Are the Dropout Models in MOOC General: Perspective from Big Data

*

***Abstract*—Online education has become the first choice for millions of internet users around the world to get access to knowledge and skills. With online courses such as Massive Online Open Courses(MOOCs) flourishing, the problem of high dropout rate emerges. Many researches aim to implement dropout prediction to mark learners with high risk of dropout in the early stage of a course. Among these researches, the employment of features concerned with learning behaviors combined with machine learning techniques proves effective to obtain decent prediction results. Therefore, this method is trending in recent years. However, the majority of existing researches only test prediction based on one or two courses, making people wonder whether there could be a universally applicable prediction model for different types of courses. In this paper, we propose a study on performance of different prediction models with the perspective of big data. We collect the background data from 20 representative courses in different subjects on a popular MOOC platform in China, and test the effectiveness of the existing predict models. The results show that there does not exist a universally feasible prediction method that can work for all the courses. A future research line could focus on designing adaptive models that can work in different courses with different students, etc.**

***Index Terms*—online courses, performance, prediction, big data, machine learning**

I. Introduction

With more than ten years of development, online education has become an important way for internet users around the world to learn knowledge and skills. Among them, the online courses, represented by massive open online courses (MOOC), have gained the attention and use of hundreds of millions of learners [1]. MOOC refers to an open course designed to provide educational content to a large number of participants through online platforms and provide free access. Since the concept was put forward in 2008, MOOC has evolved various forms from the initial college courses for undergraduate students, and the target population has also been expanding. However, while MOOC appeals to a huge amount of learners, it also comes with a higher dropout rate which could be up to 80-95% [2].To deal with this problem, the researchers have done a lot of work based on the data of online learning. These work aim to find the possible factors that lead to the learner's choice to dropout, predict whether the learner may withdraw from the course, and to intervene as soon as possible according to the results of prediction.

Some researchers try to establish the connection between learner's dropout with the online learning platforms they use, arguing that experience of platforms using could influence the interest and patience of learners, leading to withdrawing from learning. Espada et al. [3] quantitatively analyze the user experience of several MOOC platforms from aspects of content quality, kinds of materials, and technical features. Comparison across different platforms has been made, presenting the difference in accessibility, web-page loading, page-and-content search, etc. In addition, some researchers also focus on the quality of courses and resources, pointing out the poor design of course materials and problematic way of teaching could have a negative influence on students' will of continuing learning. Hemberg et al. [4] investigate the relation between learning design and learner behavior in a computational thinking MOOC, categorize MOOC resources with a learning design lens. Through classification of resources like videos, practices and problem sets, they prove that variations of learner behavior are related to different resource types, and a fine arrangement of course resources could lead to lower probability of dropout for learners.

Besides the insights of environment and resource impact mentioned above, the activities of a learner implicate his or her ending, i.e. whether he or she chooses to continue learning or withdraw in half. Therefore, a huge amount of work based on data of learner behavior, making it the most heated direction of learner dropout prediction in recent years. Various features have been extracted from the learning behavior data. These features are categorized as demographic, video-related, exercise-related, forum-related and platform use variables in a generalized fashion [5]. Early researches focus on one of these types of features [6] [7], and later multiple types of behavior-related features are combined to establish the learning patterns of learners, based on which they could make more precise predictions [8] [9]. Despite the differences in types of variables employing in feature selection, more and more researchers agree that machine learning methods could obtain decent results when confronted with prediction problems [10]. The practice of establishing a predictive model using Logistic Regression, Decision Tree and Support Vector Machine and other machine learning tools is trending among researches in recent years.

With huge amount of researches on dropout prediction from different insights, almost half of the articles tested predictive models in only one MOOC course. This is a limitation of the results since it is not possible to determine how general the predictive models are. Motivating by answering this question, we investigate the performance of different prediction models with the perspective of big data. In our study, we collect the background data from a popular MOOC platform in China, including platform use activities, video events and logs related to the exercises; and forum records, which provide relevant information not only about learner's attitude or sentiment, but also about social interactions. We select 20 representative courses in different subject (engineering, literature and so on), and study the effectiveness of the existing predict models. The results show that there does not exist a universally feasible prediction method that can work for all the courses. In other words, the existing model is sensitive to the activities styles of learners. A future research line could focus on designing adaptive models that can work in different courses, with different students, etc.

II. Methodology

In this section, we will explain the methodology of our experiment. We would like to address the following four questions in our research:

RQ1: Which type of features could lead to most accurate results in learners' performance prediction in online courses?

RQ2: Which prediction method is the most effective method in learners' performance prediction in online courses?

RQ3: Is it feasible to apply learning data in previous semesters to predict learners' performance in the latest semester?

RQ4: Is there a generally applicable model for performance prediction in different courses?

To answer these questions, we carry out the following experiment. First, we will describe the source of our data. Then, we will proceed to the part of data pre-processing and feature extraction. After that, we will move to experiment design, and explain our experiment step by step.

A. Dataset

The data used in our research is obtained from XueTangX online learnin platform. As a partner of edX, XueTangX was launched in 2013 and has become one of the largest online courses providers in China. We selected our data carefully among over 2,000 semesters of raw log data from this platform based on number of learners as well as ratio of certificate earners for each semester. We finally determined 33 courses and 51 semesters from different departments and categories, covering a wide selection of Art, Science, Engineering, Business, and Medicine. These courses are consecutively opened from 2016 to 2018 divided in a series of semesters.

We extract our data from forum activities, records of exercises and video watching logs of learners registered in these courses and semesters, as well as their final performance, i.e. whether they obtained the certificates or not. In total, there are around 15,290,000 time-stamped activity logs and the total size of data is 3.5 Gigabytes.

B. Labels

Our aim is to predict learners' learning performance during different courses, using features that could well reflect their behavioral patterns, and methods proved effective in learner performance prediction, and determine whether there exists a universally applicable model for different courses. In order to achieve this goal, we first needed to decide what kind of result is the one we would like to predict.

The courses on XueTangX provide learners with certificates to prove that they have successfully passed the courses they learned, and records of certificates are conserved in log data. However, during our investigation of XueTangX, we found that the certificate information may not precisely reflect the efforts and achievements of learners, especially their performance in the learning process. XueTangX has an automatic grading system for learners. The final grades of learners were from the accumulation of their usual learning activities. For instance, if they finish one exercise, the grades of this exercise will immediately be calculated into the final grades. If they finished watching a video or learning a tutorial, they will also obtain marks accordingly. This means that the learning activities of learners could directly affect their final grades and thus certificates earning. This influence is conveyed through calculating mechanism rather than learners using knowledge they acquired in the courses and applying them in a final exercise or exam, diminishing the learners' own efforts and intelligence while perceiving knowledge from online courses. Due to the direct causal relationships of learning activities and final grades, we came to the conclusion that certificate earning could not be considered as a suitable prediction label.

We turned to learners' grades in their exercises. As a course goes, the activities of taking exercises are parallel to watching videos and sending posts. In this case, the grades of exercises could solely reflect learners' efforts and knowledge states. For each course, we calculate the full marks of all exercises and the total grades that each learner earned in all exercises. If the total grade of a learner could exceed the 60 percent of full marks, we consider this learner as 'pass', otherwise 'fail'. By this means, we acquired the labels we wanted in performance prediction of online courses.

C. Features and Data Pre-Processing

Based on existing research and the situation of our data, we decide the following 9 features should be extracted and used for prediction. These features are categorized under different activity types: *forum*, *exercise* and *video*. Table I exhibits these features and their categories.

TABLE I: Features Selected in Experiment

Forum	total_posts main_posts comments avg_length
Videos	per_open per_vtotal per_compl avg_rep avg_pause

1) Forum:

total_posts: The number of posts that a learner posted in the forum in one course, whether the type of post is main threads or comments.

main_posts: The number of main posts that a learner posted in the forum in one course.

comments: The number of comments that a learner posted under main posts in the forum in one course.

avg_length: The average length of post content that a learner posted in the forum in one course.

2) Video:

per_open: The percentage of videos that a learner opened to all videos in one course.

per_vtotal: The percentage of a learner's video watching time to total video time in one course.

per_vcompl:The percentage of videos that a learner completed watching to all videos in one course.

avg_rep: The average repetitions of a video in all videos that a learner watched.

avg_pause: The average pauses of a video in all videos that a learner watched.

These features are obtained from different log data. It is notable that the learners recorded in these data are not exactly the same. Video log data commonly has the most learners, while the learners recorded in forum activity log data are fewer. This phenomenon does have a relatively reasonable explanation. A lot of learners choose to join one online course simply for knowledge and skills taught in the videos rather than course certificates. Therefore, they are very likely to only watch videos, meanwhile complete no exercises and send no posts in the forum. In this case, there is no wonder that they only leave their marks in video logs but vanish in exercise and forum data.

To solve this, we tried to explore the meaning of features. All the features we extracted are statistically calculated from log data, based on numbers of videos watched, exercises taken, posts sent in forum, and so on. If there is no record in one activity, the concerning data would be 0, thus mathematically the corresponding features calculated from these data would be set as 0 after extraction. Therefore, we decided to use the value 0 to make up the missing features of learners who are only active in one or two activities among forum, exercise and video. Through this method, we complete the features for each learner in each course.

D. Experiment Design

Our aims could be further specified from research questions as follows:

1) Determine which type of features performs best in learner performance prediction.

2) Determine which prediction method performs best in learner performance prediction.

3) Determine which semester's data performs best in learner performance prediction.

4) Investigate whether there exists a universally applicable best-performed combination of feature type, prediction method and semester data among different courses.

To achieve these aims, we decided to break down our experiment into four minor experiments in a series.

1) Experiment 1:

Under the circumstance of one semester in one course, compare the prediction performance among different feature types, and determine the best-performed feature type.

There are four types of features that we extracted and concluded from our data: *forum*, *video*, and *mix* which represents the combination of former two feature types. For each feature type, we use six prediction methods to carry out experiments: Decision Tree(DT), Random Forest(RF), Supporting Vector Machine(SVM), Logistic Regression(LR), Neural Network(MLP) and XGBoost(XGB). Therefore, we have {*forum*, *video*, *mix*} x {DT, RF, SVM, LR, MLP, XGB}, i.e. 18 sets of experiments in total for one semester in one course for one epoch. In order to minimize the randomness in experiment results, we chose to repeat each experiment ten times and consider the average prediction outcome as final result, thus in fact 180 sets of experiments for one semester in one course.

2) Experiment 2:

Under the circumstance of one semester in one course, compare the prediction performance across different prediction methods, and determine the best-performed method. Based on the existing researches, we selected four prediction methods : Decision Tree(DT), Random Forest(RF), Supporting Vector Machine(SVM), Logistic Regression(LR) and Neural Network(MLP). These methods are picked because they are the most common methods used in MOOC performance prediction research, and proved their effectiveness in different work.

The experiment of prediction methods could be combined with the first experiment on feature types. While we tested the effectiveness of feature types in course performance prediction, the same results could be applied to analyze the effectiveness of different prediction methods at the same time. Therefore, for one semester in one course, we also have 180 sets of experiments in total.

To better evaluate the performance of each feature type and prediction method in experiment 1 and 2, further processing of results was necessary. For experiment 1, we chose to calculate the average performance of four prediction methods for each single feature type within a semester's data. For experiment 2, we chose the similar measure to calculate the average performance of four feature types for each single prediction method within a semester.

3) Experiment 3:

With best-performed feature type and prediction method, use the data of previous semesters in one course to predict the performance of learners in latest semester. In order to make comparison comprehensive, aside from single semester, we decided to add new collections of last two semesters, last three semesters and so on. Among these collections of semesters, we compare their effectiveness in performance prediction, and determine which collection is best-performed.

The number of experiments for one course depends on the number of semesters. Take a course with four semesters as an

TABLE II: Comparison of Different Feature Types

Category	Course	Feature Type		
		forum	video	mix
arts	course-v1:BIFT+1301990078	0.5	**0.514585533**	0.509232626
	course-v1:JNUX+jnu2017002	0.498937952	0.543531225	**0.544853884**
	course-v1:TsinghuaX+00690242X	0.581640466	0.56504603	**0.589877108**
	course-v1:TsinghuaX+40670453X	0.662539215	0.648352738	**0.704634972**
	course-v1:TsinghuaX+00670122X	0.5577559	0.529474488	**0.590230973**
language	course-v1:TsinghuaX+20220332X	0.506417202	0.587058166	**0.60175041**
	course-v1:TsinghuaX+60250101X	**0.532849585**	0.51320103	0.520010671
	course-v1:TsinghuaX+AP000008X	0.533743031	**0.625699604**	0.612312342
	course-v1:TsinghuaX+30640014X	0.717773293	0.644851817	**0.752674669**
science	course-v1:CQCET+20180102	0.49995096	0.507811041	**0.510768373**
	TsinghuaX/10450012X/2015_T2	0.50035638	0.511229232	**0.532411693**
engineering	course-v1:TsinghuaX+00740123X	0.53411384	0.54748811	**0.570662994**
	course-v1:TsinghuaX+30240233X	0.550851267	0.522578768	**0.577175318**
	course-v1:SWPU+3615001035	0.49852968	0.537507594	**0.557806383**
	course-v1:QHU+QHU2017072701X	0.586525533	0.584834492	**0.613406521**
business	course-v1:TsinghuaX+JRFX01	0.513943156	0.521502535	**0.531705182**
	course-v1:TsinghuaX+80512073X	0.576245825	0.556523676	**0.59790055**
	course-v1:CIE+CIE2016004	0.515936036	**0.516324088**	0.499217187
	course-v1:CUPB+100722C004	0.572194947	0.507062219	**0.605906429**
	course-v1:MITx+15_390x_2015_T2	**0.567673422**	0.510497377	0.560854844
	course-v1:TsinghuaX+80511503X	0.539563811	0.592896626	**0.616227772**
medical	course-v1:BSU+70119052	0.511577305	0.560542002	**0.566035059**
	course-v1:TsinghuaX+400182X	0.631411226	0.595205342	**0.675745367**
	course-v1:CDUTCM+030687	0.496158145	0.565409911	**0.596465855**
	course-v1:TsinghuaX+34000888X	0.598556743	0.569128195	**0.653590394**
social science	course-v1:HIT+13GH16000700	0.579491629	0.672679898	**0.708025018**
	course-v1:TsinghuaX+30700313X	0.595473527	0.545716885	**0.620679862**
	course-v1:TsinghuaX+2018031901X	0.522299886	0.516803452	**0.530772164**

example. We label the semesters as Term 1, Term 2, Term 3 and Term 4. The data from Term 4 is used for test dataset. For the previous three semesters, the collection of training datasets is listed as {{Term 1}, {Term 2}, {Term 3}, {Term 2, Term 3}, {Term 1, Term 2, Term 3}}. Still, we repeat three times for each experiment. Therefore, for a course with four semesters, there are 15 sets of experiments in total.

4) Experiment 4:

After finishing the previous three steps, compare the effectiveness of prediction across different courses, and determine whether there exists a common pattern. This pattern could be summarized as a best-performed combination of feature type, prediction method and collection of previous semesters.

In this part, we not only put an effort to decide whether there exist a commonly applicable pattern or model among different courses, but also experimentally cross-predict the learners' performance in different courses, i.e. use one course's data to predict another course's performance. By this measure, we would like to have a glimpse on the probability that different courses share a common potential feature, which could derive from their commonness of online courses beneath their differences on the surface.

III. Results and Analysis

In this section, we will demonstrate the results of our experiments, and carry out analysis based on them. We employ Area Under Curve(AUC) to evaluate the effectiveness of different research factors, i.e. feature type, prediction method and collection of semesters.

A. RQ1: Which type of features could lead to most accurate results in learners' performance prediction in online courses?

The result of first experiment is shown in Table II. It is clear that in most courses, *mix* outperformed *forum* and *video*. This could be explained that since *mix* is a combination of the two former single features, it has a more complex scope of learners' behaviors during the learning period, and could better reflect the pattern of learners who are only active in one activity of them. This leads to its higher accuracy and robustness compared with other two single features. Therefore, we determine that *mix* is the best-performed feature type in this experiment.

B. RQ2: Which prediction method is the most effective method in learners' performance prediction in online courses?

The result of second experiment is showed in Table III. As is shown in the table, it is clear that XGB obtains best performance in 15 courses out of 28 in total, presents an advantage of efficiency as well as stability over other prediction methods. DT, RF and LR follow as they each scores best performance in 5, 4 and 4 courses. SVM and MLP perform worst. Both of them achieve best result in none of these courses. It is notable that within the scope of each course, the difference of performance among all methods are very narrow in most cases, meaning they could all have close efficiency when it comes to predict learners' final performance in online courses. Still, based on the actual performance, we determine that XGB is best-performed in this experiment.

For experiment 1 and 2, when we utilize *mix* features and XGB method to predict learners' performance in these courses,

TABLE III: Comparison of Different Prediction Methods

Category	Course	Prediction Method					
		DT	RF	SVM	LR	MLP	XGB
arts	course-v1:BIFT+1301990078	**0.522779677**	0.50325167	0.5	0.5	0.5	0.521604971
	course-v1:JNUX+jnu2017002	0.561685105	**0.562586558**	0.500330173	0.505810266	0.5	0.54423402
	course-v1:TsinghuaX+00690242X	0.571566063	0.578949286	0.577252465	0.580010045	0.575592658	**0.589756695**
	course-v1:TsinghuaX+40670453X	0.64689182	0.669392419	0.675583967	**0.68638235**	0.678602535	0.674200758
	course-v1:TsinghuaX+00670122X	0.561435347	0.562361717	0.551540039	**0.570555321**	0.538332718	0.570697578
language	course-v1:TsinghuaX+20220332X	0.597212163	0.588978896	0.506711275	0.573094211	0.524896993	**0.599558015**
	course-v1:TsinghuaX+60250101X	0.540863167	**0.545931955**	0.500949445	0.499434294	0.5	0.544943709
	course-v1:TsinghuaX+AP000008X	0.681750618	0.665739401	0.507152876	0.506855946	0.5	**0.682011111**
	course-v1:TsinghuaX+30640014X	0.706181281	0.71963739	0.668993429	0.688068179	0.709708386	**0.738010892**
science	course-v1:CQCET+20180102	0.514848372	0.50538283	0.5	0.5	0.5	**0.516829547**
	TsinghuaX/10450012X/2015_T2	0.51787587	**0.542323347**	0.5	0.5	0.5	0.527795393
engineering	course-v1:TsinghuaX+00740123X	0.596026003	0.584580602	0.503116494	0.511349784	0.502801724	**0.60665528**
	course-v1:TsinghuaX+30240233X	**0.569302122**	0.551571712	0.551381347	0.550546329	0.517734914	0.560674282
	course-v1:SWPU+3615001035	0.538481845	0.533569907	0.515083524	**0.557969493**	0.505176158	0.537406389
	course-v1:QHU+QHU2017072701X	0.573906103	0.58703569	0.605918368	**0.607048695**	0.602405917	0.593218319
business	course-v1:TsinghuaX+JRFX01	0.544826195	0.545520422	0.499655388	0.497939764	0.5	**0.546359974**
	course-v1:TsinghuaX+80512073X	0.574592131	0.578045097	0.572871584	0.57921997	0.57551758	**0.581093737**
	course-v1:CIE+CIE2016004	**0.534862573**	0.516899815	0.5	0.49974359	0.5	0.511448644
	course-v1:CUPB+100722C004	0.616976877	**0.624521705**	0.506125356	0.50329539	0.503717949	0.615689912
	course-v1:MITx+15_390x_2015_T2	0.570372607	0.568051562	0.520498863	0.542658802	0.5	**0.576469452**
	course-v1:TsinghuaX+80511503X	0.569974533	0.588898488	0.579273358	0.583010936	0.578152392	**0.598066712**
medical	course-v1:BSU+70119052	0.587207967	0.584389212	0.498795181	0.51606404	0.5	**0.589852331**
	course-v1:TsinghuaX+400182X	**0.71355835**	0.677283805	0.587294065	0.616311378	0.5	0.710276273
	course-v1:CDUTCM+030687	0.594552491	0.598749322	0.49962963	0.523160169	0.5	**0.599976208**
	course-v1:TsinghuaX+34000888X	0.618686908	0.620243189	0.599370896	0.606248682	0.562884942	**0.635116047**
social science	course-v1:HIT+13GH16000700	**0.777091565**	0.746712212	0.552872011	0.547664794	0.550298366	0.745754143
	course-v1:TsinghuaX+30700313X	0.597812429	0.597965471	0.576268023	0.571158359	0.572185064	**0.608351202**
	course-v1:TsinghuaX+2018031901X	0.534598376	0.529804139	0.509822686	0.519188189	0.5	**0.546337613**

which is the best-performed combination as shown in the results, 17 out of 28 courses present AUC over 0.60. Given that our research have not fine-tuned each algorithm every time in order to guarantee the fairness across different prediction methods and feature types, this result shows a well prospect of predict learners' performance during online courses using data concerning with their learning behaviors.

C. RQ3: Is it feasible to apply learning data in previous semesters to predict learners' performance in the latest semester?

The result of third experiment is showed in Table IV. From the feasibility perspective, utilizing data of previous semesters to predict learners' performance of latest semester should not be determined as useless, since the assessment by AUC could reach over 0.55 in all courses' best cases. Throughout IV, using data of a single previous semester is best-performed in most courses. In a total of 9 courses, only 1 course achieve their best with combination of multiple historic semesters' data, while the rest all obtained best performance while employing the data from a single historic semester. However, there is no clue for other general conclusions as the results show great randomness. The performance of prediction could achieve near 100 percent, meanwhile drop down to around 0.55 in different courses. This means the effectiveness of historic semester prediction varies a lot in different courses, implicating that the metrics of each semester may not be consistent within one course. In addition, courses under one category cannot be determined to share a common pattern when it comes to best semester collection for prediction. For courses under 'arts', they seem to achieve their best prediction result while utilizing the last semester. However, such pattern cannot be summarized in other categories. Overall, great volatility shows across all courses from prediction results to best semester collection, and we could declare that there is no universally applicable solution for different courses in historic semester prediction.

D. RQ4: Is there a generally applicable model for performance prediction in different courses?

We finally dig a little in this experimental direction. Utilizing data of one course to predict learners' performance of another course is able to achieve around 0.70 or even more in AUC, meanwhile could drop down to below 0.50. Since the randomness that the result shows, we could only conclude that for certain combination, the reason that they achieve higher AUC is simply that they have larger datasets, which promote their accuracy. In this case, the uniqueness of different courses clearly overwhelms the potential commonness among them.

IV. CONCLUSIONS

In this paper, we investigated whether there exists a universally applicable prediction process of learners' final performance in online courses. We observe from three perspectives: feature types, prediction methods and historic semesters. Also, we had a go on cross-prediction across different courses. Our conclusions could be summarized as follows:

- From the perspective of feature type, *mix* proved to be best-performed in most courses.

TABLE IV: Prediction Using Historic Semesters

Category	Course	Number of Semesters	Best Semester Collection	AUC
arts	course-v1:JNUX+jnu2017002	2	\	0.569317956
	course-v1:TsinghuaX+00670122X	2	\	0.643306452
	course-v1:TsinghuaX+00690242X	3	1 from latest semester	0.627041311
	course-v1:TsinghuaX+30640014X	3	1 from latest semester	0.807710036
business	course-v1:TsinghuaX+80512073X	4	2 from latest semester	0.556817579
	course-v1:TsinghuaX+JRFX01	3	last 2 semesters	0.553571429
medicine	course-v1:TsinghuaX+34000888X	3	2 from latest semester	0.761904762
engineering	course-v1:TsinghuaX+00740043X	4	2 from latest semester	0.632537408
social science	course-v1:HIT+13GH16000700	2	\	0.997572816

TABLE V: Cross-Prediction among Different Courses

Performance	Training Course	Test Course	AUC
Top 5	course-v1:QHU+20171225+2018_T1	course-v1:TsinghuaX+10610183_2X+2017_T2	0.738686841
	course-v1:QHU+20171225+2018_T1	course-v1:LUIBE+2018030701+2018_T2	0.71673601
	course-v1:SCUT+144001+2018_T2	course-v1:TsinghuaX+10610183_2X+2017_T2	0.713978689
	course-v1:LUIBE+2018030701+2018_T2	course-v1:TsinghuaX+10610224X+2017_T2	0.698373237
	course-v1:LUIBE+2018030701+2018_T2	course-v1:TsinghuaX+10610183_2X+2017_T2	0.698259781
Bottom 5	course-v1:GIT+1400000003+2017_T2	course-v1:ChongqingUX+CHEM10000X+2018_T2	0.488353268
	course-v1:SCUT+144001+2018_T2	course-v1:GIT+1400000003+2017_T2	0.487058642
	course-v1:TsinghuaX+10610183_2X+2017_T2	course-v1:GIT+1400000003+2017_T2	0.457797227
	course-v1:GIT+1400000003+2017_T2	course-v1:TsinghuaX+10610183_2X+2017_T2	0.433950388
	course-v1:GIT+1400000003+2017_T2	course-v1:HEBUT+20180207001+2018_T3	0.416147008

- From the perspective of prediction methods, XGBoost outperforms other prediction methods.
- From the perspective of historic semesters, there is no clear pattern among different courses.
- As for the cross-prediction experiment, there is no evidence that one course could have a significant impact on learners' performance prediction in other courses.

In summary, a universally feasible prediction method that can work for all the courses does not exist. The uniqueness of different courses overwhelms the potential commonness, and existing models are sensitive to the learners' activities during the learning process. Still, we could draw the conclusion that within the scope of prediction for a single course, the combination of different activities data is a wise and robust choice, and XGBoost is more suitable in this work than other prediction methods in most cases. Our future work could base on these discoveries, and focus on adaptive models that can work in different courses, with different types of learners in online learning environment. For historic semester prediction, since current results haven't enlightened us with a clear common pattern, we attempt to turn to other thinking. For example, we could choose two courses in which majority of learners overlap, and observe whether current prediction combinations are portable in this case. Also, it is important to extend our conclusions to other platforms with different characteristics, and see whether they still suit under other circumstances.

References

[1] D. Shah, By The Numbers: MOOCs in 2018, 2018.

[2] C. G. Brinton, M. Chiang, ”Social learning networks: A brief survey”, 2014 48th Annu. Conf. Inf. Sci. Syst., pp. 1-6, 2014.

[3] J. P. Espada, C. C. Rodríguez, V. García-Díaz and R. González Crespo, ”Method for analysing the user experience in MOOC platforms,” 2014 International Symposium on Computers in Education (SIIE), Logrono, 2014, pp. 157-162.

[4] E. Hemberg, S. Biswas, A. Bajwa, N. Law and U. O'Reilly, ”Categorizing Resources and Learners for a Finer-Grained Analysis of MOOC Viewing & Doing,” 2019 IEEE Learning With MOOCS (LWMOOCS), Milwaukee, WI, USA, 2019, pp. 116-121.

[5] P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Muñoz-Merino and C. D. Kloos, ”Prediction in MOOCs: A Review and Future Research Directions,” in IEEE Transactions on Learning Technologies, vol. 12, no. 3, pp. 384-401, 1 July-Sept. 2019.

[6] C. G. Brinton, S. Buccapatnam, M. Chiang, H. V. Poor, ”Mining MOOC clickstreams: Video-watching behavior vs. in-video quiz performance”, IEEE Trans. Signal Process., vol. 64, no. 14, pp. 3677-3692, Jul. 2016.

[7] M. Klüsener, A. Fortenbacher, ”Predicting students' success based on forum activities in MOOCs”, Proc. Int. Conf. Intell. Data Acquisition Adv. Comput. Syst.: Technol. Appl., vol. 2, pp. 925-928, Sep. 2015.

[8] B. Hong, Z. Wei and Y. Yang, ”Discovering learning behavior patterns to predict dropout in MOOC,” 2017 12th International Conference on Computer Science and Education (ICCSE), Houston, TX, 2017, pp. 700-704.

[9] L. Qiu, Y. Liu and Y. Liu, ”An Integrated Framework With Feature Selection for Dropout Prediction in Massive Open Online Courses,” in IEEE Access, vol. 6, pp. 71474-71484, 2018.

[10] F. Dalipi, A. S. Imran and Z. Kastrati, ”MOOC dropout prediction using machine learning techniques: Review and research challenges,” 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, 2018, pp. 1007-1014.