9-2021

Enhancing project based learning with unsupervised learning of project reflections

Hua Leong FWA
Singapore Management University, hlfwa@smu.edu.sg

Enhancing Project Based Learning with Unsupervised Learning of Project Reflections

Hua Leong, Fwa*
Nanyang Polytechnic, Singapore
keith_fwa@nyp.edu.sg

ABSTRACT

Natural Language Processing (NLP) is an area of research and application that uses computers to analyze human text. It has seen wide adoption within several industries but few studies have investigated it for use in evaluating the effectiveness of educational interventions and pedagogies. Pedagogies such as Project based learning (PBL) centers on learners solving an authentic problem or challenge which leads to knowledge creation and higher engagement. PBL also lends itself well in plugging the gap between what is taught in classrooms and applying the knowledge gained to the real working environment. In this study, we seek to investigate how we can use NLP techniques to uncover insights into and enhance our PBL process. Both topic modelling and sentiment analysis techniques are applied to analyze final year students' reflections written as part of their final year project module. We described the entire process from text cleansing, pre-processing, modelling to visualization and evaluated the use of Latent Dirichlet Allocation and Attention Based Aspect Extraction for topic modelling. The results or visualizations which we derived from the topic and sentiment models showed that we can both effectively infer the key topics as reflected by our learners and extract actionable insights on the PBL process.

CCS CONCEPTS

• **Information systems**; • **Information systems applications**; • **Decision support systems**;

KEYWORDS

datasets, neural networks, project-based learning

ACM Reference Format:
Hua Leong, Fwa*. 2021. Enhancing Project Based Learning with Unsupervised Learning of Project Reflections. In *2021 5th International Conference on Digital Technology in Education (ICDTE 2021), September 15–17, 2021, Busan, Republic of Korea.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3488466.3488480

1 INTRODUCTION

Project-based learning (PBL) is a student-centered pedagogy that involves learners solving an authentic problem or challenge. PBL manifests from the theory of social constructivism [27] which posits that knowledge is constructed from interactions with the More Knowledgeable Others (MKOs) and that one can transcend the boundaries of knowledge or the Zone of Proximal Development through interactions with the MKOs. In addition, PBL captures the learners' interest and invokes them to engage in serious thinking as they apply and connect their knowledge in the resolution of real-world problems. In PBL, learners create their own knowledge as opposed to the case in didactic instruction where they are provided with the required knowledge by the teacher. In the process of completing a project, the learners not only acquire 21st century skills such as teamwork and communications but also learn to solve open, sophisticated, and ill-structured problems in an interdisciplinary fashion. PBL is thus postulated to lead to deeper knowledge and higher engagement of the learners and these benefits are established in a number of studies [3, 4, 6, 12].

In recent years, there have been much debate and effort to bridge the gap between higher education and the workplace. The common criticism is that the universities and colleges are not adequately preparing their graduates for the industry and there needs to be more done to contextualize what is taught and learnt to the real working environment. PBL which predicates on the application of knowledge into a real-life project, lends itself well in plugging this gap and thus gaining prominence in today's classroom.

With the prevalence of PBL, it is imperative that educators also reflect on the effectiveness of PBL and how can it be improved. Previous research [3] which reflects on effectiveness of and possible improvements to PBL have used surveys and interviews. The study by [8] investigated the effectiveness of wikis for PBL in higher education through acquiring learners feedback using the research instruments of questionnaires and interviews. Research instruments such as questionnaires and interviews are usually administered to a sample of the target population. It would however be intractable if researchers were to scale it up to entire population owing to the amount of resources required to code and interpret interview results by hand. An automated tool would be necessary if we are to scale up the analysis.

Natural Language Processing (NLP), an area of research and application that explores how computers can be used to understand and manipulate human text or speech [7] offers a panacea to the constraints of these traditional research instruments. Within the education domain, NLP has been used for purposes such as predicting learner's performance [25], automated essay scoring [1, 26], short answer scoring [16] as well as mining student's feedback [9]. [19] also employed sentiment analysis on student's feedback for evaluation of teaching quality. There is thus a growing interest in the application of NLP to improve T&L and possibly even to gain greater insights into the T&L process more efficiently. Although a

ICDTE 2021, September 15–17, 2021, Busan, Republic of Korea

ACM ISBN 978-1-4503-8499-5/21/09…$15.00
https://doi.org/10.1145/3488466.3488480

number of studies [2, 21, 25] have applied NLP in education, a gap in the literature exists on the use of NLP to evaluate effectiveness of pedagogies such as PBL.

In our context, our learners (with an average age of 19) in the institution are tasked to work on a real-world project as part of their final year curriculum for a period of 12 weeks. The projects that our learners worked on range from projects that industry has commissioned our institution to implement to problem statements from industry or research problems that our researchers are working on. We have also adopted agile software development practices and used the Microsoft Azure DevOps cloud platform as the tool to enhance our learners' software development efficiency. Microsoft Azure DevOps is a Software as a service (SAAS) cloud platform that provides an end-to-end DevOps toolchain for developing and deploying software.

Getting learners to reflect on their PBL journey is an integral part of PBL. By reflecting on the relationship between problem solving and learning, learners relate their prior knowledge to their new knowledge and understand how their problem-solving techniques can be better re-applied should another similar context occurs. Our learners are similarly required to input their reflections in the form of wikis on their final year project experience before they complete the PBL module. The wiki repository accumulated over batches of final year students offers a sizeable and rich source to be tapped for uncovering learners' opinions and sentiments on the PBL process. Currently, sporadic manual analysis is performed by individual lecturers who are interested to find out what are their learners' main areas of concerns relating to their final year project. Automated text analytics techniques such as topic modelling and sentiment analysis can potentially be applied here to gain greater insights into key topics and the related sentiments as reflected by the learners. We can then examine the topics in detail to identify potential enhancements to our PBL pedagogy.

Within the field of machine learning, there are 2 main types of tasks: supervised, and unsupervised learning. Supervised learning requires a labelled data set and it works by learning the function that maps between the inputs and the desired output or ground truth. In contrast, unsupervised learning does not require a labelled data set and its goal is to infer the natural structure within the data set. One limitation of supervised learning is the extensive manual effort required to annotate the data set and this labelling of ground truth would have to be repeated if new data is periodically added to the set and the model has to be updated. For our context, new learners' reflections and new projects are continually added into the wiki repository and conditions for the PBL process may have changed as well, thus making it impractical for us to apply supervised learning techniques. This thus led us to use unsupervised NLP techniques for the purpose of this study.

2 RELATED STUDIES

[24] applied structural topic modelling, an unsupervised learning NLP technique for analysis of Massively Online Open Courses (MOOC) pre-course surveys, discussion forum threads and course evaluations. Their analysis of pre-course surveys revealed the motivations driving participants to sign up for the MOOC courses and how it differs across gender and age groups while analysis of discussion forums and course feedback allow the authors to identify potential student gaps in knowledge and course improvements respectively.

[10] applies topic modelling and sentiment analysis on learners' course feedback as well. The authors evaluated both rule based and machine learning classifiers to extract suggestions from the textual comments of course feedbacks. The focus of their paper is on the techniques for implementing an automated system for analysing course feedback though. In contrast, our study elaborates on how the results from the analysis can be interpreted and the insights uncovered from the analysis which closes the loop from analysis to intervening for enhanced learning outcomes.

[17] used a popular unsupervised topic model algorithm – Latent Dirichlet Allocation (LDA) to automatically extract key topics from the posts in Stack Overflow. Stack Overflow is a popular question and answering forum used by about 50 million people monthly to learn and share knowledge on technology topics. The key objective of the study is to use text analytics to identify practical workplace knowledge to bridge the skills gap between college and industry. The topics identified by LDA represents challenges that are faced by practitioners of Computer Science in workplace and thus postulated to be useful in deriving the key concepts that would have to be covered in classrooms. This study corroborates the potential for the use of unsupervised topic modelling algorithm to extract key topics for possible enhancements to the teaching of Computer Science.

3 METHODOLOGY

To address the gaps identified, we aim to answer the following research questions for this study:

- How can we design and apply topics modelling and sentiment analysis techniques to automatically analyse the PBL reflection logs?
- What are the key topics mined from the learners' reflections on PBL?
- Can we uncover actionable insights relating to PBL from the mining of learners' reflections?

To answer the research questions, we applied topic modelling and sentiment analysis of our learners' reflections wikis collected from 290 learners for 42 different projects over a 6 years period from year 2014 to 2020. There was a total of 3,405 sentences constituting 79,349 words for the entire reflection corpus.

Our overall processing workflow for topic modelling and sentiment analysis of our learners' reflections is shown in Fig 1.

We have developed the entire system using Python and deployed our machine learning models using Azure Machine Learning Studio on the cloud. The web application that we developed for the visualization of the topics and sentiments analysis is hosted in a Docker container on Microsoft Azure Cloud. The benefits of using cloud computing [28] include zero up-front capital cost e.g. for purchase of computer servers, the ability of the cloud resources to auto-scale with the computing demands and the ability to leverage on common cloud services e.g. authentication which eliminates the need to re-build some of these common services each time a new system is to be developed.

The reflection wikis will have to be first extracted from the individual project sites in our institution's Azure DevOps site. Next,

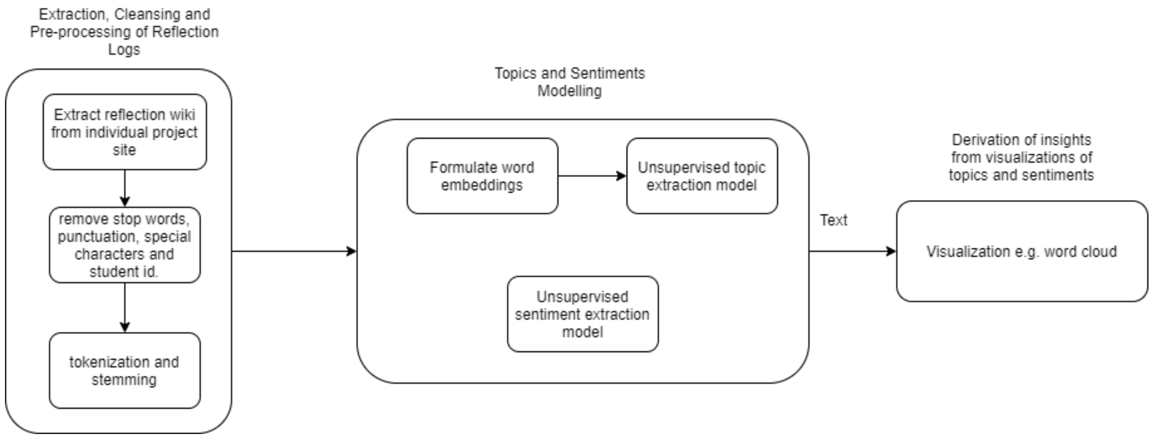

Figure 1: Processing workflow for Topic Modelling and Sentiment Analysis of learners' reflections

Figure 2: Word clouds showing keywords in each mined topics from ABAE

stop words, punctuation and special characters which include Hypertext Mark-up Language (HTML) tags are removed. The stripped text is then tokenized into individual word and stemmed into its basic form using Porter Stemmer [22]. Stop Words removal, tokenization and stemming is done using the Python NLTK library [15].

For the unsupervised topic extraction model, we have experimented with both Latent Dirichlet Allocation (LDA) [5] and Attention-based Aspect Extraction (ABAE) [11] models where aspect here is analogous to topic. The PyTorch library is used in the construction of the ABAE model. ABAE works by passing the word embeddings into an attention- based encoder to derive sentence level embeddings. The sentence embeddings are then reconstructed with the aspect embeddings. The training objective for the model is to minimize the reconstruction error and, in the process, learn the aspects or topics.

Word embeddings are vector representations of words such that two vectors $\vec{v}$ and $\overrightarrow{v'}$ corresponding to the words t and t', are close in an abstract space of N dimensions if they are of similar contexts. For the ABAE model, we passed in pre-trained word embeddings as inputs to the model. Pre-trained word embeddings captures both semantics and syntactic meanings of words as they are trained on large data sets. They are frequently used to boost the performance of NLP models. For the pre-trained word embeddings, we

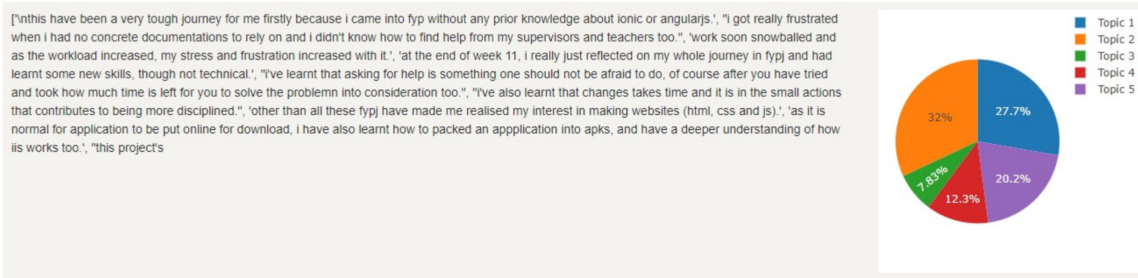

Figure 3: Topics distribution for individual reflection

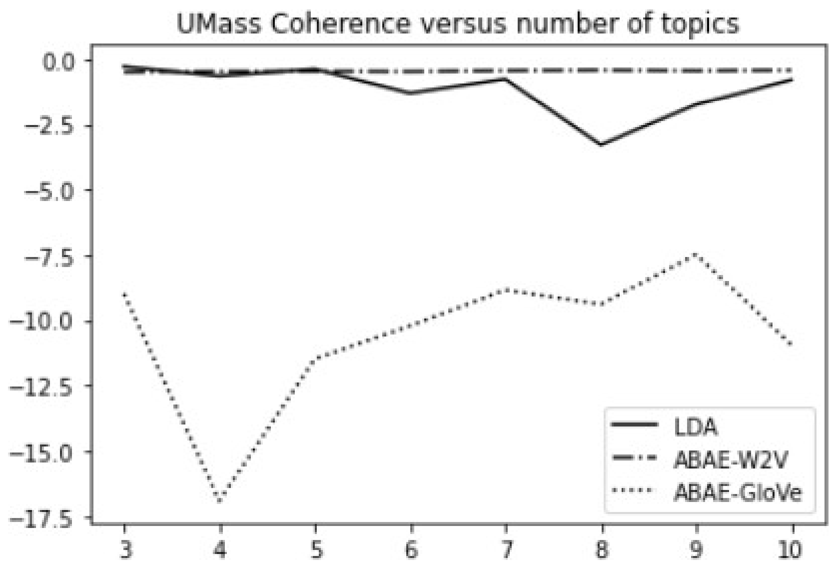

Figure 4: UMass coherence scores versus number of topics

have evaluated both Word2Vec [18] and Global Vectors for word representation (GloVe) embeddings [20].

Fig 4 shows the UMass coherence scores for LDA, ABAE(Word2Vec) and ABAE(GloVe). Running the 3 models across the number of topics (K=3 to K=10), we see that ABAE(GloVe) has the best UMass coherence scores across the topics. We used GloVe embeddings in this study as it has the lowest UMass coherence score [23] among the 3 models that we compared, indicating a higher degree of semantic similarity between high scoring words in a topic. To select the best K, we run the ABAE(GloVe) model for K=3 to K=10. Although K=4 offers the best coherence score, we selected K=5 as we found the topics to be more interpretable as compared to K=4.

We have used Valence Aware Dictionary and sEntiment Reason (VADER) library [13] to extract the sentiments for each sentence of the learners' reflection text. VADER is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. Being rule-based, we thus do not need to have a labelled dataset for training the model. VADER outputs the positive, negative and neural classification score but we only extract the positive and negative polarity score for each reflection sentence.

Finally, we built a Python web application using Plotly Dash [14] and deployed it in a Microsoft Azure Cloud Docker container to visualize the results of both the topic and sentiment models.

4 RESULTS AND DISCUSSIONS

4.1 Visualizations

A word cloud visualization for the various topics is shown in Fig 2. The key terms for each topic are displayed for each topic's word cloud. The word clouds are graphical representations that give greater prominence to words with a higher weight. In Fig 2, the size of the word in the word cloud denotes the weight of the word in the identified topic.

We can surmise from Fig 2 that Topic 1 can be summarized by the systems, software, services and support available to the learners during the period of the project. Topic 2 seems to be relating more to the people and feelings part. Topic 3 contains keywords of their friends' names and terms describing discipline e.g. lazy, hard work and scold and can thus be summarized as relating more to friends and discipline. Topic 4 contains the technical terms such as "Linux,

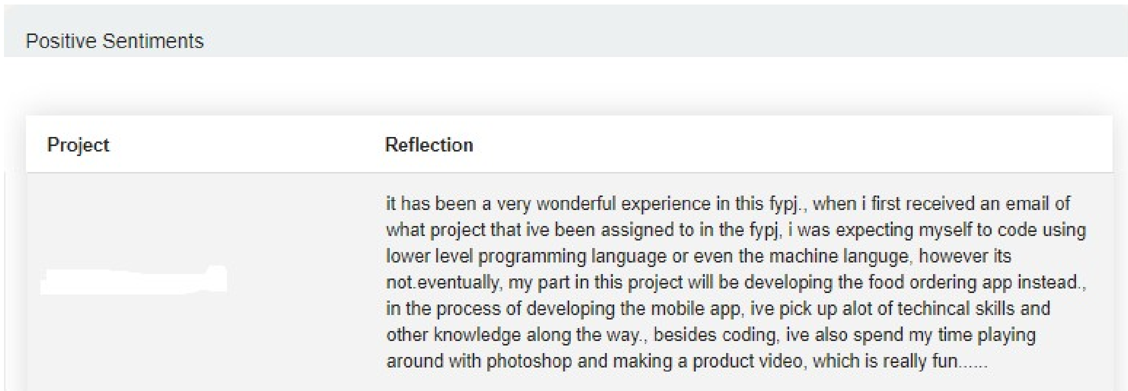

Figure 5: Top N positive sentiment reflections

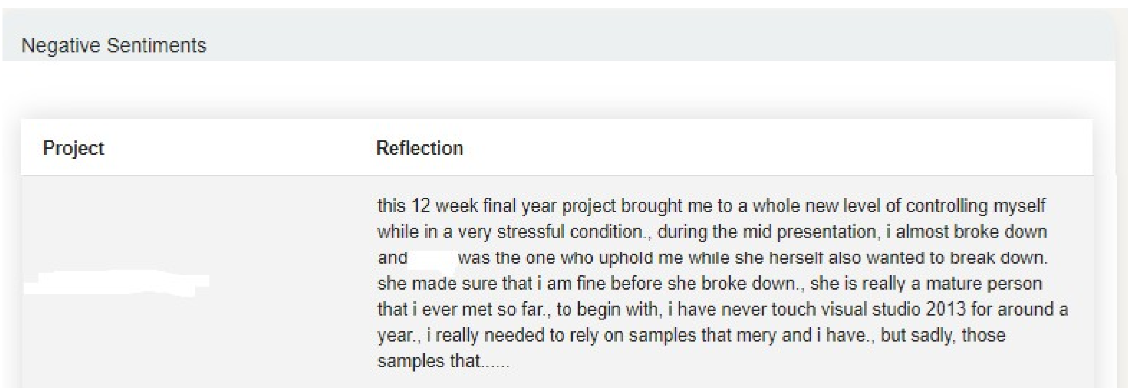

Figure 6: Top N negative sentiment reflections

PHP, JavaScript" and can thus be classified as the technical aspect. Topic 5 seems to be more about factors relating to perception and capability of self or individual.

A sample of the proportion of topics for the individual reflections is depicted in Fig 3. The topic proportions for individual reflection is calculated as the average of the topic weights over the sentences that make up the individual reflection. This provides us with a quick view of the topic composition for each reflection and allows us to investigate further into reflections that may have a higher proportion in a particular topic of interest. One finding from the visualization is the topic distribution does not differ by much among majority of the reflections. Another point to note is although topic 3 relating to friends and discipline has the lowest distribution, there is an overlap between topic 2 which relates to people and feelings and topic 3 as friends are considered people as well.

Fig 5 and Fig 6 shows the top N positive and negative reflections respectively. The sentiment score for each sentence is given by the VADER sentiment library and the scores are then averaged across the sentences that make up the reflection. A slider control allows the user to vary the value of N.

By filtering for the top n negative and positive reflections, we discovered that most of the reflections are positive. Looking into majority of the negative reflections, we find that the sentiment score is moderated by the factor of friends i.e. friends support may have moderated their negativity as can be seen from the example negative sentiment reflection in Fig 6

Filtering into the positive reflections unfortunately does not reveal much though as we were not able to pinpoint the commonality among the positive reflections. Drilling further into sentences with high positive sentiments as seen in Fig 7, again the common theme among them is the support from friends, partners and supervisors.

For the sentences with negative polarity as shown in Fig 8, majority of them reflected on knowledge gaps in programming and difficulties faced with resolving programming bugs.

4.2 Research Questions

Relating to the first research question, we have proposed and implemented an unsupervised NLP system using topic modelling, sentiment analysis and visualization for enhanced insights into the key topics and sentiments as reflected by our learners in their PBL journey.

As for the second research question, from the word clouds for the various topics, we can effectively infer the key topics surrounding PBL. From the topic distribution for the various reflections, we can also conclude that topic 1, 2 and 5 constitute the highest proportion for most of the reflections. This indicates that most of the learners

sentence
filter data...
i would like to wish the next batch good luck in this project and hopefully they can compl
i was pleasantly surprised at how friendly my partner, , was and i appreciate his re
but im lucky enough to have partner like as she is willing to help me whenever i need
lastly, i would like to thank mr , our co-supervisor, for giving good suggestions, expl
i'd also like to thank my partner, , for giving me coding tips on swift and checking i
lastly, i would like to thank to my teacher, who have being helping, support and gi
has really inspired and enabled me to think creatively and think critically, and to
i would also like to thank my teacher for giving me this opportunity to learn effectively
i am really happy to work with them on this project and i hope that the next batch of stud
from the beginning of this project i was determined to give my best effort in order to ens
i would like to sincerely thank my supervisor for her patience, support, help an
i would also like to say thanks to my partner , she had been very nice, supportive
i would also want to thank to all my friends who have been encouraging me when i encounter

Figure 7: Top N positive sentiment sentences

sentence
filter data...
this allow me to know where the error lies and tackle the problem instead of not knowing w
i really thank for being there with me all the time, when i was happy, angry, exhaust
chose wrong course from the start and lost motivation for the entire 3 years to study, lea
(upon which if anything had gone horridly wrong would leave you with a solid wall of error
also, i am not a highly motivated person and when i faced failure or errors, i will tend t
but sadly, those samples that we have were useless for the first 7 weeks since we were usi
with little and no knowledge in android studio i had a hard time trying to figure out ever
with this skill, i'm able to know what kind of error am i facing, which line of codes has
nothing impressive, screwed up presentation, gave up in the end and lost interest in schoo
we did really bad for our interim presentation and i felt really depressed for almost the
however i regret my lack of skill in coding and regret being unable to finish the pcmdb ad

Figure 8: Top N negative sentiment sentences

are concerned with the system, software, services and support (topic 1), people and feelings (topic 2) and perception and capability of themselves (topic 5). Since they are most mentioned, to efficiently enhance the learning in PBL, we can implement interventions that are related to these three topics. We can possibly investigate into the current available support for our PBL process.

For the last research question on actionable insights on PBL from mining of the learners' reflections (on a sentence level), friends' support is a common theme that has moderated the negativity of the reflections. Similarly, friends and supervisor support contributed to the positivity of the reflections. The topic distributions for the reflections also corroborated the importance of the friends' support factor. The motivation of the learner in PBL contributes significantly to the amount of learning that they can derive from the process and enhancing the engagement and motivation of the learners leads to the attainment of the desired learning outcomes. One possible intervention may be to create more opportunities for productive discussions and interactions among learners and their peers. As to the other finding on the knowledge gaps in programming and difficulties faced with resolving programming bugs, we will enhance our curriculum to further strengthen our learners' programming and debugging skills.

5 CONCLUSION

In this study, we have designed and implemented a cloud-based system that uses topic modelling and sentiment analysis for the mining of our learners' reflection logs for their final year project module for uncovering insights and enhancements to our PBL process. We have described the process flow for the data extraction, pre-processing, modelling and visualization for the system. For the topic models, we have evaluated both LDA and ABAE topic models and used ABAE as the final model for the analysis. We have also developed a containerized web application to visualize the results of the reflection text analysis.

From the visualizations that were built, we can effectively infer the key topics as reflected by our learners on the PBL process. Most of our learners are concerned with the system, software services and support, their feelings or emotions and the people they interact with as well as perceptions of self and their own ability. We were also able to extract insights on PBL from mining of the learner's reflections. A potential enhancement to our PBL that we can possibly implement is to improve the engagement and motivation of our learners by possibly creating more opportunities for discussions and interactions between them and their peers. We can also possibly investigate and extend the current support available for our learners since it is most mentioned.

In summary, we have proposed and implemented an unsupervised text analytics system to uncover insights from our learners' reflection logs to enhance our PBL process. The insights uncovered from the topic and sentiment models allow us to identify and evaluate potential interventions to enhance our PBL pedagogy.

REFERENCES

[1] Laura K Allen and Danielle S McNamara. 2015. You Are Your Words: Modeling Students' Vocabulary Knowledge with Natural Language Processing Tools. International Educational Data Mining Society (2015).
[2] Laura K Allen, Erica L Snow, and Danielle S McNamara. 2015. Are you reading my mind? Modeling students' reading comprehension skills with Natural Language Processing techniques. In Proceedings of the fifth international conference on learning analytics and knowledge. 246–254.
[3] Brigid JS Barron, Daniel L Schwartz, Nancy J Vye, Allison Moore, Anthony Petrosino, Linda Zech, and John D Bransford. 1998. Doing with understanding: Lessons from research on problem-and project-based learning. Journal of the learning sciences 7, 3-4 (1998), 271–311.
[4] Stephanie Bell. 2010. Project-based learning for the 21st century: Skills for the future. The clearing house 83, 2 (2010), 39–43.
[5] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. the Journal of machine Learning research 3 (2003), 993–1022.
[6] Noel Capon and Deanna Kuhn. 2004. What's so good about problem-based learning? Cognition and instruction 22, 1 (2004), 61–79.
[7] Gobinda G Chowdhury. 2003. Natural language processing. Annual review of information science and technology 37, 1 (2003), 51–89.
[8] Samuel Kai Wah Chu, Yin Zhang, Katherine Chen, Chi Keung Chan, Celina Wing Yi Lee, Ellen Zou, and Wilfred Lau. 2017. The effectiveness of wikis for project-based learning in different disciplines in higher education. The internet and higher education 33 (2017), 49–60.
[9] Venugopal Dhanalakshmi, Dhivya Bino, and Abinaya M Saravanan. 2016. Opinion mining from student feedback data using supervised learning algorithms. In 2016 3rd MEC international conference on big data and smart city (ICBDSC). IEEE, 1–5.
[10] Swapna Gottipati, Venky Shankararaman, and Jeff Rongsheng Lin. 2018. Text analytics approach to extract course improvement suggestions from students' feedback. Research and Practice in Technology Enhanced Learning 13, 1 (2018), 1–19.
[11] Ruidan He, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2017. An unsupervised neural attention model for aspect extraction. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, 388–397.
[12] Cindy E Hmelo-Silver. 2004. Problem-based learning: What and how do students learn? Educational psychology review 16, 3 (2004), 235–266.
[13] Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the International AAAI Conference on Web and Social Media, Vol. 8. AAAI Press.
[14] Plotly Technologies Inc. 2015. Collaborative data science. Montreal, QC. https://plot.ly
[15] Edward Loper and Steven Bird. 2002. Nltk: The natural language toolkit. arXiv preprint cs/0205028 (2002).
[16] Jiaqi Lun, Jia Zhu, Yong Tang, and Min Yang. 2020. Multiple data augmentation strategies for improving performance on automatic short answer scoring. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34. 13389–13396.
[17] Ingrid Marçal, Rogério Eduardo Garcia, Danilo Eler, and Ronaldo Celso Messias Correia. 2020. A Strategy to Enhance Computer Science Teaching Material Using Topic Modelling: Towards Overcoming The Gap Between College And Workplace Skills. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education. 366–371.
[18] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. arXiv preprint arXiv:1310.4546 (2013).
[19] Zarmeen Nasim, Quratulain Rajput, and Sajjad Haider. 2017. Sentiment analysis of student feedback using machine learning and lexicon based approaches. In 2017 international conference on research and innovation in information systems (ICRIIS). IEEE, 1–6.
[20] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 1532–1543.
[21] Sarah Elizabeth Petersen and Mari Ostendorf. 2007. Natural Language Processing Tools for Reading Level Assessment and Text Simplication for Bilingual Education. Citeseer.
[22] Martin F Porter. 1980. An algorithm for suffix stripping. Program 14, 3 (1980), 130–137.
[23] Michael Röder, Andreas Both, and Alexander Hinneburg. 2015. Exploring the space of topic coherence measures. In Proceedings of the eighth ACM international conference on Web search and data mining. 399–408.
[24] Justin Reich, Dustin Tingley, Jetson Leder-Luis, Margaret E Roberts, and Brandon Stewart. 2015. Computer-assisted reading and discovery for student generated text in massive open online courses. Journal of Learning Analytics 2, 1 (2015), 156–184–156–184.
[25] Carly Robinson, Michael Yeomans, Justin Reich, Chris Hulleman, and Hunter Gehlbach. 2016. Forecasting student achievement in MOOCs with natural language processing. In Proceedings of the sixth international conference on learning analytics and knowledge. 383–387.
[26] Kaveh Taghipour and Hwee Tou Ng. 2016. A neural approach to automated essay scoring. In Proceedings of the 2016 conference on empirical methods in natural language processing. 1882–1891.
[27] Lev Vygotsky. 1978. Interaction between learning and development. Readings on the development of children 23, 3 (1978), 34–41.
[28] Lizhe Wang, Gregor Von Laszewski, Andrew Younge, Xi He, Marcel Kunze, Jie Tao, and Cheng Fu. 2010. Cloud computing: a perspective study. New generation computing 28, 2 (2010), 137–146.