11-2022

Investigating Bloom's cognitive skills in foundation and advanced programming courses from students' discussions

Joel Jer Wei LIM
Singapore Management University, joel.lim.2018@scis.smu.edu.sg

GOTTIPATI Swapna
Singapore Management University, SWAPNAG@smu.edu.sg

Kyong Jin SHIM
Singapore Management University, kjshim@smu.edu.sg

Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Investigating Bloom's Cognitive Skills in Foundation and Advanced Programming Courses from Students' Discussions

Joel J. W. LIM, Swapna GOTTIPATI* & Kyong Jin SHIM
School of Computing and Information Systems, Singapore Management University, Singapore
joel.lim.2018, *swapnag,kjshim@smu.edu.sg

Abstract: Programming courses provide students with the skills to develop complex business applications. Teaching and learning programming is challenging, and collaborative learning is proposed to help with this challenge. Online discussion forums promote networking with other learners such that they can build knowledge collaboratively. It aids students open their horizons of thought processes to acquire cognitive skills. Cognitive analysis of discussion is critical to understand students' learning process. In this paper, we propose Bloom's taxonomy based cognitive model for programming discussion forums. We present machine learning (ML) based solution to extract students' cognitive skills. Our evaluations on compupting courses show that ensemble model performs better with an average F1-score of 76%.

Keywords: Discussion forum, bloom's taxonomy, cognitive skills, programming posts, machine learning

1. Introduction

Learning to write computer programs has never been an easy task largely due to abstraction, critical thinking, analytical abilities, etc. (Bosse & Gerosa, 2016). Foundational programming courses teach students how to develop console applications, web applications, and other applications with simple features. Whereas advanced programming courses such as Object Oriented Programming and Advanced Web Application Development, train students to develop more complex and robust software applications. Discussion forums for programming courses are part of collaborative learning that provide an interactive learning environment for cognitive participation in the form of questions and answers.

Cognitive analysis of discussions is critical to understanding students' learning processes (Garrison, 2003). Six levels of Bloom's taxonomy (Bloom, 1956) are widely used for evaluating the cognitive levels of learner's knowledge. The cognitive skills analysis in programming courses is not similar to non-programming courses due to the nature of posts and learning objectives (Cabo, 2015). To support cognitive analysis for programming courses, Fuller (2007) developed a Computer Science specific learning taxonomy. Analysing programming posts aid instructor to prepare effective teaching methods to improve student's performance (Dunlosky, 2013).

Manually analyzing discussions to understand the cognitive skills of students in programming courses is a tedious job. Several researchers have proposed automated content analyses of discussions using machine learning techniques (Ntourmas, 2019; SHAH, 2021). In this paper, we take a similar approach. First of all, we define Bloom's taxonomy based cognitive model on discussion posts for programming courses. Secondly, we develop a solution model based on machine learning to analyze the content. In particular, each post is classified into a Bloom's cognitive level using feature modelling and classification algorithms. Finally, we apply the solution model to the foundational and advanced programming courses to extract and compare students' cognitive skills and interactions. One key contribution of this research work is the framework that aligns Bloom's cognitive skills to programming discussion forums. Second contribution is ML based solution design for classifying discussion posts.

2. Related Work

Discussion forums are special platforms where students exhibit both cognitive and social skills. Schellens (2005) proposed a social media cognitive taxonomy based on the users' behaviour in

discussion forums. It is based on task-related contributions and cognitive dimensions, and they are categorized as new information, explicitation and evaluation. Since these categories are limited and not very useful for several disciplines, several researchers used the idea of applying modified Bloom's taxonomy to discussion forums. For example, Stump et al. (2013) proposed modified Bloom's framework to analyze the cognitive skills of students in MOOC forums specific to the Engineering discipline - Circuits and Electronics. Wong et al. (2015) proposed a cognitive classification framework based on Bloom's taxonomy for discussion forums and applied it to a MOOC forum related to art discipline - Introduction to Art. Automated classification of forum posts can be achieved with machine learning models (Ntourmas, 2019, Shah, 2021, & Khodeir, 2021). In this paper, we propose a Bloom's taxonomy-based framework for discussion forums related to programming courses. To automatically classify the posts, we adopt machine learning techniques used by previous works. We further analyze posts and apply visual models for pattern discovery and comparison analysis.

3. Framework for Cognitive Analysis of Programming Discussions

Previous studies (Kovanović et al. 2016) have pointed to utilizing Bloom's taxonomy to evaluate students' understanding of topics in online discussion forums. Based on these studies, Wong et al. (2015) proposed a cognitive classification framework for discussion forums. We integrated both works and Table 1 shows our proposed framework to classify the programming discussion posts.

Table 1. *Cognitive classification of posts in programming discussion forums*

Label	Non-programming course forums	Programming course forums	Bloom Levels
0	Message that is irrelevant to the concepts such as acknowledgement or greeting, etc.	A post that is irrelevant to the concepts such as acknowledgement or greeting, etc.	Non-Cognitive
1	Message must translate, construe, interpret, or extrapolate information. Message must describe, list, or name factual information.	A question post with no detailed explanation. A reply post with the direct answer with no explanation or pointer to another post.	Understand & Remember
2	Message must exploit information and put the resulting knowledge into action	A reply to the question using only theory without explanation but provides links to resources. A post with questions related to concepts and with clear explanations.	Apply
3	Message must appraise or relate information to the real world	A reply to the question with an attempt to explain, paraphrase and contextualize the concepts.	Evaluate
4	Message must deduce, scrutinize, or survey information	A reply that evaluates the pros and cons of the answer or solution in context to the question.	Analyze
5	Message must formulate, generate, restructure, or combine information.	A post that summarizes various concepts discussed with respect to question or related to course topic.	Create

In our preliminary data analysis, we observed non-programming related posts and these posts are still useful to understand the interaction levels of the students in a course. Therefore, we propose Level-0 to capture such posts. We also combined two levels, understand, and remember as, unlike programming assessments, in the context of discussion forums, it is difficult to distinguish both the levels. To explain this challenge, we use example posts as shown in Table 2. For level 5, we used the example from developer forums since it is more detailed.

Table 2. *Cognitive levels and example posts*

Label	Bloom's Level	Programming course forums

0	Non Cognitive	<@U0193N90XN0> ikr :joy: Post 1- Happy Birthday <@U017X1BHQP6>!:wink: Post 2- got it!\ Post 3- now in school with XXX practicing WAD
1	Understand & Remember	Post 1- hi! does anyone know how to prevent the underline just for 'j'? Post 2- Perhaps you could try not including float for #side-bar
2	Apply	hi anyone know why such behaviour is displaying? I am currently using font-size:150% for the button which by right should display 1.5x bigger than the rest of the font and it looks fine when rendered as a website but when rendered in mobile version, the font shrinks and looks smaller in comparison to the rest of the text (Web Vs Mobile-Iphone 5)
3	Evaluate	you could try making the larger image visible only when the size is >md and making it hidden once size <md. likewise you can make the smaller image visible only when size is <md and making it hidden once size is more than md. you can refer to <https://getbootstrap.com/docs/4.0/utilities/display/> under hiding elements

4	Analyze	Perhaps, you can try adding an id to the and add in the style :
		li-design {
		text-decoration: none;
		border-bottom: dotted;
		}
		<li id = "li-design">Handsome
		So that the entire will be displayed with that specific style.
		Hope this helps!

5	Create	Xcode supports source code for the programming languages **C, C++**, **Objective-C**, **Objective-C++**, **Java**, **AppleScript**, **Python**, **Ruby**, ResEdit (Rez), and **Swift**, with a variety of programming models, including but not limited to **Cocoa**, Carbon, and **Java**. Third parties have added support for GNU Pascal,[5] Free Pascal,[6] Ada,[7] C#,[8] Perl,[9] and D.[10] Source : h ttps://en.wikipedia.org/wiki/Xcode Which one would you want to add ? I've never ventured in those areas. But there are XCode extensions. https://developer.apple.com/documentation/xcodekit/creating_a_source_editor_extension And a few implementation tutorials https://www.vadimbulavin.com/xcode-source-editor-extension-tutorial/ https://code.tutsplus.com/tutorials/how-to-create-an-xcode-source-editor-extension--cms-26772

4. Solution Design

The solution design overview consists of two key stages: Data cleaning and, Post classification.

Data Cleaning: Instead of the traditional discussion forums in Learning Management Systems the faculty used Slack which is a popular messaging application. It was originally built for businesses. However, by March 2020, over 3,000 colleges and universities had adopted Slack in their classrooms (Slack, 2020). We first parsed JSON files from Slack into an excel workbook and then extracted specific fields; student message, reply Message, Replying target, Discussion ID, Links, date, and time of post. As the first step of normalizing the post, we tokenized it, converted it into lowercase, removed emojis and removed stop words using NLP library and regex (Kulkarni, 2019). We removed posts with a single word. We consider both textual contextualized information and post characteristics such as length, to classify posts into cognitive categories.

Discussion Post Classification: The second stage of the solution design is categorizing the posts into Bloom's cognitive levels. We first generate the post features using Linguistic Inquiry and Word Count (LIWC) (Tausczik, 2010) which provides seven categories in relation to cognitive ability. The post types range from a set of simple language to complex words. Dale-Chall readability score provides the measure of comprehension difficulty when reading text (McClure, 1987). We then train and test the classifiers on the post and choose the classifier with the best performance, in this case, F1-score. The key features used in training the classifier are post tokens, LIWC cognitive dimensions and Dale–Chall readability score. This study explores the use of BERT (Devlin et al. 2019), AdaBoost (Schapire, 2013) and XGBoost (Chen & Guestrin, 2016) algorithms to build the classifier. We also use the ensemble method to ensemble all three algorithms.

5. Experiments and Results

In this section, we first describe the data preparation followed by the results and analysis of the solution model on discussion forums from foundation and advanced level programming courses.

5.1 Background of the Courses

In our school, all students must complete two web application development courses. Web Application Development course (WAD I) is a foundation course that equips students with the knowledge and skills to develop database-driven web applications using HTML, PHP, and MySQL. Upon successful completion of WAD I, the advanced course, Web Application Development course (WAD II) teaches students how to develop well-styled and responsive web applications that provide rich user experiences using HTML, CSS, Bootstrap, JavaScript, and Vue.js.

5.2 Cognitive Levels Classification

The team collected 816 posts from both the courses, and we labeled random 773 posts (300 for training and 443 posts for testing). For labeling, the members used rule-based approach to assign cognitive levels to each post (Anderson et al. 2005) and we took the common label. The team used cognitive framework described in Table 1. Three coders were trained on how to label the discussion posts. Table 3 shows F1-score results of four classification models and Ensemble produced best performance.

Table 3. *Classifier performance - F1-Score comparison on testing data*

Course	BERT	XGBoost	AdaBoost	Ensemble
WAD I	0.37	0.66	0.61	**0.8**
WAD II	0.15	0.59	0.54	**0.72**

Most of the posts in Slack consist of short replies which create an impediment that prevents the BERT from fully contextualizing and associating words. The wide range of topics discussed in the Slack forum further dilutes word associations and creates further difficulty for the model to effectively pick keywords. As such the overall accuracy of the Slack models has fallen to 37% and 15% for WAD I and WAD II respectively. While the BERT is largely successful in its performance in multiple NLP-related tasks, its performance is significantly dependent on support from a large dataset.

For WAD I, the overall accuracy of the XGBoost and AdaBoost are 66% and 61% respectively. There is a notable 29% and 24% increase in accuracy from the XGBoost and AdaBoost compared to the BERT. For WAD II, the overall accuracy of the XGBoost and AdaBoost model are 59% and 54% respectively. There is a notable 44% and 39% increase in accuracy from the XGBoost and AdaBoost compared to BERT. The overall accuracy of the ensemble model for each forum dataset is higher than the attained accuracy from each individual three machine learning classifiers. Figure 1 shows the precision, recall and F1-score comparisons.

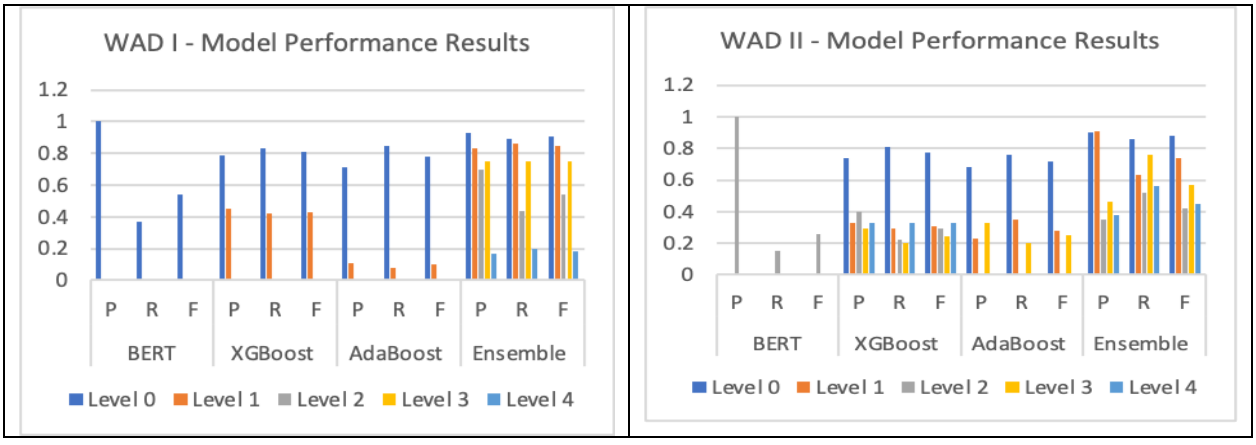

Figure 1. Precision, recall & F1-score comparison of the classifiers for each level.

From Figure 1, we observe that the ensemble model has performed better on the levels. We also notice that level 5 is not extracted. This is due to the smaller data sets for this level which we will explain next. In the final step of creating a cognitive classifier, we use ensemble models to label the remaining data. In total, we have 816 posts which we use for comparative analysis in the next section.

5.3 Comparison Analysis between Foundation and Advance Programming Courses

Figure 2 depicts the cognitive analysis among the participants in both the forums.

Level	WAD I	WAD II
0	156	242
1	98	177
2	19	40
3	25	42
4	6	11

(a) Comparison by number of posts

Level	WAD I	WAD II
0	**51.3%**	47.3%
1	32.2%	**34.6%**
2	6.3%	**7.8%**
3	8.2%	8.2%
4	2.0%	**2.1%**

(b) Comparison by percentage of posts

Figure 2. Cognitive analysis comparison on full dataset.

From Figure 2, we observe that although the students from advanced programming post extensively compared to the foundation level, the cognitive levels are statistics by proportion are very similar. We observe that most simple questions and answers (level 1) are provided for both courses. From the comparative analysis, we observe that the higher order thinking skills are not evident in the posts which agrees with the findings by Johnson & Fuller (2006). We have noticed that in developer forums, higher order cognitive levels are evident and this is due to the type of the questions. Therefore, the instructors should post questions that require higher order skills so that the students will try to answer and instructors can analyse the gaps in the knowledge.

5.4 Limitations

Our proposed solution worked well for the discussion forum within the chosen information systems programming courses. Our experiments show that ensemble model performs better than the three machine learning models. However, this may not be true for other courses. Therefore, we can improve the solution model to consider the code related keywords to improve the performance of the models and to generalize across all programming courses. The second limitation is the discussion forum settings. We used Slack which provides data extraction in JSON format which outputs more clean data. This may not be true for other discussion forums and the extraction process may end up with noise in the data. To handle the noise (e.g. HTML tags), the researchers need to adopt extensive cleaning techniques. Further, the data used in this study is a limitation for the model performance for higher order cognitive skills classification. To overcome this, we can enhance the data by using the programming posts from Stack Overflow which is a popular forum among software developers. An interesting future work is to align the cognitive levels to the teaching concepts. The summarization based on the teaching concepts and cognitive statistics enables the teaching faculty to understand which programming concepts the students are finding difficult. We are currently working on the cognitive level summaries and student-level personalized journey dashboard that can be shared with the students.

6. Conclusion

Discussion forums play a key role in collaborative learning and aid in students' cognitive skills development. This paper presents a Bloom's taxonomy-based framework for discussion forums related to programming courses. Our solution design adopts existing machine learning techniques to automatically classify discussion forum posts. We evaluated the solution model on discussion posts of foundation and advanced programming courses from the computing curriculum. We analysed students' cognitive levels for both courses. The main findings show that students post both cognitive and non-cognitive messages, and most of the posts are at the lower cognitive level.

Acknowledgements

We would like to thank all the students who participated in this research.

References

Anderson, L. W. (2005). Objectives evaluation and the improvement of education. *Studies in Educational Evaluation*, 31(2): 102-113.Availableat: https://doi.org/10.1016/j.stueduc.2005.05.004

Bloom, B. S. Taxonomy of Educational Objectives: The Classification of Educational Goals. New York: Longmans, Green, 1956. Print.

Bosse, Y. & Gerosa, M. A. (2016). Why is programming so difficult to learn? Patterns of Difficulties Related to Programming Learning. *ACM SIGSOFT Software Engineering Notes*. 41(6).

Cabo, C. (2015). Quantifying student progress through Bloom's taxonomy cognitive categories in computer programming courses. *ASEE*. 122. 10.18260/p.24632.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. ArXiv, abs/1810.04805.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public interest, 14(1), 4-58.

Fuller, U. & Johnson, C. & Ahoniemi, T. & Cukierman, D., Hernán-Losada, I., Jacková, J., Lahtinen, E., Lewis, T., Thompson, D., Riedesel, C. & Thompson, E. (2007). Developing a computer science-specific learning taxonomy. *ACM SIGCSE Bulletin*. 39. 152-170. 10.1145/1345375.1345438.

Garrison, D. (2003). Cognitive presence for effective asynchronous online learning: The role of reflective inquiry, self-direction and metacognition. *Elements of Quality Online Education: Practice and Direction*. 4.

Johnson, C. & Fuller, U. (2006). Is Bloom's taxonomy appropriate for computer science?, *presented at Baltic Sea Conference on Computing Education Research: Koli Calling*.

Khodeir, N. A. (2021) Bi-GRU Urgent Classification for MOOC Discussion Forums Based on BERT. *IEEE Access*, vol. 9, pp. 58243-58255, 2021, doi: 10.1109/ACCESS.2021.3072734.

Kulkarni, A., & Shivananda, A. (2019). Natural language processing recipes. Apress.

McClure, G. (1987). "Readability formulas: Useful or useless. (an interview with J. Peter Kincaid.)". IEEE Transactions on Professional Communication. 30: 12–15. S2CID 13157772.

Ntourmas, A., Avouris, N., Daskalaki, S., & Dimitriadis, Y. (2019) Comparative Study of Two Different Mooc Forums Posts Classifiers: Analysis and Generalizability Issues. *IISA* 1-8, doi: 10.1109/IISA.2019.8900682.

Schellens, T. & Valcke, M. (2005) Collaborative learning in asynchronous discussion groups: What about the impact on cognitive processing? *Computers in Human Behavior*, 21(6), 957-975.

Schapire, R. E. (2013). Explaining adaboost. In *Empirical inference* (pp. 37–52). Springer.

Shah, U., Rambhia, R., Kothari, P., Ramesh, R., & Banerjee, G. (2021) Automatic Classification of MOOC Forum Messages to Measure the Quality of Peer Interaction. *ICCE, 2021*.

Shim, K. J., Gottipati, S., & Lau, Y. M. (2021). Flip & slack – Active flipped classroom learning with collaborative slack interactions. *ICCE 2021*. November 22-26. 121-131.

Slack. (2020, March 16). Distance learning thrives in Slack. https://slack.com/blog/collaboration/distance-learning-in-slack

Stump, G. S., DeBoer, J., Whittinghill, J., & Breslow, L. (2013, December). Development of a framework to classify MOOC discussion forum posts: Methodology and challenges. *In NIPS Workshop on Data Driven Education*, 1-20.

Tausczik, Y. R. & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*. 29, 24–54.

Wong, J.-S. & Pursel, B. & Divinsky, A. & Jansen, J. (2015). Analyzing MOOC discussion forum messages to identify cognitive learning information exchanges. *Proceedings of the Association for Information Science and Technology*. 52. 1-10. 10.1002/pra2.2015.145052010023.