, average sentiment, and emotional valence, then the predicted score is presumably correct. Whereas, considering the *specificity* (1.00), the result shows that if the model estimates or outputs what a students' score is; taking into account the students' gender, then there is a 100% good chance that the score is truly correct.

Finally, we determined the *accuracy* of the model, which represents the total number of the correct predictions by the model (which can be found on the diagonal-axis values of the confusion matrix – see: Table 8) divided by the total number of the test dataset that was utilized to predict the scores ($n = 300$), i.e., $(TP + TN)/(TP + TN + FP + FN) = > (165 + 28 + 25 + 16 + 9 + 15 + 9 + 8 + 9 + 6 + 10)/300$, is equals to 1.00 (100%). While, the *error-rate* which represents how often is the model classification wrong, i.e., $(FP + FN)/(TP + TN + FP + FN) = > (0 + 0)/300$, is equals to 0.00 (0%). We also combined the *precision* and *recall* into an F1-score, which means the harmonic mean of the precision and recall, i.e., $(2 \times \text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) = > (2 \times 1.00 \times 1.00)/(1.00 + 1.00)$ is equals to 1.00 (100%). Moreover, the Cohen's kappa coefficient metric (p expected is 0.9) (Carpentier et al., 2017), which measures how good the model predictions are compared to random guessing or assignment is equals to 1.00.

Thus, going by the high and acceptable values of the precision, recall, specificity, accuracy, F1-score, and error-rate by the model, which was uttermostly observed for the majority of the k -values (63.6%) in the cross-validation of the method; we concluded that the KNN textual data classification approach by using the EPDM + ML, is a good predictor and efficient method to determine what the students' recommendation scores for the teachers would be taking into account the students' gender, average sentiment, and emotional valence parameters.

9 Discussion

This study introduced the EPDM + ML model (Fig. 1), as an extension of the educational process mining and data mining model proposed in Okoye et al. (2020). This was done to show the technical and scholastic ways on how to utilize pieces of evidence idiosyncratically drawn from educational dataset to inform and improve the teaching quality and performances for the stakeholders (teachers and students). The model (EPDM + ML) was developed through the amalgamation of the Text mining and Machine learning technique we grounded on the descriptive decision theory (Baucells & Katsikopoulos, 2011; Chandler, 2017) which studies the

rationale behind the decisions that users (e.g., students) are disposed to make by means of the textual data quantification and statistical analysis. Studies that have looked into the text mining (e.g., sentiment analysis) method and its main application within the different studied contexts, have shown that machine learning techniques can be a good predictor of the students' feedback and/or recommendation of the teachers' performances or outcomes (Abu Zohair, 2019; Altrabsheh et al., 2014; De Fortuny et al., 2013; Dey et al., 2016; Litman & Forbes-Riley, 2004; Ofli et al., 2016). For instance, whereas Altrabsheh et al (2014) applied machine learning algorithms such as the Naive Bayes (NB), Complement Naive Bayes (CNB), Maximum Entropy (ME), and Support Vector Machines (SVM) to analyze the real-time students feedback. This study employs the *k*-nearest neighbour (KNN) machine learning method (Abu Alfeilat et al., 2019; Ghosh et al., 2020; Viji et al., 2020; Wong & Yeh, 2019), which have shown its usefulness for classification problems, to predict the students' evaluation of teaching and their recommendations for the teachers within the setting of higher education. In practice, the main purpose of the machine learning technique such as the one defined in this paper (EPDM + ML) is to predict what can happen across the dataset by learning the characteristics and/or relationships of some subsets or analyzed variables/parameters. Thus, by determining the number of process instances executed in turns of the *k*-values (see: Tables 8 and 9) in the available SET data, and by referencing the quantified variables (i.e., *ave_sentiment* and *emotional_valence*); we predicted the recommendation of scores for the teachers by the students based on the most occurring labels in the *k* nearest ones (Abu Alfeilat et al., 2019; Abu Zohair, 2019; Dey et al., 2016). The results show that the KNN text classification model (EPDM + ML) is a good predictor and useful technique that can be used to determine what the students' recommendation scores for the teachers would be taking into account the students' gender, average sentiment, and emotional valence, as reported and explained in detail in the earlier section.

It is important to mention from the findings on how the students evaluate their teachers' performance and recommendation through the ECOA SET survey, as shown in Tables 1 and 3, and Figs. 2 and 3; that a greater proportion of the students' comments that we analyzed ($n=85,378$) were considered to be neutral, i.e., equal to zero and largely interpreted to be positive in nature considering the comments which have shown the average sentiments and emotional valence. Thus, we note for the sentiment analysis that 76.4% (i.e., 65,219 out of 85,378) of the provided comments were classified as neutral, whilst for the emotional valence, 88.2% (i.e., 75,280 out of 85,378) were classed as neutral, respectively. On the other hand, for the comments which came out to contain some sort of either positive or negative sentiment, i.e., 20,159 out of 85,378 (23.6%), and emotional valence (11.8%), i.e., 10,098 out of 85,378; we found that the female students recommended the teachers by taking into account the sentiments ($p=.000$), whilst the male students appeared to be slightly borderline in terms of the emotions and sentiment when doing so with the closest non-significant values being $p=.056$ for emotional valence, and $p=.077$ for average sentiment, respectively. Also, when considering the different genders of the students in terms of the average sentiments scores; we note for the neutral comments (i.e., comments with zero values) that there was 35,296 out of 45,294 (77.9%) for males, and 29,923 out of 40,084 (74.7%) for the females. Whereas for

the remainder comments which shows to contain some form of sentiment, there was 9998 out of 45,294 (22.1%) for the males, and 25.3% (10,161 out of 40,084) for the females, respectively. Likewise, considering the students' gender in terms of the emotional valence scores; the males showed a total of 89.2% (40,394 out of 45,294), and females 87.0% (34,886 out of 40,084) for the neutral comments. Whereas for the positive/negative emotional valence scores, they showed a total of 10.8% (4900 out of 45,294) for males, and 13.0% (5198 out of 40,084) for females, respectively.

9.1 Implications of this study

Concerning the implications of this study, both in practice and the wider spectrum of scientific research and/or educational technologies, in particular, we note that prior studies have looked at the effect that educational technologies, such as the teaching analytical methods, have on the teaching perspectives and experiences for the students (Boring, 2017; Engen, 2019; Gallego-Arrufat et al., 2019; Gomes & Ma, 2020; Gordillo et al., 2019; Ndukwe & Daniel, 2020; Silva et al., 2019). Whereas some of the existing studies argued that the students evaluation of teaching (SET) may not necessarily be the most effective way of determining the teachers' teaching performances and assessment (Boring, 2017; Gomes & Ma, 2020), other studies have also highlighted the early indicators or success factors that have been achieved over the years, particularly through the use of the educational technologies and data to support the teaching–learning processes and development (Bowdre, 2020; Clark et al., 2020; Engen, 2019; Hilliger et al., 2020; Kori et al., 2018; Oyedotun, 2020; Raffaghelli et al., 2020; Silva et al., 2019). To note, Engen (2019) mentioned that, there is now more than ever, the necessity for ample understanding of the emerging methods and digital technologies, both in terms of what can be called the cultural-to-social aspects for use by the educators, in fostering an effective teachers-students learning processes and experience (Dimitriadis et al., 2021). Whereas, Ndukwe and Daniel (2020) explored the broad conception of usefulness and importance of the teaching analytics (TA) within the Education domain. Their review study (Ndukwe & Daniel, 2020) premeditated on establishing a framework to help describe and inform the different aspects of TA in education by developing a model that allows educationalists to gain farther insights into how the method (TA) can aid the stakeholders (e.g., teachers) to advance/enhance the several teaching dimensions, pedagogies, and outcomes in practice.

On the other hand, Gomes and Ma (2020) argued that through *engagement*, for instance, by measuring helpfulness and students' expectations, that the educators may find an alternative to SET, and its implications for practice within the different contexts or educational domain. In theory, the authors (Gomes & Ma, 2020) notes that alternative methods to students evaluation of teaching must involve observing or studying the students' emotional state or affective outcomes, thus, engaging the paradigm of the disconfirmations, by arguing helpfulness (e.g., emotional wellbeing of the students or educational support that are provided beyond the traditional classroom settings) to mean overall satisfaction for the said stakeholders (educators, teachers, students, etc.). In the same vein, by studying

the sentiment and emotional valence (intensities) of the comments provided by the students in SET towards the teachers' recommendations, which forms part of the main contributions of this paper; it is deemed to be a useful method towards achieving, both in theory and in practice, the aforementioned objectives and alternatives to SETs, particularly in leu and aftermath of the recent Covid-19 outbreak that have impacted the teaching–learning processes (IEEE, 2020b; Viner et al., 2020), and contingency plans by the Educators (Bao, 2020; Kummitha, 2020; Lin & Wang, 2021; Ma et al., 2021; TEC, 2020; Woolliscroft, 2020) in ensuring that the students are learning effectively through the several educational technologies that are used to facilitate the continuous teaching and learning, and students well-being in the diaspora. Moreover, the results of our study shows that an adequate understanding and analysis of the different factors, such as the sentiments and emotions expressed by the students, as well as, how to leverage that information to not only understand how the students evaluates the teachers by considering the gender differences or preconceptions, but also, in predicting what the students' recommendation of the teachers' performances or assessments score would be; stands to be a major contribution as it concerns efficient application, monitoring and management of the teachers-students learning processes and experiences, socio-technical and general well-being (Al-Maskari et al., 2021; Çevik & Bakioğlu, 2021; Dimitriadis et al., 2021; Garcez et al., 2021; Petersoni et al., 2018; Rapanta et al., 2020; UNESCO, 2020).

Furthermore, the main driver for this research, and the series of experimentations conducted in this paper are as follows; first, there is now more than ever an increasing need to recuperate or reinstate the teachers' and students' learning processes/experiences through emerging and innovative (state-of-the-art) methods such as the TA, following the rapidly changing educational environment, curricula, and ecosystem, or yet, what could be called the post-Covid-19 Education era or backdrops (Bao, 2020; IEEE, 2020a, b; UNESCO, 2020, 2021; Woolliscroft, 2020). Second, data about the students' evaluations of teaching (SET) are now captured and stored at an unprecedented rate within the several educational information systems and databases, which can be leveraged to provide adequate measures or solution to not only monitor, but to foster and/or ensure that the students are learning effectively. Perhaps, this has now become inevitable, in the post-Covid 19 era, particularly in connection to the many educational technologies that are being put in place by the different HEIs to help foster the teaching–learning processes for the stakeholders. In the same vein, we introduced the EPDM + ML model and its underlying analysis and implementation, to help bridge the identified gaps and challenges both in literature and in practice. Besides, the method (EPDM + ML) can be adopted by the Educationalists, Process innovators, Technologists, and Policy-makers in preparedness and/or advancement of the several educational activities and initiatives that underlie the present-day teaching and learning processes, as well as, provision of valuable and effective support for the said stakeholders at large (e.g., the teachers, students, educational community, etc.).

Data-driven methods, such as the EPDM and Machine learning technique represented in this study cannot be fully described without placing emphasis on the concept of *datafication* (Cerratto Pargman & McGrath, 2021; Prinsloo, 2017;

Raffaghelli et al., 2020; Renz & Hilbig, 2020; Slade & Prinsloo, 2013; Webb et al., 2018). The “datafication” theory or practices acknowledges different ethical considerations surrounding the outcomes or results of the data mining and/or machine learning techniques. Thus, the study deemed it necessary to discuss some of the related ethical implications of the applied methods particularly as it concerns the sociotechnical perspective on data usage within the educational context, that at the same time underlies the method of this paper. For example, existing studies that looked into the ethical challenges and perceptions in use of data within the higher education context (Slade & Galpin, 2012; Slade & Prinsloo, 2013) emphasized the need to collect and analyze educational datasets under conditions that ensure trust among the different stakeholders (e.g., HEIs, teachers, students, etc.). Having said that, the procedures or various steps and stages of analysis and handling of the educational data utilized for this study, were performed within the social structure and moral purposes/standpoints of the technical expertise (Perrotta & Williamson, 2018; Prinsloo, 2017; Prinsloo & Slade, 2017; Slade & Galpin, 2012; Slade & Prinsloo, 2013). Moreover, the resultant predictions and conclusions of the study have been made bearing in mind the veracity and variability of the captured datasets, and the need for appropriate procedures for carrying out the data-driven segmentation and diversification (Perrotta & Williamson, 2018). These also comprised taking into account the ethical pedagogies requiring the higher educational institutions to provide contextualized technical solutions that hypothetically aim to improve the effectiveness and quality of the teaching–learning processes or practice in the diaspora (Slade & Galpin, 2012; Slade & Prinsloo, 2013). Therefore, whilst the socio-technical standpoint of the different higher education institutions and the underlying (methodological) algorithmic decision-making (De Fortuny et al., 2013) and recommendation systems (Prinsloo, 2017; Prinsloo & Slade, 2017) such as the EPDM + ML model proposed in this paper, tends to offer a huge potential. The study urges that we must also recognize the ethical challenges and risks that are complementary to the Data-driven methods or datafied-Education per se (Cerratto Pargman & McGrath, 2021; Hilliger et al., 2020). Perhaps, there is no harm but instead, an ample opportunity for more intrinsic studies to span, when we acknowledge the threats and implications in applying the new and emerging technology-focused (algorithmic) data-driven decision-making (De Fortuny et al., 2013) methods, and pedagogical practices within the higher education domain or teaching–learning settings. This is especially *feasible*, and at the same time *sustainable*, when suitable and sufficient measures like sensing, processing, acting, and learning are subordinately put into place alongside the developed and implemented technologies and innovations (Prinsloo, 2017; Prinsloo et al., 2012; Renz & Hilbig, 2020).

In summary, this study aimed to identify the glitches and opportunities with teaching analytics and technologies that can be used to support the learning processes for the users especially as it concerns the students’ evaluation of teaching within the higher education context. It studied the extent to which factors such as sentiments and emotions expressed by the students in SET impact their recommendation of the teachers considering both gender constructs. In the efforts to provide answers to the research questions and objectives; we ruminated the new trends and use of state-of-the-art technologies such as the Text mining and Machine learning

towards effective teaching analytics and educational process innovation. To this end, the study proposed the educational process and data mining plus machine learning model (EPDM+ML) that proves to be effective with a high level of accuracy and efficacy towards the contextual analysis of data collected from SET in a setting of higher education. The results of the method (EPDM+ML) can be applied by educators or HEIs to understand and achieve an improved educational process/management through the data-driven and/or technology-focused solutions. However, while the authors believe that the proposed method and work therein are suitable for contextual analysis of the educational data and ample understanding of the teachers-students learning processes and/or perspectives based on the SET; this may also come with some limitations or threats to the validity of the study. For example, although the study introduced a conceptual framework and approach for analyzing the different sets of descriptive and quantifiable datasets about the SET through its proposed method, there could potentially exist or emerge other ways to approach this. The threats may also be related to the velocity, volume, variety, vagueness, and variability of the several educational datasets that are collected at an increasing rate within the education domain, or yet, modern-day educational settings. Moreover, there could also be bigger areas and analytical compositions or components that may have not been yet addressed considering the scope of this paper. Perhaps, this is because the congruence of the text mining and machine learning techniques are emerging technologies or practicalities within the educational domain, and there are not too many methods or educational design-frameworks that considers both approaches in the current literature. Henceforth, this research represents as an added incentive, both in terms of theoretical and methodological know-hows, towards a more rigorous and robust researches to come particularly within the wider areas of Educational technologies, Teaching competences, and/or Teaching analytical methods that can be employed for higher education and process management.

10 Conclusion

In this study, the authors shows that (educational) data collected from students' evaluation of teaching (SET) can be contextually analyzed using technologies such as the Text mining and Machine learning techniques. The methods (Text mining and Machine learning) can be utilized to extract and provide valuable information that can not only be used to understand the teachers-students learning processes, but can also be leveraged to drive the several educational processes forward. For this purpose, the study proposed the Educational Process and Data Mining plus Machine Learning model (EPDM+ML) that was designed based on amalgamation of the text mining and machine learning classifications to analyze the SET data. Technically, the text mining method was applied to understand the extent or intensities (polarization) of the sentiments and emotions expressed by the students when recommending the teachers in the captured SET. While the machine learning classification model was built to predict what the students' recommendation for the teachers may be based on the extracted or quantified data (average sentiment and emotional valence)

by considering the students' gender. Theoretically, this study demonstrated that the contemporary idea of applying methods such as the Text mining and Machine learning for educational purposes, is a promising practice or teaching pedagogy. This is due to the fact that the method (text mining and machine learning) can be used to provide a more robust and contextual analysis of the several educational datasets, e.g., the SET, and in consequence, employed by the Educators to not only understand the different patterns or relationships that exist within the datasets, but also utilized to improve the teaching–learning processes at large. The study has applied the EPDM+ML model using the case of SET data collected within a higher education setting to illustrate the application of the different functional elements or components of the proposed method. In practice, the study assumes that the Educationalists must take the additional responsibilities of applying the EPDM+ML model in understating the different activities that underlie the educational processes or teaching practices/performance evaluations in their differ contexts. This would not only guarantee or warrant an efficient approach towards the understanding of the teachers-students experiences, and how well to effectively improve on them. But also, by doing so, the educators would have consequently ensured to put in place a robust and effective teaching analytical method useful for educational process innovation and management. Future works can adopt the proposed model, text mining, and machine learning approach presented in this study, to analyze the various datasets collected about the students' learning processes in the different contextual domains. The further studies can also focus on reconstructing or modifying the proposed model (EPDM+ML) to include other components or functionalities that may have not already been introduced in this paper.

Acknowledgements The authors would like to acknowledge the technical and financial support of Writing Lab, Institute for Future of Education, Tecnologico de Monterrey, in the publication of this work. We would also like to acknowledge the Institutional Effectiveness Department, National Region ECOA, Tecnologico de Monterrey, Mexico for provision of the datasets used for the analysis in this study.

Author contributions KO designed the study and led the data analysis, interpretation, and writing of the article. AAP contributed in the data cleaning, interpretation, translation, and discussion of the results. CCZ contributed in the data cleaning, interpretation and discussing the results of the analysis. JGA provided and clarified the data used for the analysis. JE provided insights to the research. SH supervised the research and is the corresponding author. All authors contributed to the manuscript and have read and approved the final version.

Funding This research was funded by the host institution.

Data availability The datasets used and analyzed during this study are available from the corresponding author on request.

Code availability The codes used in performing the text mining analysis and machine learning classification model in this study are available from the corresponding author on request.

Declarations

Conflict of interests The authors declare that they have no competing interests.

References

- Abu Alfeilat, H. A., Hassanat, A. B. A., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., Eyal Salman, H. S., & Prasath, V. B. S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: A review. *Big Data*, 7(4), 221–248. <https://doi.org/10.1089/big.2018.0175>
- Abu Zohair, L. M. (2019). Prediction of Student's performance by modelling small dataset size. *International Journal of Educational Technology in Higher Education*. <https://doi.org/10.1186/s41239-019-0160-3>
- Al-Maskari, A., Al-Riyami, T., & Kunjumammed, S. K. (2021). Students academic and social concerns during COVID-19 pandemic. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-021-10592-2>
- Alao, V. M., Lansangan, J. R. G., & Barrios, E. B. (2019). Estimation of semiparametric mixed analysis of covariance model. *Communications in Statistics Simulation and Computation*. <https://doi.org/10.1080/03610918.2019.1694152>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37(April 2018), 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- 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*. <https://doi.org/10.1186/s41239-019-0137-2>
- Altrabsheh, N. (2016). *Sentiment analysis on students' real-time feedback*. University of Portsmouth.
- Altrabsheh, N., Cocca, M., & Fallahkhair, S. (2014). Sentiment analysis: Towards a tool for analysing real-time students feedback. In *2014 IEEE 26th International Conference on Tools with Artificial Intelligence*.
- Ariza-López, F. J., Rodríguez-Avi, J., Alba-Fernández, M. V., & García-Balboa, J. L. (2019). Thematic accuracy quality control by means of a set of multinomials. *Applied Sciences*, 9(20), 4240. <https://doi.org/10.3390/app9204240>
- Badri, M., Abdulla, M., Kamali, M., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly ac-credited business program in the UAE. *International Journal of Educational Management*, 20(1), 43–59. <https://doi.org/10.1108/09513540610639585>
- Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies*, 2(2), 113–115. <https://doi.org/10.1002/hbe2.191>
- Barton, E. A., & Dexter, S. (2020). Sources of teachers' self-efficacy for technology integration from formal, informal, and independent professional learning. *Educational Technology Research and Development*, 68, 89–108. <https://doi.org/10.1007/s11423-019-09671-6>
- Baucells, M., & Katsikopoulos, K. V. (2011). Descriptive models of decision making. In J. J. Cochran & L. A. Cox (Eds.), *Wiley encyclopedia of operations research and management science*. Wiley. <https://doi.org/10.1002/9780470400531.eorms0249>
- Benkowitz, A., Parkes, S., Bardy, H., Myler, K., Peters, J., Akhtar, A., Keeling, P., Preece, R., & Smith, T. (2019). Using student data: Student-staff collaborative development of compassionate pedagogic interventions based on learning analytics and mentoring. *Journal of Hospitality, Leisure, Sport and Tourism Education*, 25(January), 100202. <https://doi.org/10.1016/j.jhlste.2019.100202>
- 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*, 259–264. <https://doi.org/10.1109/DEST.2009.5276726>
- Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. <https://doi.org/10.1002/widm.1230>
- 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>
- Bowdre, P. R. (2020). The use of predictive analytics to shift the culture of academic advising toward a focus on student success. *Journal of Education & Social Policy*. <https://doi.org/10.30845/jesp.v7n3p3>

- Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., & F., Wong FM, . (2014). Learning about social learning in moods: From statistical analysis to generative model. *IEEE Transactions on Learning Technologies*, 7(4), 346–359. <https://doi.org/10.1109/TLT.2014.2337900>
- Carpentier, M., Combescore, C., Merlini, L., & Perneger, T. V. (2017). Kappa statistic to measure agreement beyond chance in free-response assessments. *BMC Medical Research Methodology*, 17(1), 1–8. <https://doi.org/10.1186/S12874-017-0340-6>
- Cech, T., Spaulding, T., & Cazier, J. (2018). Data competence maturity: Developing data-driven decision making. *Journal of Research in Innovative Teaching & Learning*, 11(2), 139–158. <https://doi.org/10.1108/JRIT-03-2018-0007>
- Cerratto Pargman, T., & McGrath, C. (2021). Be careful what you wish for! Learning analytics and the emergence of data-driven practices in higher education. *Stockholm University Press*. <https://doi.org/10.16993/BBK.1>
- Çevik, M., & Bakioğlu, B. (2021). Investigating students' E-Learning attitudes in times of crisis (COVID-19 pandemic). *Education and Information Technologies*. <https://doi.org/10.1007/s10639-021-10591-3>
- Chandler, J. (2017). Descriptive Decision Theory. In *Stanford Encyclopedia of Philosophy* (Winter 20).
- Clark, P. (2015). *The green paper needs big data*. Times Higher Education.
- Clark, J.-A., Liu, Y., & Isaias, P. (2020). Critical success factors for implementing learning analytics in higher education: A mixed-method inquiry. *Australasian Journal of Educational Technology*, 36(6), 89–106. <https://doi.org/10.14742/AJET.6164>
- Cover, T. M. (1968). Estimation by the nearest neighbor rule. *IEEE Transactions on Information Theory*, 14(1), 50–55. <https://doi.org/10.1109/TIT.1968.1054098>
- 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), 1–14. <https://doi.org/10.1145/3152892>
- Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.12230>
- De Fortuny, E. J., Martens, D., & Provost, F. (2013). Predictive modeling with big data: Is bigger really better? *Big Data*, 1(4), 215–226. <https://doi.org/10.1089/big.2013.0037>
- De Quincey, E., Kyriacou, T., Briggs, C., & Waller, R. (2019). Student centred design of a learning analytics system. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3303772.3303793>
- Dehghani, A., Glatard, T., & Shihab, E. (2019). Subject Cross Validation in Human Activity Recognition. In *ACM Conference '17, July 2017, Washington, DC, USA*. <https://doi.org/10.1145/nnnnnnn.nnnnnn>
- Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment analysis of review datasets using Naïve Bayes' and KNN classifier. *International Journal of Information Engineering and Electronic Business*, 8, 54.
- Dimitriadis, Y., Martínez-Maldonado, R., & Wiley, K. (2021). Human-centered design principles for actionable learning analytics. *Research on E-Learning and ICT in Education*. https://doi.org/10.1007/978-3-030-64363-8_15
- Dollinger, M., & Lodge, J. M. (2018). Co-Creation strategies for learning analytics. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3170358.3170372>
- 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*. <https://doi.org/10.1186/s41239-019-0153-2>
- EEOA. (2013). *Student Opinion Survey (EOA)*—(Encuesta de opinión de los alumnos). Retrieved from <https://portalrep.itesm.mx/va/encuestas/1.htm>
- El-Halees, A. (2011). Mining opinions in user-generated contents to improve course evaluation. *Software Engineering and Computer Systems*. https://doi.org/10.1007/978-3-642-22191-0_9
- Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal-Wallis analysis. *Computer Methods and Programs in Biomedicine*, 102(1), 75–80. <https://doi.org/10.1016/j.cmpb.2010.11.002>
- 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*. <https://doi.org/10.3916/C61-2019-01>

- Er, E., Gómez-Sánchez, E., Dimitriadis, Y., Bote-Lorenzo, M. L., Asensio-Pérez, J. I., & Álvarez-Álvarez, S. (2019). Aligning learning design and learning analytics through instructor involvement: A MOOC case study. *Interactive Learning Environments*, 27(5–6), 685–698. <https://doi.org/10.1080/10494820.2019.1610455>
- 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 of Computing Education*, 18(4), 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*. <https://doi.org/10.1007/s10798-019-09520-w>
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*. <https://doi.org/10.1504/IJTEL.2012.051816>
- Ferguson, R., & Clow, D. (2016). Learning analytics community exchange: Evidence hub. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/2883851.2883878>
- Frey, B. B. (2018). *Kruskal-Wallis Test*. The SAGE encyclopedia of educational research, measurement, and evaluation. SAGE Publications Inc. <https://doi.org/10.4135/9781506326139.n377>
- Gallego-Arrufat, M., Torres-Hernández, N., & Pessoa, T. (2019). Competence of future teachers in the digital security area. Competencia de futuros do-centes en el área de seguridad digital. *Comunicar*. <https://doi.org/10.3916/C61-2019-05>
- Garcez, A., Silva, R., & Franco, M. (2021). Digital transformation shaping structural pillars for academic entrepreneurship: A framework proposal and research agenda. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-021-10638-5>
- 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*. <https://doi.org/10.1080/00313831.2019.1649718>
- Ghosh, C., Saha, S., Saha, S., Ghosh, N., Singha, K., Banerjee, A., & Majumder, S. (2020). Machine learning based supplementary prediction system using K nearest neighbour algorithm. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3517197>
- Gomes, M., & Ma, W. (2020). Engaging expectations: Measuring helpfulness as an alternative to student evaluations of teaching. *Assessing Writing*, 45, 100464. <https://doi.org/10.1016/j.asw.2020.100464>
- 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*. <https://doi.org/10.3916/C61-2019-09>
- Hernández, P. A. (2013). Factores que inciden en la evaluación del desempeño docente por los alumnos de nivel superior en la Universidad TecMilenio, campus Ciudad Juárez / Factors for Teacher Performance Assessment for Upper Level Students at University TecMilenio, Campus Ciudad Juárez. *Nósis Revista De Ciencias Sociales*, 22(43–2), 188–225. <https://doi.org/10.20983/NOESIS.2013.2.5>
- Herodotou, C., Hlosta, M., Boroowa, A., Rienties, B., Zdrahal, Z., & Mangafa, C. (2019a). Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology*, 50(6), 3064–3079. <https://doi.org/10.1111/bjet.12853>
- Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., & Hlosta, M. (2019b). A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-019-09685-0>
- Hilliger, I., Laet, T. De, Henríquez, V., Guerra, J., Ortiz-Rojas, M., Zuñiga, M. Á., Baier, J., & Pérez-Sanagustín, M. (2020). For learners, with learners: Identifying indicators for an academic advising dashboard for students. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12315 LNCS (pp. 117–130). https://doi.org/10.1007/978-3-030-57717-9_9
- Holmes, W., Nguyen, Q., Zhang, J., Mavrikis, M., & Rienties, B. (2019). Learning analytics for learning design in online distance learning. *Distance Education*, 40(3), 309–329. <https://doi.org/10.1080/01587919.2019.1637716>
- IEEE. (2020a). *Active data is fueling cities' responses to COVID-19*. IEEE Transmitter. Retrieved from <https://transmitter.ieee.org/how-active-data-is-fueling-cities-responses-to-covid-19/>
- IEEE. (2020b). *How COVID-19 is affecting industry 4.0 and innovation*. IEEE Transmitter. Retrieved from <https://transmitter.ieee.org/how-covid-19-is-affecting-industry-4-0-and-the-future-of-innovation/>


- 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*. <https://doi.org/10.1186/s41239-019-0155-0>
- Kalaivani, P. (2013). Sentiment Classification of Movie Reviews by supervised machine learning approaches. *Indian Journal of Computer Science and Engineering*, 4(4), 285–292.
- 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), 1–19. <https://doi.org/10.1145/3183343>
- 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>
- Kummitha, R. K. R. (2020). Smart technologies for fighting pandemics: The techno- and human- driven approaches in controlling the virus transmission. *Government Information Quarterly*. <https://doi.org/10.1016/j.giq.2020.101481>
- LALA. (2020). *Building Capacity to use Learnig Analytics to Improve Higher Education in Latin America (LALA Project)*. Retrieved from <https://www.lalaproject.org/>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- 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/10.1177/0010880405275966>
- Lawrenz, S., Sharma, P., & Rausch, A. (2019). The significant role of metadata for data marketplaces. In *International Conference on Dublin Core and Metadata Applications* (pp. 95–101).
- Lin, Y.-J., & Wang, H. (2021). Using virtual reality to facilitate learners' creative self-efficacy and intrinsic motivation in an EFL classroom. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-021-10472-9>
- Litman, D. J., & Forbes-Riley, K. (2004). Predicting student emotions in computer-human tutoring dialogues. <https://doi.org/10.3115/1218955.1219000>
- Ma, K., Chutiyami, M., Zhang, Y., & Nicoll, S. (2021). Online teaching self-efficacy during COVID-19: Changes, its associated factors and moderators. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-021-10486-3>
- Mackness, J., Fai, S., Mak, J., & Williams, R. (2010). The ideals and reality of participating in a MOOC. In *Connectivism and Connective Knowledge, Proceedings of the 7th International Conference on Networked Learning 2010* (pp. 266–275).
- Mahmoud, M., Dafoulas, G., Abd ElAziz, R., & Saleeb, N. (2020). Learning analytics stakeholders' expectations in higher education institutions: A literature review. *The International Journal of Information and Learning Technology*, 38(1), 33–48. <https://doi.org/10.1108/IJILT-05-2020-0081>
- Mangaroska, K., & Giannakos, M. (2019). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, 12(4), 516–534. <https://doi.org/10.1109/TLT.2018.2868673>
- Martens, M., Hajibayova, L., Campana, K., Rinnert, G. C., Caniglia, J., Bakori, I. G., Kamiyama, T., Mohammed, L. A., Mupinga, D. M., & Oh, O. J. (2020). “Being on the wrong side of the digital divide”: Seeking technological interventions for education in Northeast Nigeria. *Aslib Journal of Information Management*, 72(6), 963–978. <https://doi.org/10.1108/AJIM-05-2020-0172>
- Mayer-Schönberger, V., & Cukier, K. (2014). *Learning with big data: The future of education*. Houghton Mifflin Harcourt.
- Medne, A., Lapina, I., & Zeps, A. (2020). Sustainability of a university's quality system: Adaptation of the EFQM excellence model. *International Journal of Quality and Service Sciences*, 12(1), 29–43. <https://doi.org/10.1108/IJQSS-09-2019-0108>
- Montemayor-Gallegos, J. E. (2002). *Reliability and validity of the opinion survey carried out to the students to evaluate and provide feedback on the performance of the ITESM-Single Edition teachers*. Instituto Tecnológico y de Estudios Superiores de Monterrey. Retrieved from <http://hdl.handle.net/11285/567554>
- Mourad, M. (2017). Quality assurance as a driver of information management strategy: Stakeholders' perspectives in higher education. *Journal of Enterprise Information Management*, 30(5), 779–794. <https://doi.org/10.1108/JEIM-06-2016-0104>
- Muldner, K., Burleson, W., Van De Sande, B., & Vanlehn, K. (2011). An analysis of students' gaming behaviors in an intelligent tutoring system: Predictors and impacts. *User Modeling and User-Adapted Interaction*, 21(1–2), 99–135. <https://doi.org/10.1007/s11257-010-9086-0>

- 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*. <https://doi.org/10.1186/s41239-018-0093-2>
- Ndukwe, I. G., & Daniel, B. K. (2020). Teaching analytics, value and tools for teacher data literacy: A systematic and tripartite approach. *In International Journal of Educational Technology in Higher Education*, 17(1), 22. <https://doi.org/10.1186/s41239-020-00201-6>
- Nganji, J. T. (2018). Towards learner-constructed e-learning environments for effective personal learning experiences. *Behaviour and Information Technology*. <https://doi.org/10.1080/0144929X.2018.1470673>
- Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*, 100(December 2018), 298–304. <https://doi.org/10.1016/j.chb.2018.12.019>
- Ofli, F., Meier, P., Imran, M., Castillo, C., Tuia, D., Rey, N., Briant, J., Millet, P., Reinhard, F., Parkan, M., & Joost, S. (2016). Combining human computing and machine learning to make sense of big (Aerial) data for disaster response. *Big Data*, 4(1), 47–59. <https://doi.org/10.1089/big.2014.0064>
- Okoye, K., Arrona-Palacios, A., Camacho-Zuñiga, C., Hammout, N., Luttmann-Nakamura, E., Escamilla, J., & Hosseini, S. (2020). Impact of Students Evaluation of Teaching: A Text Analysis of the Teachers Qualities by Gender. *International Journal of Educational Technology in Higher Education (ETHE)*, 17, 1–27.
- Oyedotun, T. D. (2020). Sudden change of pedagogy in education driven by COVID-19: Perspectives and evaluation from a developing country. *Research in Globalization*, 2, 100029. <https://doi.org/10.1016/j.resglo.2020.100029>
- 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. *Educational Technology and Society*, 17, 49.
- Papamitsiou, Z., & Economides, A. A. (2019). Exploring autonomous learning capacity from a self-regulated learning perspective using learning analytics. *British Journal of Educational Technology*, 50(6), 3138–3155. <https://doi.org/10.1111/bjet.12747>
- Payne, A. (2006). *Achieving excellence in customer management*. *Handbook of CRM*. Elsevier BH.
- Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial intelligence in education: Challenges and opportunities for sustainable development*. United Nations Educational, Scientific and Cultural Organization.
- 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>
- Petersoni, A., Dumontii, H., Lafuente, M., & Lawiii, N. (2018). *Understanding innovative pedagogies: Key themes to analyse new approaches to teaching and learning*. Retrieved from https://www.oecdilibrary.org/education/understanding-innovative-pedagogies_9f843a6e-en. Accessed 7 July 2021.
- Pettersson, F. (2020). Understanding digitalization and educational change in school by means of activity theory and the levels of learning concept. *Education and Information Technologies*, 26(1), 187–204. <https://doi.org/10.1007/s10639-020-10239-8>
- 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>
- Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room—The obligation to act. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3027385.3027406>
- Prinsloo, P., Slade, S., & Galpin, F. (2012). Learning analytics: Challenges, paradoxes and opportunities for mega open distance learning institutions. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/2330601.2330635>
- Raffaghelli, J. E., Manca, S., Stewart, B., Prinsloo, P., & Sangrà, A. (2020). Supporting the development of critical data literacies in higher education: Building blocks for fair data cultures in society.

- International Journal of Educational Technology in Higher Education*, 17(1), 58. <https://doi.org/10.1186/s41239-020-00235-w>
- Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online University teaching during and after the covid-19 Crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education*, 2(3), 923–945. <https://doi.org/10.1007/s42438-020-00155-y>
- Renz, A., & Hilbig, R. (2020). Prerequisites for artificial intelligence in further education: Identification of drivers, barriers, and business models of educational technology companies. *International Journal of Educational Technology in Higher Education*, 17(1), 14. <https://doi.org/10.1186/s41239-020-00193-3>
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3, 12.
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wires Data Mining and Knowledge Discovery*. <https://doi.org/10.1002/widm.1355>
- Roscoe, J. T. (1975). *Fundamental research statistics for the behavioral sciences* (2nd ed.). Berlin: Holt, Rinehart, and Winston.
- Rstudio. (2020). *RStudio—RStudio*. Retrieved from <https://rstudio.com/products/rstudio/>. Accessed 20 April 2020.
- Salinas, J. O. G., & Martínez, G. M. F. (2018). Características Personales y Práctica Docente de Profesores Universitarios y su Relación con la Evaluación del Desempeño. *Revista Iberoamericana De Evaluación Educativa*, 11(2), 9–33. <https://doi.org/10.15366/RIEE2018.11.2.001>
- Sánchez-Mena, A., Martí-Parreño, J., & Miquel-Romero, M. J. (2019). Higher education instructors' intention to use educational video games: An fsQCA approach. *Educational Technology Research and Development*, 67, 1455–1478. <https://doi.org/10.1007/s11423-019-09656-5>
- 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>
- 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*. <https://doi.org/10.3916/C61-2019-03>
- Slade, S., & Galpin, F. (2012). Learning analytics and higher education: Ethical perspectives. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/2330601.2330610>
- Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510–1529. <https://doi.org/10.1177/0002764213479366>
- TEC. (2018). *TEC21 Modelo Educativo*. Tecnológico de Monterrey (ITESM). Retrieved from <http://modelotec21.itesm.mx/files/folletomodelotec21.pdf>. Accessed 14 August 2020.
- TEC. (2020). *HyFlex + Tec |The Flexible Digital Plus Model and Virtual-InPerson Learning*. Tecnológico de Monterrey. Retrieved from <https://tec.mx/es/noticias/nacional/institucion/tec-iniciara-semester-remota-y-migrara-paulatinamente-presencial>. Accessed 19 June 2020.
- Tondeur, J., Scherer, R., Siddiq, F., Baran, E., & Tondeur JoTondeur, J. (2020). Enhancing pre-service teachers' technological pedagogical content knowledge (TPACK): A mixed-method study. *Educational Technology Research and Development*, 68, 319–343. <https://doi.org/10.1007/s11423-019-09692-1>
- Tóth, Z. E., & Surman, V. (2019). Listening to the voice of students, developing a service quality measuring and evaluating framework for a special course. *International Journal of Quality and Service Sciences*, 11(4), 455–472. <https://doi.org/10.1108/IJQSS-02-2019-0025>
- Tseng, C., Chou, J., & Tsai, Y. (2018). Text mining analysis of teaching evaluation questionnaires for the selection of outstanding teaching faculty members. *IEEE Access*, 6, 72870–72879. <https://doi.org/10.1109/ACCESS.2018.2878478>
- 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. (2014). *Global citizenship education: Preparing learners for the challenges of the 21st century*. UNESCO Digital Library. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000227729>. Accessed 24 August 2020.
- UNESCO. (2015). *Competency based education. Learning portal—planning education for improved learning outcome*. Retrieved from <https://learningportal.iep.unesco.org/en/library/competency-based-education>. Accessed 18 February 2020.

- UNESCO. (2020). *Covid-19 Education: From disruption to recovery. School closures caused by Coronavirus (Covid-19)*. Retrieved from <https://en.unesco.org/covid19/educationresponse>. Accessed 10 September 2020.
- UNESCO. (2021). *Global Education Coalition*. Retrieved from <https://en.unesco.org/covid19/educationresponse/globalcoalition>. Accessed 18 August 2021.
- van der Aalst, W. M. P. (2016). *Process mining: Data science in action*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-49851-4>
- Viji, C., Beschi Raja, J., Ponnagall, R. S., Suganthi, S. T., Parthasarathi, P., & Pandiyan, S. (2020). Efficient fuzzy based K-nearest neighbour technique for web services classification. *Microprocessors and Microsystems*, 76, 103097. <https://doi.org/10.1016/j.micpro.2020.103097>
- Viner, R. M., Russell, S. J., Croker, H., Packer, J., Ward, J., Stansfield, C., Mytton, O., Bonell, C., & Booy, R. (2020). School closure and management practices during coronavirus outbreaks including COVID-19: A rapid systematic review. *The Lancet Child and Adolescent Health*, 4(5), 397–404. [https://doi.org/10.1016/S2352-4642\(20\)30095-X](https://doi.org/10.1016/S2352-4642(20)30095-X)
- 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*. <https://doi.org/10.1186/s41239-019-0163-0>
- Wang, B., Ogburn, E. L., & Rosenblum, M. (2019). Analysis of covariance in randomized trials: More precision and valid confidence intervals, without model assumptions. *Biometrics*, 75(4), 1391–1400. <https://doi.org/10.1111/biom.13062>
- Webb, M. E., Prasse, D., Phillips, M., Djordje, M. K., Angeli, C., Strijker, A., Carvalho, A. A., Andresen, B. B., Dobozy, E., & Laugesen, H. (2018). Challenges for IT-enabled formative assessment of complex 21st Century skills SLO-National institute for curriculum development. *Technology, Knowledge and Learning*, 23, 441–456. <https://doi.org/10.1007/s10758-018-9379-7>
- 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)*. Retrieved from <https://www.cs.cmu.edu/~mwmen/papers/edm2014-camera-ready.pdf>. Accessed 27 July 2020.
- 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), 1–16. <https://doi.org/10.1145/3343195>
- 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>
- Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, 6(2), 53–69. <https://doi.org/10.18608/jla.2019.62.4>
- Wong, T.-T., & Yeh, P.-Y. (2019). Reliable Accuracy Estimates from k-fold Cross Validation. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/tkde.2019.2912815>
- Woolliscroft, J. O. (2020). Innovation in response to the COVID-19 pandemic crisis. *Academic Medicine: Journal of the Association of American Medical Colleges*. <https://doi.org/10.1097/ACM.00000000000003402>
- Xiong, Z., Cui, Y., Liu, Z., Zhao, Y., Hu, M., & Hu, J. (2020). Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation. *Computational Materials Science*, 171, 109203. <https://doi.org/10.1016/j.commatsci.2019.109203>
- 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>
- Zhou, X., Su, M., Liu, Z., Hu, Y., Sun, B., & Feng, G. (2020). Smart tour route planning algorithm based on naïve Bayes interest data mining machine learning. *ISPRS International Journal of Geo-Information*, 9(2), 112. <https://doi.org/10.3390/ijgi9020112>

Authors and Affiliations

Kingsley Okoye¹ · Arturo Arrona-Palacios¹ · Claudia Camacho-Zuñiga² · Joaquín Alejandro Guerra Achem³ · Jose Escamilla⁴ · Samira Hosseini^{1,5} 

✉ Samira Hosseini
samira.hosseini@tec.mx

- ¹ Writing Lab, Institute for Future of Education, Office of the Vice President for Research and Technology Transfer, Tecnologico de Monterrey, CP 64849 Monterrey, Nuevo Leon, Mexico
- ² School of Engineering and Sciences, Tecnologico de Monterrey, Toluca Campus, Toluca, Mexico
- ³ Office of the Vice-Rector for Academic and Educational Innovation, Vice-Rector for Professional, Tecnologico de Monterrey, 64849 Monterrey, Nuevo Leon, Mexico
- ⁴ Institute for Future of Education, Office of the Vice President for Research and Technology Transfer, Tecnologico de Monterrey, CP 64849 Monterrey, Nuevo Leon, Mexico
- ⁵ School of Engineering and Sciences, Tecnologico de Monterrey, CP 64849 Monterrey, Nuevo Leon, Mexico