International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137

Volume – 07, Issue – 11, November 2022, PP – 35-41

Ml-Based Classification System as a Mapping Technique of Massive Open Online Courses

Ayse Kok Arslan

Abstract: One of the most problematic areas in online learning is how to improve the quality assurance of digital learning systems. Analysis and classification of MOOCS is a difficult task, given the variability of MOOC structures, contents, designs, platforms, providers, and learner profiles. To overcome this challenge, this study aims to propose an automatic and large-scale ML-based classification system for MOOCs according to their learning objectives by making use of the six cognitive levels of Bloom's taxonomy. As a result of the research, a representation and a detailed analysis of the dataset for experimentation with the different models are provided. Further research can focus on the privacy implications of the current control on developments of AI taking into account creativity, and innovation which can hardly be performed by machines.

Keywords: digital learning systems, learning objectives, Massive Open Online Courses (MOOCs), Bloom's taxonomy, machine learning, software, training software, education, elearning, AI, computing, technology

1. Introduction

Given the cognitive level of learning objectives (LOs),a few scholars[1-3] proposed an evidence-based massive online open course (MOOC) ontology, that served as a standard to unify the representation of MOOCS and facilitate interoperability between MOOCs platforms.

The aim of the study is to enrich this ontology with metadata about learning objectives classified according to Bloom's taxonomy. This also helps to automatically extract semantically rich descriptive metadata from different MOOC providers and integrate this metadata into a repository.

As the aim is the cognitive classification of MOOCs according to their learning objectives, this study recommends to adopt the basic architecture of BERT and then to add an output layer for the classification. The output layer can be either a simple classifier like softmax or a more complicated network like the bidirectional Bi-LSTM.

The theoretical foundation of Bloom's taxonomy is most appropriate for this study context since it covers the different levels of cognitive learning and allowed for classifying learning objectives according to six hierarchical levels.

The study concludes with a mapping MOOC learning objectives and Bloom's taxonomy levels based on a cognitive classification of MOOCs.

2. Review of Existing Studies

Various theoretical frameworks have been adopted for the evaluation and classification of MOOCs. [1] associated good learning with quality learning as it was critical to meet the characteristics of good learning in order to accomplish effective learning. [2] based the 12-dimensional assessment framework, as well as the 7Cs for learning design framework on this principle[4].

The existing research has addressed one of the following [4]:

- Frameworks developed for quality assurance that are generalist and lack means to operationalize the review of MOOCs' quality.
- Case studies that detailed the design of individual MOOCs to highlight best practices and pedagogical models which are not based on a well-defined evaluation framework.
- Descriptive frameworks that were intended for designing MOOCs from scratch.
- Evaluation frameworks that dealt with several dimensions including the pedagogical.

No research has been done on the automatic classification of LOs. Machine learning has been used most often, followed by the rule-based approach. The deep learning approach has been used less often; only the ANN architecture has been tested in this context. BERT was used in a single study for cognitive classification purposes [6].There has been some research comparing BERT and other machine learning or deep learning models as well [4].

Some scholars[10, 11] used machine learning for a large analysis of MOOCs. However, the number of MOOCs they examined remained limited, and their data collection methods were manual[12]. These researchers [12]used machine learning for the analysis of about 20 MOOCs. Nevertheless, the result of their clustering cannot be generalized given the limited number of MOOCs they analyzed.

3. Overview of BERT-based Model

BERT is the state-of-the-art technique in NLP[14, 17]and it has demonstrated its performance on small datasets.

During pre-training, the model is trained on a large unlabeled corpus. The model is then fine-tuned, starting with the pre-trained parameters and refining all parameters with task-specific labeled data.

A simple transformer consists of an encoder that reads text input and a decoder that generates a task prediction. BERT requires only the encoder depicted in Fig. 1 as the objective is to develop a model of the language representation.

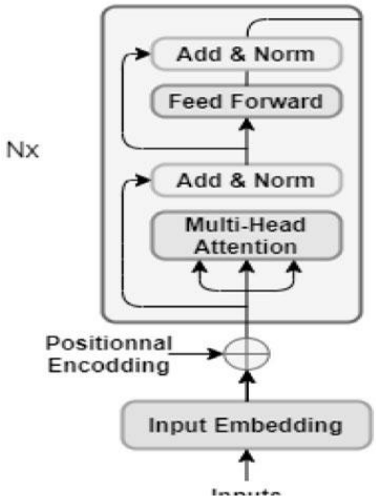

Fig. 1. The BERT Encoder:

Values: C – 'At the end' …; T_1,- 'of the module' T_2,-'the learner' …, *'will be able' 'to analyze* …; E_{CLS} – 'The module'…; E_1 , E_2–'aims to equip the learner', …, E_N – …; E_{SEP} – …; *CLS* –' the module's objective' …; tok_1, tok_2 – "is to enhance' …, tok_N – "analytical thinking skills'…– …

BERT represents a single sentence or a pair of sentences as a sequence of tokens with the following characteristics [11]:

- The first token in the sequence is *CLS*.
- When there is a pair of sentences in the sequence, they are separated by the token *SEP*.
- For a given token, its input representation is constructed by summing the corresponding token, position, and segment embeddings (Fig. 2).

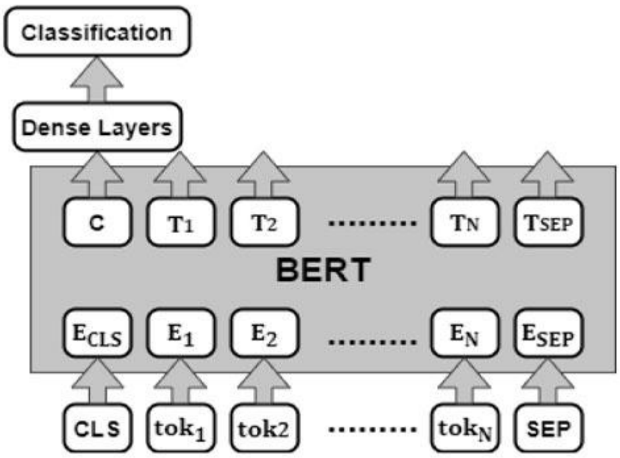

Fig. 2. BERT Architecture:

Values: C – 'At the end' …; T_1,- 'of the module' T_2,-'the learner' …, *'will be able' 'to analyze* …; E_{CLS} – 'The module'…; E_1 , E_2_'aims to equip the learner', …, E_N – …; E_{SEP} – …; *CLS* –' the module's objective' …; tok_1, tok_2 _ "is to enhance' …, tok_N – "analytical thinking skills'…– …

Moreover, BERT fine-tuning involves training a classifier with different layers on top of the pre-trained BERT transformer to minimize task-specific parameters. Fine-tuning for a specific task can be done using several approaches, either by fine-tuning the architecture or by fine-tuning different hyper-parameters such as the learning rate or the choice of the best optimization algorithm [15, 19].

If the classification problem is multi-class, the output layer is based on a softmax activation layer. An example has been provided below:

$$\sigma\left(\vec{z}\right) = \frac{e^{zi}}{\sum_{j=1}^{K} e^{zj}},$$

Where $\vec{z} = \left(z_1; \ : \ : \ : \ ; \ z_K\right);$ z_i values are the elements of the input vector to the softmax function; K is the number of classes in the multi-class classifier. The output node with the highest probability is then chosen as the predicted label for the input.

For preprocessing, one can simply clean the text of non-alphabetic characters and converted it to lower case (Fig. 3).

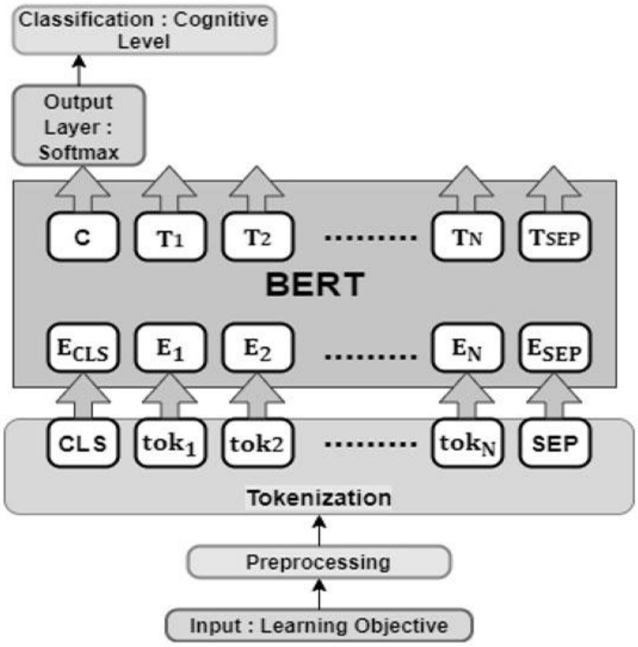

Fig. 3. BERT-Based Fine-Tuning Architecture:

Values: C – 'At the end' …; T_1,- 'of the module' T_2,-'the learner' …, *'will be able' 'to analyze* …; E_{CLS} – 'The module'…; E_1 , E_2_'aims to equip the learner', …, E_N – …; E_{SEP} – …; *CLS* –' the module's objective' …; tok_1, tok_2 _ "is to enhance' …, tok_N – "analytical thinking skills'…– …

The fully connected layer took the output of BERT's 12 layers and transformed it into the final output of six classes that represented the six cognitive levels of Bloom's taxonomy (Fig. 4). This layer consists of three dense layers.

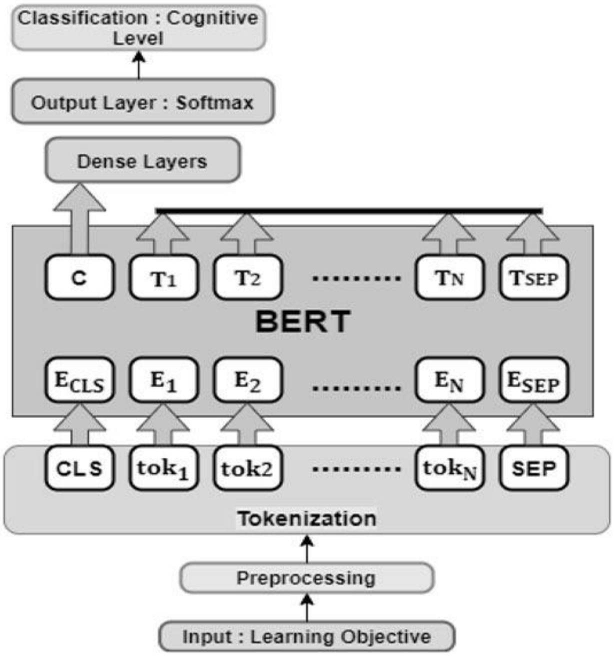

Fig. 4. BERT With Fully Connected Layers Architecture:

Values: *C* – 'At the end' …; T_1,- 'of the module' T_2,-'the learner' …, *'will be able' 'to analyze* …; E_{CLS} – 'The module'…; E_1 , E_2 _'aims to equip the learner', …, E_N – …; E_{SEP} – …; *CLS* –' the module's objective' …; tok_1, tok_2 _ "is to enhance' …, tok_N – "analytical thinking skills'…– …

In previous architectures, the *CLS* output was the only one used as input for the classifier. In this architecture, one can use all the outputs of the last transformer encoder as inputs to an LSTM or Bi-LSTM recurrent neural network as shown in Fig. 5.

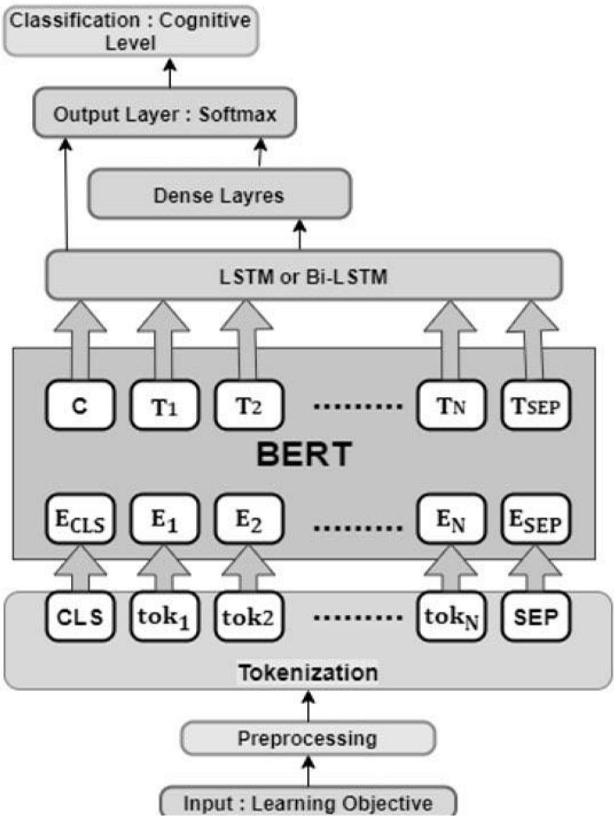

Fig. 5. BERT With Fully Connected Layers and Deep Network Layers Architecture:

Values: *C* – 'At the end' …; T_1,- 'of the module'T_2,-'the learner' …, *'will be able' 'to analyze* …; E_{CLS} – 'The module'…; E_1, E_2 'aims to equip the learner', …, E_N – …; E_{SEP} – …; *CLS* –' the module's objective' …; tok_1, tok_2 – "is to enhance' …, tok_N – "analytical thinking skills'…– …

The next section provides a representation and a detailed analysis of the dataset for experimentation with the different models.

4. Methodology

4.1. Data Analysis

Researchers can start by collecting LOs (Table 1) from the MOOCs providers, Coursera, and edX, and then manually annotate them based on Bloom's taxonomy action verbs list. However, this could lead to ambiguity about the actual meaning of the required cognition [22-27].

Table 1
The Distribution of LOs in Dataset

Cognitive level	2394	Example
Knowledge (remembering)	400	Describe the concept of modular programming and the uses of the function in computer programming
Comprehension (understanding)	400	At the end of this module, the learner will be able to classify clustering algorithms based on the data and cluster requirements
Application (applying)	400	Apply a design process to solve object-oriented design problems
Analysis (analyzing)	400	Analyze the appropriate quantization algorithm
Evaluation (evaluating)	394	Compare the semantic and syntactic ways encapsulation
Synthesis (creating)	400	Create a Docker container in which you will implement a Web application by using a flask in a Linux environment

4.2. Evaluation Metrics

Given a dataset with an approximately balanced number of samples from all classes, one can use the accuracy measure to evaluate the performance of a model and compare it with other models[28].

Accuracy is the sum of true positive (TP) and true negative (TN) items divided by the sum of all other possibilities:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

Where *TP* – True Positives; *TN* – True Negatives; *FN* – False Negatives;*FP* – False Positives.

4.3. Environment Setup

One can use the Google Colab and Tensorflow environment as well as Keras Tensorflow to build the BERT models. Keras TensorFlow is an open-source mathematical software library used for machine learning applications. It has tools to run on graphic processing units, which can significantly reduce training and inference times on some models. Keras is a high-level APi for TensorFlow and has a modular and easily extensible architecture, and it allows users to create sequential models or a graph of modules that can be easily combined.

The library contains many different types of neural layers, cost, and activation functions. We implemented different fine-tuning strategies of BERT on Tensorflow Hub (TFHub). TFHub provides a way to try, test, and reuse machine learning models.

4.4. Implementation Details

For the implementation of the models adopted, one can use the Keras Layer function of Tensorflow Hub to build our BERT layer. Then, one can tokenize text based on the variables of this layer. This would allow having the first input of BERT model, (e.g: input_word_ids). Next, one can add the embeddings of position input_mask and segments segment_ids.

4.5. Discussion

The rise of AI makes it impossible to ignore a serious debate about its future role in our lives. Given complex algorithms designed by programmers that can transmit their own biases an in-depth discussion is critical to promote, and develop knowledge and wisdom.

There is a need for further research on the privacy implications of the current control on developments of AI taking into account creativity, and innovation which can hardly be performed by machines.

5. Conclusions

This study proposes an automatic and large-scale ML-based classification system for MOOCs according to their learning objectives by making use of the six cognitive levels of Bloom's taxonomy. During the course of the research, it is shown that analyzing learning objectives (LOs) associated with modules and programs can further enhance the quality of digital learning system. As a result of the research, a representation and a detailed analysis of the dataset for experimentation with the different models are provided.

References

[1]. Sebbaq, H., &Faddouli, N. E. E. (2021, January). MOOCs semantic interoperability: Towards unified and pedagogically enriched model for building a linked data repository. *International Conference on Digital Technologies and Applications*. Springer, 621-631.

[2]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805. https://arxiv.org/abs/1810.04805

[3]. Abduljabbar, D. A., & Omar, N. (2015). Exam questions classification based on Bloom's taxonomy cognitive level using classifiers combination. *Journal of Theoretical and Applied information Technology*, *78*(3), 447. http://www.jatit.org/volumes/Vol78No3/15Vol78No3.pdf

[4]. Conole, G. (2014, April). The 7Cs of learning design: A new approach to rethinking design practice. *Proceedings of the Ninth international Conference on Networked Learning* (pp. 502-509). Edinburgh, Scotland. https://www.lancaster.ac.uk/fss/organisations/netlc/past/nlc2014/abstracts/pdf/conole.pdf

[5]. Davis, D., Seaton, D., Hauff, C., &Houben, G. J. (2018, June). Toward large-scale learning design: Categorizing course designs in service of supporting learning outcomes. *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 1-10). https://doi.org/10.1145/3231644.3231663

[6]. Kopp, M., & Lackner, E. (2014). Do MOOCs need a special instructional design? *Proceedings of Sixth international Conference on Education and New Learning* (EDULEARN14; pp. 7138-7147). Barcelona, Spain. https://library.iated.org/view/KOPP2014DOM

[7]. Molenda, M. (2003). in search of the elusive ADDiE model. *Performance improvement*, *42*(5), 34-37. http://www.damiantgordon.com/Courses/DT580/in-Search-of-Elusive-ADDiE.pdf

[8]. Osman, A., & Yahya, A. A. (2016). Classifications of exam questions using linguistically-motivated features: A case study based on Bloom's taxonomy. https://www.researchgate.net/publication/298286164_Classifications_of_Exam_Questions_Using_Linguistically-Motivated_Features_A_Case_Study_Based_on_Blooms_Taxonomy

[9]. Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980. https://arxiv.org/abs/1412.6980

[10]. Major, C. H., & Blackmon, S. J. (2016). Massive open online courses: Variations on a new instructional form. *New Directions for institutional Research*, *2015*(167), 11-25. https://doi.org/10.1002/ir.20151

[11]. Das, S., Das Mandal, S. K., &Basu, A. (2020). identification of cognitive learning complexity of assessment questions using multi-class text classification. *Contemporary Educational Technology*, *12*(2), Article ep275. https://doi.org/10.30935/cedtech/8341

[12]. Haris, S. S., & Omar, N. (2012, December). A rule-based approach in Bloom's taxonomy question classification through natural language processing. *Seventh international Conference on Computing and Convergence Technology* (iCCCT; pp. 410-414). institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/abstract/document/6530368

[13]. Conole, G. (2016). MOOCs as disruptive technologies: Strategies for enhancing the learner experience and quality of MOOCs. *Revista de Educación a Distancia*, *50*(2). http://dx.doi.org/10.6018/red/50/2

[14]. González-Carvajal, S., & Garrido-Merchán, E. C. (2020). *Comparing BERT against traditional machine learning text classification*. arXiv preprint arXiv:2005.13012. https://arxiv.org/abs/2005.13012

[15]. Merrill, M. D. (2012). *First principles of instruction*. John Wiley & Sons. https://digitalcommons.usu.edu/usufaculty_monographs/100/

[16]. Margaryan, A., Bianco, M., & Littlejohn, A. (2015). instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77-83. https://doi.org/10.1016/j.compedu.2014.08.005

[17]. Xing, W. (2019). Exploring the influences of MOOC design features on student performance and persistence.*Distance Education, 40*(1), 98-113.https://doi.org/10.1080/01587919.2018.1553560

[18]. Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

[19]. Nevid, J. S., & McClelland, N. (2013). Using action verbs as learning outcomes: Applying Bloom's taxonomy in measuring instructional objectives in introductory psychology. *Journal of Education and Training Studies*, *1*(2), 19-24. https://doi.org/10.11114/jets.v1i2.94

[20]. Omar, N., Haris, S. S., Hassan, R., Arshad, H., Rahmat, M., Zainal, N. F. A., &Zulkifli, R. (2012). Automated analysis of exam questions according to Bloom's taxonomy. *Procedia: Social and Behavioral Sciences*, *59*, 297-303. https://doi.org/10.1016/j.sbspro.2012.09.278

[21]. Mohammed, M., & Omar, N. (2020). Question classification based on Bloom's taxonomy cognitive domain using modified TF-iDF and word2vec. *PLOS ONE 15*, e0230442. https://doi.org/10.1371/journal.pone.0230442

[22]. Grandini, M., Bagli, E., & Visani, G. (2020). *Metrics for multi-class classification: An overview*. arXiv preprint arXiv:2008.05756. https://arxiv.org/abs/2008.05756

[23]. Pardos, Z. A., & Schneider, E. (2013). *AiED 2013 workshops proceedings* (Vol. 1). http://people.cs.pitt.edu/~falakmasir/docs/AiED2013.pdf

[24]. Quintana, R. M., & Tan, Y. (2019). Characterizing MOOC pedagogies: Exploring tools and methods for learning designers and researchers. *Online Learning*, *23*(4), 62-84. https://doi.org/10.24059/olj.v23i4.2084

[25]. Rosselle, M., Caron, P. A., &Heutte, J. (2014, February). A typology and dimensions of a description framework for MOOCs. in *Proceedings of European MOOCs Stakeholders Summit 2014,* (eMOOCs 2014; pp. 130-139). Lausanne, Switzerland. Proceedings document published by Open Education Europa (www. openeducationeuropa. eu). https://hal.archives-ouvertes.fr/hal-00957025/

[26]. Sebbaq, H., El Faddouli, N. E., & Bennani, S. (2020, September). Recommender system to support MOOCs teachers: Framework based on ontology and linked data. *Proceedings of the 13th international Conference on intelligent Systems: Theories and Applications,* Article 18. https://doi.org/10.1145/3419604.3419619

[27]. Ting Fei, Wei Jyh Heng, Kim ChuanToh, & Tian Qi. (2003). Question classification for e-learning by artificial neural network. *Proceedings of the Joint Fourth international Conference on information, Communications and Signal Processing, and the Fourth Pacific Rim Conference on Multimedia* (pp. 1757-1761). institute of Electrical and Electronics Engineers. Singapore. https://doi.org/10.1109/iCiCS.2003.1292768

[28]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., &Polosukhin, i. (2017). Attention is all you need. *31st Conference on Neural information Processing Systems* (NiPS 2017; pp. 5998-6008). Long Beach, USA. https://arxiv.org/abs/1706.03762

[29]. Yousef, A. M. F., Chatti, M. A., Schroeder, U., &Wosnitza, M. (2014, July). What drives a successful MOOC? An empirical examination of criteria to assure design quality of MOOCs. *14th international Conference on Advanced Learning Technologies* (pp. 44-48). institute of Electrical and Electronics Engineers. https://doi.org/10.1109/iCALT.2014.23

[30]. Yusof, N., Hui, C. J. (2010). Determination of Bloom's cognitive level of question items using artificial neural network. *10th international Conference on intelligent Systems Design and Applications* (iSDA; pp. 866-870). institute of Electrical and Electronics Engineers. Cairo, Egypt. https://doi.org/10.1109/iSDA.2010.5687152

Ayse Kok Arslan
Researcher
Silicon Valley
Oxford Northern California Alumni
San Francisco, Northern California, United States
Phone: +16692908776
ORCID number: 0000-0002-5159-4019