12-7-2022

Applying Sentiment Analysis and Machine Learning Algorithms on Students' Reflections to Identify an Effective Teaching Strategy as a Factor of Learning Successes

Pornpat Sirithumgul
Department of Computer Engineering, Rajamangala University of Technology Phra Nakhon, pornpat.s@rmutp.ac.th

Pimpaka Prasertsilp
School of Science and Technology, Sukhothai Thammathirat Open University, pimpaka.pra@stou.ac.th

Applying Sentiment Analysis and Machine Learning Algorithms on Students' Reflections to Identify an Effective Teaching Strategy as a Factor of Learning Successes

Research-in-progress

Pornpat Sirithumgul
Department of Computer Engineering
Rajamangala University of Technology Phra Nakhon
Thailand
Email: pornpat.s@rmutp.ac.th

Pimpaka Prasertsilp
School of Science and Technology
Sukhothai Thammathirat Open University
Thailand
Email: pimpaka.pra@stou.ac.th

Abstract

This paper proposes an idea of applying sentiment analysis and machine learning algorithms to analyze over 30 thousand students' opinions, from a review site on 2,183 instructors teaching at the department of computer science in 43 universities in California, United States. In this proposal paper, we propose using a sentiment analysis method to analyze and classify students' opinions into 3 classes, namely positive, negative, and neutral. Further, we suggest applying machine learning algorithms on the classified opinions of students and some other objective data from the review site to identify an effective teaching strategy as a crucial factor of students' learning successes. Eventually in the discussion, we propose a theoretical framework, based on the expected outcome of this research to strengthen the theory of technology-mediated pedagogy.

Keywords: Sentiment Analysis, Text Analytics, Machine Learning Algorithms, Technology-Mediated Pedagogy, Teaching Strategies

1 Introduction

This current research is aimed to explore environmental factors that support a use of pedagogical technology. Under the same umbrella of this current research, we have invented a knowledge expert system (Sirithumgul et al., 2022) that could automatically generate questions for measuring students' knowledge. At run time, this system (see Figure 1) would be used by an instructor to generate questions that would be further sent to students via a learning management system (LMS). In the other way around, when students submit their answers to the questions via LMS, the knowledge expert system would analyze the answers, and then generate a report informing the instructor about students' strength and weaknesses related to the lesson taught in class.

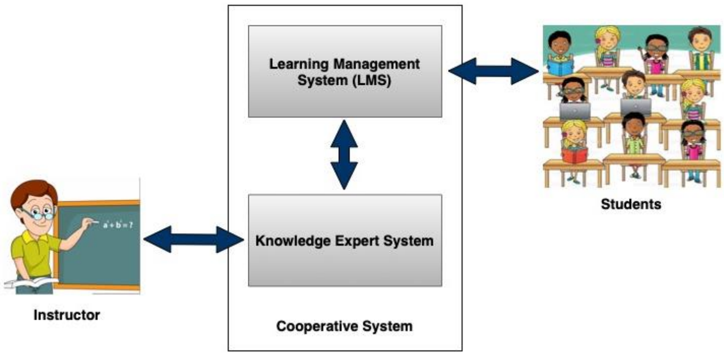

Figure 1: Pedagogical Technology in this Research

We aim to conduct this current research because we realize that the users – both instructor and student will not be able to get the most benefit from the knowledge expert system if we do not define a suitable environment for the system to be deployed. We started this research from reviewing the literature (Biggs, 1996; De Silva, 2019; Deibl et al., 2018; Khumalo, 2018; Kumar & Pande, 2017; Henderson et al., 2017; Nayar & Venkatasubramanian, 2018; Okoro et al., 2021; Oliver & Herrington, 2003; Rowe et al., 2020; and Trigwell & Prosser, 2014) in the area of technology-mediated pedagogies (TMPs) and found that an environment of the system should be defined in terms of teaching strategies that fit well with the system. We also found, based on the literature (Biggs, 1996; De Silva, 2019; Deibl et al., 2018; Khumalo, 2018; Kumar & Pande, 2017; Henderson et al., 2017; Nayar & Venkatasubramanian, 2018; Okoro et al., 2021; Oliver & Herrington, 2003; Rowe et al., 2020; and Trigwell & Prosser, 2014) that the crucial key of aligning a technology with teaching strategies was students' feedbacks.

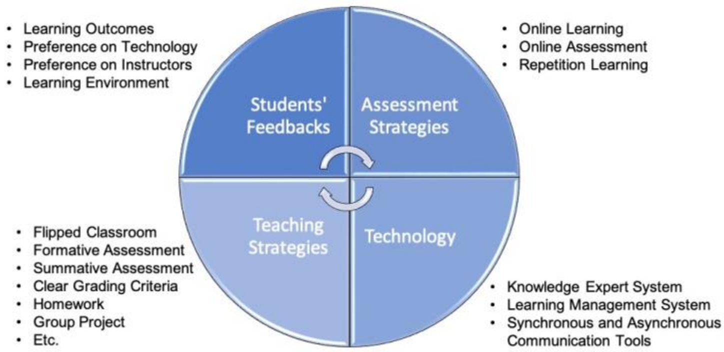

Figure 2: Cooperating Among Pedagogical Technologies, Assessment Strategies, Teaching Strategies and Students' Feedbacks

We have learned at the current stage (see Figure 2) that the knowledge expert system we invented (Sirithumgul et al., 2022) could work cooperatively with other technologies, including LMS and communication tools. Students could use all these cooperative technologies to learn, relearn and take multiple sets of exams repeatedly until they were proved an achievement of their studies. However, what we have not yet known is about what teaching strategy acceptable by students and well congruent with the technologies we have. We therefore design this current research to find a teaching strategy through analyzing students' feedbacks on instructors and learning environments constructed by the instructors.

In order to answer the research question *"What is the teaching strategy acceptable by students and congruent with the existing technologies?"*, we plan to use an inductive method that can identify an effective teaching strategy from students' opinions. Technically, this current research would employ the two AI techniques — sentiment analysis and machine learning to find the essence from textual data. Different than the previous studies (Quille and Bergin, 2016; Hung et al., 2017; Ko and Leu, 2021), this current research does not mean to use the AI techniques to examine successful attributes for students based on their own thoughts/behaviors, e.g., self-efficacy or self-regulation. Rather, we aim to use the techniques to reveal the role of instructors as a successful attribute for students.

This paper would elaborate our research design organized in four sections as follows.

1. *Section 'Dataset'* describes characteristics of data we plan to use in this research. These data are from the review site *'Rate My Professors'* that basically asks students about their opinions on instructors and classes they have taken.
2. *Sections 'Fundamental Techniques' and 'Research Method'* mainly describe the two techniques we aim to apply on the data, namely sentiment analysis and machine learning. Specifically in section *'Fundamental Techniques'*, we describe background knowledge about sentiment analysis and machine learning, and in section *'Research Method'*, we describe a step-by-step application of the two techniques on the data.
3. *Section 'Discussions'* is about our discussions on a teaching strategy as an expected outcome of this research. We also describe how the teaching strategy would be a part of a theoretical framework that supports a pedagogy technology and promotes students' learning successes.

2 Dataset

In this research, we aim to use data from the review site *'Rate My Professors'* (RMP)[1] in which students can rate, compliment, and voice their complaints on the instructors of classes they have taken. The data on this site are actually students' answers to the open-ended and closed-ended survey questions. For open-ended questions, students are allowed to write their reviews on instructors, mainly about teaching styles, and their communication skills used to convey classes' material.

5-Point Likert Scale	**Binary Choices (Yes/No)**	**Tag Choices**	**Multiple Choices**
• Rate your professor • How difficult was this professor?	• Would you take this professor again? • Was this class taken for credit? • Did this professor use textbooks? • Was attendance mandatory?	• Get ready to read • Participation matters • Group projects • Gives good feedback • Lots of homework • Beware of pop quizzes • So many papers • Etc.	• Select a grade received ○ Letter grade (i.e., A+, A, A-, B+, B, B-, C+, C, C-, D+, D, D- and F) ○ Audit/No grade ○ Drop/Withdrawal ○ Incomplete ○ Not sure yet ○ Rather not say

Table 1. Four Types of Closed-Ended Questions

There are four types of closed-ended questions in RMP (see Table 1). Two questions employ 5-point Likert scales for students to rate their preferences on instructors. Four are binary questions asking students to answer a *'yes'* or a *'no'* about a class they have taken, including (1) whether students take this class for credit, (2) whether this class is mandatory, (3) whether students would take another class with the instructor of this class, and (4) whether a textbook is required in class.

There is one closed-ended question asking students to choose 3 out of 20 tags that best describe a class and/or an instructor of a class. Last question asks students about a grade they received. The students may choose a letter grade, e.g., *A*, *B*, *C*, *D*, or *'Drop/Withdrawal'*, or *'Incomplete'* to show a level of their learning success. The students may rather choose *'Audit/No grade'* if they do not take a class for credit, or rather choose *'Not sure yet'* or *'Rather not say'* in case they have not yet known their grade at the time.

[1] Data from the website https://www.ratemyprofessors.com, accessed August 8th, 2022

In this research, we plan to apply a sentiment analysis on students' answers to open-ended questions to polarize their opinions to *'positive'*, *'negative'*, or *'neutral'* categories. We would further apply machine learning algorithms on the polarized opinions, and also on students' answers to closed-ended questions to find what specific characteristics/teaching strategies of an instructor that make students impressed and lead them to a learning success.

The data we plan to use in this research would be in the scope of feedbacks from students at the department of computer science in the universities located in the state of California, USA. With the defined scope, we found in our preliminary research that there were 35,273 students' feedbacks on 2,183 instructors in 43 universities. We expect this amount of data enough for analyzing by sentiment analysis and machine learning techniques employed in this research.

3 Fundamental Techniques

This section is dedicated to describe the essence of sentiment analysis and machine learning used in this research. In part of sentiment analysis, we would describe characteristics and an application of the three sentiment analyzers, namely TextBlob, VADER, and Google Cloud NLP. In part of machine learning, we would describe machine learning algorithms being applied to find a model and attributes of the model that can classify students into successful and unsuccessful groups.

In this research, we plan to employ 3 sentiment analyzers, rather than any one of them because we want to increase validity and reliability of analyzing text. The two analyzers, namely TextBlob and VADER are actually Python libraries providing APIs for sentiment analysis. Likewise, Google Cloud NLP is a SaaS (software as a Service) that provides all basic services about natural language processing, and we would use one of the services about sentiment analysis in this research.

An output from TextBlob would be the two properties – polarity and subjectivity (Loria, 2018). A polarity is a float between [-1, 1] where -1 shows the most extreme negative sentiment, and 1 shows the most extreme positive sentiment. A subjectivity is ranged between [0, 1]. It is an index shows a sentence (or a block of sentences) is about a factuality or a personal opinion. The higher value of subjectivity, the more a sentence is about opinion. An output from VADER (Valence Aware Dictionary for sEntiment Reasoning) is the scores in four categories, namely positive, negative, neutral, and compound (Borg & Boldt, 2020). VADER would decide a sentiment from the highest scores among the scores in positive, negative and neutral categories. The scores in compound would show an intensity of a sentiment presented in the highest-score category. An output from Google Cloud NLP shows two properties – scores and magnitude ("Google Cloud NLP API," n.d.). The scores ranged from 0.25 to 1.00 present a positive sentiment, from -0.25 to 0.25 present a neutral sentiment, and from -1.00 to -0.25 present a negative sentiment. The magnitude is an index showing the strength of a sentiment. It is ranged from 0 to infinity.

Sample Sentences	**Sentiment Analysis Tools**			**Final Output**
	TextBlob	**VADER**	**Google Cloud NLP**	
"This professor is excellent!"	**Pos.** (polarity = 1.00)	**Pos.** (compound = 0.61)	**Pos.** (magnitude = 0.90)	**Pos.**
"Wow! such a clear, brilliant teacher. Every class makes you love the material as much as he does."	**Pos.** (polarity = 0.30)	**Pos.** (compound = 0.94)	**Pos.** (magnitude = 2.90)	**Pos.**
"Terrible teacher. Doesn't clearly explain things and is very unreasonable when it comes to midterms and final questions."	**Neg.** (polarity = -0.18)	**Neg.** (compound = -0.66)	**Neg.** (magnitude = 1.40)	**Neg.**
"Grade determined by 5 exams and 5 programs. Exams were not overly difficult as long as you do the given practice. Each program has to completely pass his tests or else you cannot pass."	**Neg.** (polarity = -0.15)	**Neu.** (compound = 0.54)	**Neu.** (magnitude = 1.30)	**Neu.**

Table 2. Examples of Analyzing Sentences by TextBlob, VADER, and Google Cloud NLP

Table 2 presents sample sentences being analyzed by TextBlob, VADER, and Google Cloud NLP. An output from analyzing by VADER and Google Cloud NLP is either positive (presented by the abbreviation *'Pos.'*), negative (presented by the abbreviation *'Neg.'*), or neutral (presented by the abbreviation *'Neu.'*). Different than VADER and Google Cloud NLP, TextBlob would return only positive or negative; there is no neutral returned from this analyzer. An intensity level of a sentiment analyzed by TextBlob, VADER, and Google Cloud NLP is described by *'polarity'*, *'compound'*, and *'magnitude'*, respectively. Column *'Final Output'* in Table 2 shows the majority vote of the outputs from the three analyzers. As presented in Table 2, a final output is decided to be positive, negative, or neutral from an agreement of 2 out of 3, or all 3 analyzers.

The output from sentiment analysis and the other objective data from the survey would be further processed by machine learning (ML) techniques. The supervised learning algorithms (Kotsiantis, 2007), including Decision Tree, Bayesian Network, Naïve Bayes, Support Vector Machine, Multilayer Perceptron, Logistic Regression, and K-Nearest Neighbor would be used to produce models for classifying students into the successful and unsuccessful groups. Among these models, we will eventually find out the most effective one that shows the highest true positive (TP) and true negative (TN) rates of classifying data.

We would also employ the unsupervised learning algorithms (Ko & Leu, 2021), namely K-means and Association Rule for doing a descriptive analysis on the classification model. We would search for attributes impacting on students' learning successes from clustering data by K-means, and also search for a cause-and-effect relationship of variables from mining data by Association Rule. Table 3 shows a summary of ML algorithms and the expected outcomes.

Machine Learning (ML) Techniques		
	Algorithms	**Outputs**
Supervised Learning	• Decision Tree, Bayesian Network, Naïve Bayes, Support Vector Machine, Multilayer Perceptron, Logistic Regression, K-Nearest Neighbor	• Classification Models Dividing Students into Groups
Unsupervised Learning	• K-means, Association Rule	• Attributes Impacting on Students' Successes • A Cause-and-Effect Relationship of Variables

Table 3. Machine Learning Techniques Employed in this Research

4 Research Method

This research is conducted in four steps (see Figure 3). ***Step 1***, called *'Data Gathering'*, is for gathering students' feedbacks from the website *'Rate My Professors'*. A web scraper would be the technical tool for extracting data from the web pages at this step. ***Step 2***, called *'Data Preprocessing'*, is for pre-processing the data extracted from the web pages at Step 1. At this step, a natural language processing method would be used to restore and correct textual data. Also, the three sentiment analyzers – TextBlob, VADER, and Google Cloud NLP would be applied to transform subjective opinions of students to the three polarized opinions, namely positive, negative, and neutral.

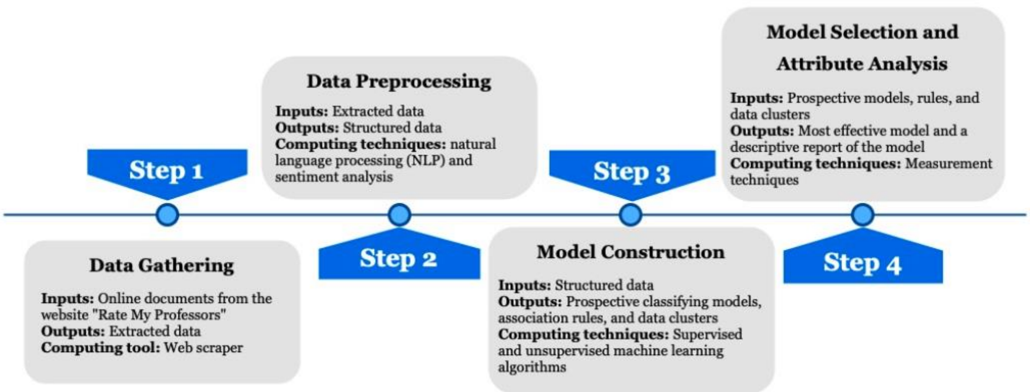

Figure 3: Four Steps Conducted in this Research

Step 3, called *'Model Construction'*, employs supervised and unsupervised machine learning algorithms to build (1) prospective models classifying students into successful and unsuccessful groups, (2) association rules revealing the relationships among variables in the dataset, and (3) data clusters showing the similarity of data in the same groups. ***Step 4***, called *'Model Selection and Attribute Analysis'*, is for choosing the most effective model from the prospective models built at Step 3. The three measurement techniques, namely accuracy, sensitivity, and specificity (Ko & Leu, 2021) would be utilized to find the best classification model. Also, at this step we would further analyze and describe the model, based on the variables presented in the top-ranked association rules, and the attributes of data in the same clusters.

5 Discussions

The outcome of this research would promote a better understanding of a teaching strategy that is admitted by students, and also helps the students to achieve their learning successes. Further, we expect to learn from the research outcome about whether there is anything else about instructors, e.g., a personality, or communication/social skills that may affect students' learning successes. We plan to use all we learn from this research to build up a theoretical framework proposed in Figure 4.

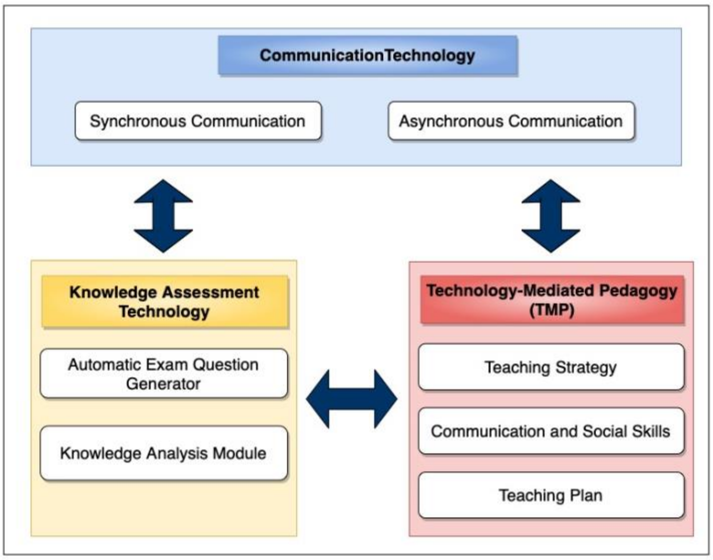

Figure 4: Theoretical Framework

As presented in the theoretical framework (see Figure 4), we define a technology-mediated pedagogy (TMP) mainly from a teaching strategy, and we plan to use the TMP together with the *'knowledge assessment technology'* – a back-end technology comprising an automatic exam question generator and a knowledge analysis module, and the *'communication technology'* – a front-end technology supporting interactions between users and the backend technology.

The interaction between the TMP and the *'knowledge assessment technology'* presented in this framework could be implied that the elements composing the TMP, including a teaching strategy, communication and social skills of instructors, and a teaching plan are adjustable, depending on students' performances being evaluated by the knowledge analysis module. As students and lessons taught in class are the two variables subject to change, this framework recommends instructors use the TMP first formed in this research as a guideline at the beginning, and further adapt it, if necessary to form a new TMP that is more suitable with their own class.

6 Conclusion

This research proposes a method for finding a teaching strategy acceptable by students, and suitable with a technology used in classroom. At the beginning, we suggest gathering students' reflections from a review site as a source of data. Further, we suggest using a sentiment analysis method to polarize students' opinions on their instructors, and then applying machine learning algorithms to find an effective factor affecting students' learning successes. Finally, we propose a theoretical framework showing a congruence between a technology and a teaching strategy constructed based on the output from this current research.

7 References

Biggs, J.B. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347-364.

Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, Volume 162, 113746.

C. -Y. Ko and F. -Y. Leu, "Examining Successful Attributes for Undergraduate Students by Applying Machine Learning Techniques," in IEEE Transactions on Education, vol. 64, no. 1, pp. 50-57, Feb. 2021, doi: 10.1109/TE.2020.3004596.

Deibl, I., Zumbach, J., Geiger, V., & Neuner, C.M. (2018). Constructive Alignment in the Field of Educational Psychology: Development and Application of a Questionnaire for Assessing Constructive Alignment. *Psychology Learning & Teaching, 17*, 293 - 307.

De Silva, Vajira. (2020). Developing Self-regulated Learning Skills in University Students Studying in the Open and Distance Learning Environment by Using KWL Method. Journal of Learning for Development. 7. 204-217. 10.56059/jl4d.v7i2.389.

Google Cloud NLP API. (n.d.). Retrieved August 15th, 2022, from https://cloud.google.com/natural-language/.

Henderson, M., Selwyn, N. & Aston, R. (2017) What works and why? Student perceptions of ‘useful’ digital technology in university teaching and learning, Studies in Higher Education, 42:8,1567-1579, DOI: 10.1080/03075079.2015.1007946

J. -L. Hung, M. C. Wang, S. Wang, M. Abdelrasoul, Y. Li and W. He, "Identifying At-Risk Students for Early Interventions—A Time-Series Clustering Approach," in IEEE Transactions on Emerging Topics in Computing, vol. 5, no. 1, pp. 45-55, Jan.-March 2017, doi: 10.1109/TETC.2015.2504239.

Khumalo, S.S. (2018). Improving Student Success Rate in Open Distance Learning Settings through the Principle of Constructive Alignment. *Trends in E-learning*.

Kumar, R.; Pande, N. Technology-mediated learning paradigm and the blended learning ecosystem: What works for working professionals? *Procedia Computer Science* 2017, 122, 1114–1123.

Loria, S. *Textblob Documentation*. 2018. Available online: https://buildmedia.readthedocs.org/media/pdf/textblob/dev/textblob.pdf. Accessed on August 15th, 2022.

Nayar, Ajitha & Venkatasubramanian, Srikirupa. (2018). Technology-Mediated Pedagogies for Skill Acquisition toward Sustainability Education. 10.5772/intechopen.74336.

Okoro, C.S., Takawira, O., & Baur, P. (2021). An assessment of tutoring performance, challenges and support during COVID-19: A qualitative study in a South African university. *Journal of University Teaching and Learning Practice*.

Oliver, Ron & Herrington, Jan. (2003). Exploring Technology-Mediated Learning from a Pedagogical Perspective. ECU Publications. 11. 10.1076/ilee.11.2.111.14136.

Quille, K., & Bergin, S. (2016). Programming: Further Factors That Influence Success. Paper presented at the Twenty-seventh Annual Workshop of the Psychology of Programming Interest Group (PPIG 2016), Cambridge, UK. Retrieved from http://www.ppig.org/library/paper/programming-further-factors-influence-success

Rowe, N., Martin, R., Buck, R., & Mabingo, A. (2020). Teaching collaborative dexterity in higher education: threshold concepts for educators. *Higher Education Research & Development, 40*, 1515 - 1529.

S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques", Informatica, vol. 31, no. 3, pp. 249-268, 2007.

Sirithumgul, Pornpat, Prasertsilp, Pimpaka & Olfman, Lorne. An Algorithm for Generating Gap-Fill Multiple Choice Questions of an Expert System. In Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS-55), Maui, Hawaii, USA. January 4 – 7, 2022.

Trigwell, K.R., & Prosser, M. (2014). Qualitative variation in constructive alignment in curriculum design. *Higher Education, 67*, 141-154.

8 Acknowledgements

This research was partially supported by National Science and Technology Development Agency (NSTDA) of Thailand [grant code FDA-CO-2562-8930-TH] and Rajamangala University of Technology Phra Nakhon, Thailand [grant code 65-1406-01/3].