SCIENCE CHINA
Information Sciences

• MOOP •

June 2021, Vol. 64 164101:1–164101:3
https://doi.org/10.1007/s11432-019-9939-2

XiaoMu: an AI-driven assistant for MOOCs

Zhengyang SONG[1], Jie TANG[3*], Tracy Xiao LIU[4], Wenjiang ZHENG[2], Lili WU[2], Wenzheng FENG[3] & Jing ZHANG[5]

[1]*Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing* 100084*, China;*
[2]*Center for Strategic Studies, Chinese Academy of Engineering, Beijing* 100088*, China;*
[3]*Department of Computer Science and Technology, Tsinghua University, Beijing* 100084*, China;*
[4]*School of Economics and Management, Tsinghua University, Beijing* 100084*, China;*
[5]*Information School, Renmin University of China, Beijing* 100872*, China*

Received 7 May 2019/Accepted 8 July 2019/Published online 26 April 2021

Citation Song Z Y, Tang J, Liu T X, et al. XiaoMu: an AI-driven assistant for MOOCs. Sci China Inf Sci, 2021, 64(6): 164101, https://doi.org/10.1007/s11432-019-9939-2

Nowadays, massive open online courses (MOOCs), are attracting widespread interest as an alternative education model. Several MOOCs platforms, such as Coursera, edX, and Udacity have been built and they provide low-cost opportunities for anyone who needs to access a massive number of courses from the worldwide top universities. However, the learning completion rate of the learners via MOOCs is extremely low compared with that in real classroom learning. Thus, this study builds a system, called XiaoMu, to design intelligent intervention strategies for the learning process of learners and improve the learning efficiency through automatic analysis of courses and learners. We deployed our system on the website1), which is one of the largest MOOCs in China.

We will introduce the key components of our XiaoMu system in the following paragraphs. First, we use a course concept extraction algorithm to extract important concepts from the scripts of course videos and discover the prerequisite relations between the extracted concepts using a prerequisite relation learning algorithm. Second, an automated video navigation module is used to help the learners easily navigate interesting segments while watching the course videos. Third, a dropout prediction module is used to give learners a kind reminder when they are at the edge of dropping out of the course. Finally, we develop a question answering module to solve the questions that the learners have during the learning process, and a active questioning module to inspire the learners to think more while watching the course videos. Figure 1 illustrates the XiaoMu system components.

Component 1 (Course concept extraction)**.** A course corpus D comprises several courses C in the same field, where each course C contains several videos v, and each v contains a title and a sequence of video scripts. We define a course concept c as a k-gram extracted from the corpus D, such that it has the characteristics of phraseness and informativeness [1]. Using D, we extract the candidate course concepts from D denoted as $T = \{c_1, \ldots, c_M\}$ and output a confidence score s_i for each candidate $c_i \in T$ to indicate the likelihood of c_i to be a course concept.

In dealing with the problem, we first tokenized the titles and scripts of the videos in D and annotated each token by part-of-speech tags. We then matched the annotated tokens using certain linguistic patterns to obtain a set of candidate concepts. We used local mutual frequencies to measure the phraseness and semantic embeddings trained on Wikipedia to measure the informativeness of the extracted candidate concepts. Subsequently, we built a course concept graph by connecting each concept pair whose cosine similarity of embeddings is above a predefined threshold. A novel iterative graph-based propagation algorithm was applied on the concept graph to rank the concepts. The top ranked concepts were then retrieved as the important concepts. The experiments on the two-course data sets showed that our method significantly outperformed all the alternative methods (+0.013 – 0.318 in terms of R-precision; $p \ll 0.01$, t-test).

Component 2 (Prerequisite relation learning)**.** We need to further discover the relations between these concepts, where one of the most important types is the prerequisite relationship, to comprehensively understand the knowledge delivered by the concepts.

Given a corpus D and its corresponding course concepts T, the objective of the prerequisite relation learning was to learn a function $P : T^2 \to \{0, 1\}$ that maps a concept pair $\langle a, b\rangle \in T^2$ to a binary value that indicates whether a is a prerequisite concept of b.

Herein, we defined several features from different aspects, and the most important one is the average position distance between two concepts when they appear in the same video [2]. We adopted a random forest as the classification model. The experiments showed that the proposed method achieves significant improvements (+5.9 by F1-score) compared with the alternative methods.

* Corresponding author (email: jietang@tsinghua.edu.cn)

1) XuetangX.com.

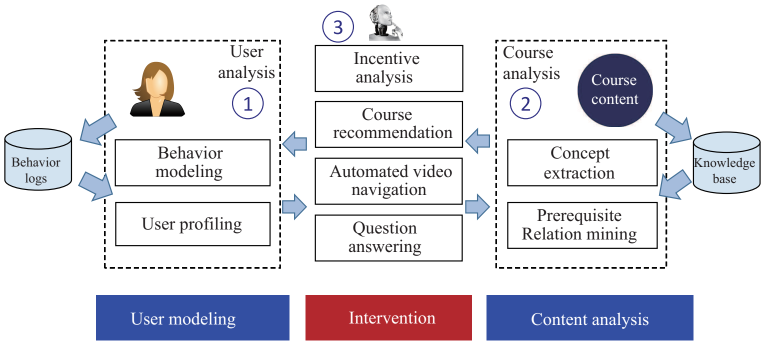

Figure 1 (Color online) Architecture of the MOOC AI system XiaoMu.

Component 3 (Automated video navigation)**.** When a user watches a video on MOOCs, jumping back and forth on the navigation bar to find certain segment positions of his/her interest is common. Thus, we define the video navigation prediction task as follows. Given a video v, a user u, and the start position in the i-th segment s_i, the objective is to train a model to maximize the probability that user u would jump back to the j-th segment s_j of video v.

We tried to solve the problem by first adopting the deterministic finite automaton to determine the common user navigation patterns in the navigation behaviors [3]. Further, we partitioned a video segment into a sequence of finite candidate navigation positions. We then defined the features from the aspects of the video content, user preference, and historical navigation behaviors. In our experiments, the factorization machine was proven as the best model with 0.73 AUC for non-science courses and 0.74 AUC for science courses.

Component 4 (Dropout prediction)**.** A large portion of learners would dropout from a course along the whole semester; thus, identifying possible dropouts and giving them incentive to continue their studies would be useful.

Given user u's learning activities $X(u,c)$ on course c and the context information $Z(u,c)$, such as u's demographics and c's category before time t, our goal herein was to predict whether u will drop out from c after t (i.e., we would like to learn function $f:(X(u,c),Z(u,c))\to y\in\{0,1\}$).

We designed a neural network model to predict the users' dropout behaviors [4]. In particular, we first extracted several statistic features, such as the watching times and homework submission times from $X(u,c)$ and $Z(u,c)$, as the model input. We then applied an embedding layer and a convolutional layer on the input to obtain an intermediate feature representation. An additional attention mechanism was used to capture the importance of the intermediate features from different learning activities.

We did a two-week intervention experiment, i.e., pop an encourage tip when the learner tends to drop out. Improvements are observed in terms of video watching time, answer correction ratio, etc. For dropout reasons, correlation analysis shows a positive correlation between the dropout probability and the number of dropout friends.

Component 5 (Question answering)**.** Learners usually have questions when they are watching videos; however, the existing online discussion forum does not provide an efficient method of finding the correct answers. Therefore, we designed an automatic question answering mechanism for learners to discover answers by themselves.

We designed different question answering mechanisms responsible for different types of user questions. Given the user's current question q and the set of candidate answer mechanisms $A=\{a_1,\ldots,a_M\}$ with each a_m as an independent answer mechanism, the target was to learn a function that can choose the best answer mechanism a^* from A that can best fit q (i.e., $a^*=f(q,A)$).

We summarized the user intents into platform usage, concept explanation, course information, and small chat according to the query, and implemented the corresponding answer mechanism for each type of intent. Given a question from a user, we first classified it into a type of intent using a trained support vector machine (SVM) model then matched it to the related question-answer pairs from the corresponding answer mechanism of the intent. The SVM model was trained based on the labeled user questions collected from the website[1)]. The question-answer pairs were collected from the website[2)] and CSDN. The F1-score of our user intent classification algorithm was approximately 0.74.

Component 6 (Active questioning)**.** Asking some questions to learners when they are watching the videos would help draw their attention and guide them to think deeper and wider. Our system would provide the corresponding answer if the user click on the presented question.

For a given video and a user, we prepared a set of manually edited candidate questions and determined the time point to present each question. We also treated the click action of a user as the positive feedback/reward to the presented question. The task aimed to determine the question to be presented that can result in a maximal cumulative reward.

We modeled this problem in a multi-armed bandit learning framework with implicit feedback [5]. The candidate questions were regarded as bandit arms, and the user clicks were regarded as rewards. Some mathematical techniques such as Thompson sampling and variational Bayesian inference are used for arm selection and model update. The empirical evaluations on collected click logs demonstrated the learning effectiveness in practice.

Summary. This study introduced several components of the XiaoMu system built on the website[1)], which has been stably running online for more than a year. Through this system, we have helped learners in multiple ways. In half of the cases, we can reduce the time users spend navigating the video to one click. We intervened with the learners having a high dropout probability to increase their video watching time, number of completed assignments, and the ratio of correct answers. We also constructed a knowledge graph based on the prerequisite relations between the course concepts extracted from the course video subscripts. From

2) Zhihu.com.

April 2017 to May 2019, the number of user queries from the course concepts, active questioning and question answering module were 85043, 17965, and 30488, respectively, which covers 101025 users from 2925 courses.

Further improvements can be made. For example, we can perform some deduction or induction on the knowledge graph built from the course concepts to answer questions more complicated than simple concept explanations. Furthermore, we can do a recommendation of papers, scholars [6, 7], or videos [8] based on the knowledge concepts contained in the current course video.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2018YFB1004503) and National Natural Science Foundation of China (Grant No. 61702522).

Supporting information Videos and other supplemental documents. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Pan L M, Wang X C, Li C J, et al. Course concept extraction in moocs via embedding-based graph propagation. In: Proceedings of the 8th International Joint Conference on Natural Language Processing, 2017

2 Pan L M, Li C J, Li J Z, et al. Prerequisite relation learning for concepts in MOOCs. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017. 1447–1456

3 Zhang H, Sun M S, Wang X C, et al. Smart jump: automated navigation suggestion for videos in MOOCs. In: Proceedings of International World Wide Web Conference Committee, 2017. 331–339

4 Feng W, Tang J, Liu T X. Understanding dropouts in MOOCs. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019

5 Qi Y, Wu Q Y, Wang H N, et al. Bandit learning with implicit feedback. In: Proceedings of the 32nd Conference on Neural Information Processing System, 2018. 7276–7286

6 Wu K, Tang J, Shao Z, et al. CareerMap: visualizing career trajectory. Sci China Inf Sci, 2018, 61: 104101

7 Tang J, Zhang J, Yao L M, et al. Arnetminer: extraction and mining of academic social networks. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008. 990–998

8 Tan X Y, Guo Y C, Chen Y S, et al. Accurate inference of user popularity preference in a large-scale online video streaming system. Sci China Inf Sci, 2018, 61: 018101