Research Article

Yamini Goel and Rinkaj Goyal*

On the Effectiveness of Self-Training in MOOC Dropout Prediction

https://doi.org/10.1515/comp-2020-0153
Received Apr 23, 2020; accepted Jun 03, 2020

Abstract: Massive open online courses (MOOCs) have gained enormous popularity in recent years and have attracted learners worldwide. However, MOOCs face a crucial challenge in the high dropout rate, which varies between 91%-93%. An interplay between different learning analytics strategies and MOOCs have emerged as a research area to reduce dropout rate. Most existing studies use click-stream features as engagement patterns to predict at-risk students. However, this study uses a combination of click-stream features and the influence of the learner's friends based on their demographics to identify potential dropouts. Existing predictive models are based on supervised learning techniques that require the bulk of hand-labelled data to train models. In practice, however, scarcity of massive labelled data makes training difficult. Therefore, this study uses *self-training*, a semi-supervised learning model, to develop predictive models. Experimental results on a public data set demonstrate that semi-supervised models attain comparable results to state-of-the-art approaches, while also having the flexibility of utilizing a small quantity of labelled data. This study deploys seven well-known optimizers to train the *self-training* classifiers, out of which, Stochastic Gradient Descent (SGD) outperformed others with the value of F1 score at 94.29%, affirming the relevance of this exposition.

Keywords: Semi-Supervised Learning, Deep Learning, Self-Training, MOOCs, Dropout Prediction

***Corresponding Author: Rinkaj Goyal:** University School of Information, Communication and Technology, Guru Gobind Singh (GGS)Indraprastha University, New Delhi 110078, India; Email: rinkajgoyal@gmail.com
Yamini Goel: University School of Information, Communication and Technology, Guru Gobind Singh (GGS)Indraprastha University, New Delhi 110078, India; Email: yaminigoel1996@gmail.com

1 Introduction

1.1 Why MOOC?

MOOC is an acronym for "Massive Open Online Course" which was initially coined by George Siemens and Stephen Downes in 2008. Due to large-scale participation with open enrollment and curriculum, it is termed as "Massive" and "Open" respectively [1, 2]. These MOOCs cover a spectrum of subjects ranging from songwriting to aerodynamics and are usually offered by the universities in association with providers such as Coursera, Edx, and Udacity, among others. Generally, MOOCs carry no tuition fee and require no pre-requisites other than student's interest in the subject and a good internet connection [3].

With a recent change in the job needs and study requirements, a shift towards MOOCs as an "alternative learning resource" has been seen in the student community. Since MOOCs bridge the gap between the industrial demands and the university curriculum, they have become an ideal source for self-development [4]. According to class-central, a MOOC aggregator, the year 2019 witnessed a global enrollment (excluding China) of 110 million students in 13,500 courses [5]. The same statistic in the previous year, however, was 101 million students (China included) in 11,400 courses offered by more than 900 universities [6]. Such a massive expansion in a single year clearly shows the growing popularity of MOOCs, which are driving a shift from conventional memorization-based education system to knowledge-driven problem-solving skills [7].

Table 1 below delineates the journey of MOOCs from open courseware to the full-fledged degrees of elite universities in online mode. As depicted in table 1, this journey, the formal open educational resources, starts with the open degrees offered by the British Open University in 1971. Massachusetts Institute of Technology (MIT) started offering lecture notes and accompanying materials (including digital video recordings in some courses) for free downloading under OpenCourseWare (OCW). The tsunami of MOOCs, however, started with an online class "Introduction to AI" offered by two professors of Stanford University, Sebastian Thrun and Peter Norvig, which has attracted

Table 1: Taxonomy of MOOCs (Adapted from [8, 9]).

Types	Description
Transfer MOOCs	MOOC is created from an existing classroom lectures
Made MOOCs	Different resources like videos, interactive materials, and other activities are exclusively developed for a MOOC.
Synch MOOCs	MOOC with a fixed start and end date.
Asynch MOOCs	These are self-paced which offer flexibility in engagement and submission.
Adaptive MOOCs	These support personalized learning by dynamically assessing user's activities including event logs.
Group MOOCs	Focuses on small group collaboration.
Connectivist MOOCs	No fixed knowledge. The participants do knowledge sharing through their interpersonal connections across a network of peers.
Mini MOOCs	Short courses that require less time.

more than 160,000 learners around the globe. MOOCs had grown to such an extent that Georgia Tech started an online Master of Science program in Computer Science in collaboration with Udacity and AT&T for a fraction of the normal cost [10]. Clark [9] outlined the MOOC taxonomy given in Table 1.

Despite being an easily accessible platform, MOOCs are surrounded by a varied number of challenges. One of the most prominent challenges being the low retention rate of the learners. The median of those who complete a MOOC is 12.6% [11]. For example, in a MOOC offered by Duke University in 2012, 7,761 students out of 12,175 registered students ever watched a video. Only 3,658 learners took at least one quiz, 346 attempted the final exam, and 313 received a certificate indicating a completion rate of approximately 3% [12]. MOOCs offered by elite universities, including MIT, Stanford, and UC Berkley, have reported dropout rates as high as 95% [13].

1.2 Why such low Completion Rate?

The high attrition rate in MOOCs has been attributed to a variety of factors that are mainly classified into Student related factors, MOOCs related factor, and Miscellaneous factors [4].

(i) Student-related factors
   (a) Lack of motivation - Lack of students' motivation is the most crucial factor for a high student dropout rate [4]. The motivation of students varies by several factors that include the development of personal and professional identity, future economic benefits, challenge and achievement, entertainment, and fun [14]. Therefore, it becomes important to explore the motives that encourage students to enroll in MOOCs. Researchers at Duke University found that enjoyment and fun were selected as the most important reason for registering in a MOOC by a large majority of students (95%) followed by a general interest in the topic selected by 87% of the students. Around 15% of students used the MOOC to help them decide if they wish to take higher education classes on the topic. In contrast, a significant minority of students (10%) enrolled because they could not afford to pursue formal education [15].
   (b) Lack of time - Watching videos, completing quizzes, and assignments requires considerable time, which students may not like to invest [15]. Many students complain that the actual time required to complete the MOOC exceeds the stated time [16].
   (c) Insufficient background knowledge and skills - Sometimes, students lack the necessary background knowledge required to understand the course content. Mainly, insufficient prior mathematics knowledge required to understand the MOOC content and complete the quizzes is observed to be an issue among learners [4, 15, 17].

(ii) MOOC-related factors
   (a) Course Design - Course design (which consists of course content, course structure and information delivery technology [18]) is considered one of the relevant factors for a high dropout rate [4, 19]. Among the three components, course content is the number one driver in determining the perceived quality of the online learning experience [20] and the most significant predictor of dropouts in MOOCs [19]. Along with course design, factors like poor course delivery, harsh criticism on students' written work from faculty, extensive lecture notes with material not significantly different

from textbooks influence the student's experience in the course negatively, thereby encouraging them to quit [21].

(b) Feelings of isolation and the lack of interactivity in MOOCs - Lack of real-time collaboration with learners in MOOCs is another factor that is said to influence student attrition rate [4] directly. Learners mention feeling isolated with poor communication with both the course instructor and fellow peers. Low interaction and poor feedback without group interactions and teamwork discouraged them from continuing the course further [19].

(c) Hidden Cost - The hidden cost could be another cause of a high dropout rate [13]. Despite MOOCs' reputation as a free resource for online education, students need to pay for their certificates [22] or sometimes buy costly textbooks recommended by lecturers [13].

(iii) Other factors

(a) Casual enrollments - Some researchers have noted that many enrollments in MOOCs are from students who have no real intention to complete them. They enroll in courses with purposes as assorted as "out of curiosity" or "with an intent to explore content and MOOCs in general" [17, 23]. Grover et al. [24] view these participations as "a by-product of the open-access nature of the courses and the novelty of the medium."

(b) Peer review - It has been noted that courses that rely purely on peer grading often suffer more in terms of students' course completion rate than others [25]. Peer grading requires more work on the learner's part, thereby adding to MOOC workload [17].

1.3 How dropout rate can be reduced?

(i) Course Design and Instruction - Despite a similar teaching method to face-to-face learning, MOOCs rely on the objectivist-individual teaching approach and lack constructivist and group-oriented approaches. Greater use of constructivist and innovative teaching may address major MOOCs related issues [26]. The instructional design quality of a MOOC is a prominent determinant and prerequisite for active learning along with the experiences of learners and other stakeholders [27]. Margaryan et al. [27] analyzed 76 randomly selected MOOCs to evaluate their instructional design quality. In their report, though MOOCs were found well-packaged, most of the MOOCs fall short of following instructional design principles. A well-designed course that focuses on interactivity includes a number of topics for discussion, accepts feedback from its learners and experts, and links sources to appropriate content [28], which promotes increased participation [21].

(ii) Interaction among learners - Sunar et al. [29] highlighted the role of learner's interactions with each other as a crucial factor in sustaining engagement in the course. They empirically investigated a correlation between completion rate in an eight-week FutureLearn MOOC and the learners' frequent interaction with other participants whom they follow (Twitter-like follow system available in FutureLearn). The integration of social tools in MOOC could be a potential solution to assist peers with problems, encourage discussions, and contribute new resources. However, the course forum is still the preferred tool in MOOCs, and the detailed impact of social tools like Slack, Telegram, Facebook, etc. on learning outcomes requires more endorsement [30].

(iii) Comfort Level with Technology - As MOOCs are online courses, they require technology to deliver content and connect to students. Learners may not be familiar with the required technologies. Therefore, the lecturers should generate ways to assist students with the delivery platform like technical training manuals or incorporating technical skills in the course content [21].

(iv) Feedback - The assessment or feedback is a crucial step in any educational learning process. Unlike classroom teaching, where faculty are able to respond to students immediately, this response is delayed in online courses. The literature review suggests that this delay should not exceed 48 hours. The more the students feel connected to the learning environment and their instructor, the less likely they are to drop out of the course [21].

2 Predictive Modelling for MOOC dropouts

Although research relating to dropouts in MOOCs was scarce initially, there is now a growing community of researchers working in the area [4]. A variety of statisti-

Table 2: Overview of relevant research on MOOC dropout prediction (Adapted from [4, 52]).

Single course MOOC			
Author	**Year**	**Dataset**	**Technique**
Moreno-Marcos et al. [53]	2020	Coursera MOOC	RF, GLM, SVM & DT
Xing and Du [54]	2019	Canvas MOOC	DL
Liu and Li [55]	2017	XuetangX MOOC	K-means
Nagrecha et al. [31]	2017	edX MOOC	DT & LR
Chen and Zang [56]	2017	Coursera MOOC	RF
Xing et al. [43]	2016	Canvas MOOC	GBN & DT
Crossley et al. [48]	2016	Coursera MOOC	NLP
Chaplot et al. [46]	2015	Coursera MOOC	ANN
Multiple course MOOC			
Author	**Year**	**Dataset**	**Technique**
Mourdi et al. [57]	2020	Stanford dataset	SVM, kNN, DT, NB, LR & Voting
Mubarak et al. [58]	2020	OULAD	LR & HMM
Chen et al. [59]	2020	HarvardX	Survival Analysis
Sun et al. [60]	2019	XuetangX dataset	RNN
Chen et al.[61]	2019	KDD Cup 2015	DT & ELM
Liao et al. [62]	2019	KDD Cup 2015	CTC
Alamri et al. [63]	2019	FutureLearn MOOCs	RF, AB, XGBoost & GB
Hassan et al. [64]	2019	OULAD	LSTM
Wen et al. [65]	2019	KDD Cup 2015	CNN
Feng et al. [66]	2019	XuetangX & KDD Cup 2015	CFIN
Cristea et al. [67]	2018	FutureLearn MOOCs	Statistical Method
Haiyang et al. [68]	2018	OULAD	TSF
Qiu et al. [69]	2018	KDD Cup 2015	FSPred: uses LR
Ardchir et al. [70]	2018	KDD Cup 2015	SVM
Wang et al. [42]	2017	KDD Cup 2015	CNN & RNN
Al-Shabandar et al. [44]	2017	HarvardX	LR, LDA, NB, SVM, DT, RF, NN & SOM
Al-Shabandar et al. [45]	2017	HarvardX	DT & NN
Vitiello et al. [71]	2017	Universidad Galileo & Curtin University MOOCs (edX)	SVM & Boosted DT
Cobos et al. [72]	2017	FutureLearn & edX	GBM, Weighted kNN, Boosted LR & XGBoost
Li et al. [51]	2016	KDD Cup 2015	MV SSL
Liang et al.[33]	2016	XuetangX dataset	SVM, LR, RF & GBDT
Wang and Chen [73]	2016	KDD Cup 2015	NSSM
Vitiello et al. [74]	2016	University of Galileo MOOCs	K-means & SVM
Tang et al. [75]	2015	MITx & HarvardX courses (edX)	DT
Yang et al. [76]	2015	Coursera MOOCs	Survival Analysis
Fei and Yeung [41]	2015	Coursera & edX	RNN & HMM

RF: Random Forest, GLM: Generalized Linear Model, SVM: Support Vector Machine, DT: Decision Trees, DL: Deep Learning, LR: Logistic Regression, GBN: General Bayesian Network, NLP: Natural Language Processing, ANN: Artificial Neural Network, NB: Naive Bayes, RNN: Recurrent Neural Network, ELM: Extreme Learning Machine, CTC: Clustering and Tensor Completion, AB: Adaptive Boost, GB: Gradient Boost, LSTM: Long Short Term Memory, CNN: Convolutional Neural Network, CFIN: Context-aware Feature Interaction Network, HMM: Hidden Markov Models, LDA: Linear Discriminant Analysis, NN: Neural Network, SOM: Self Organised Map, NSSM: Nonlinear State Space Model, kNN: k-Nearest Neighbours, TSF: Time Series Forest classification algorithm, GBM: Generalised Boosted regression Models, XGBoost: eXtreme Gradient Boosting, MV SSL: Multi-View Semi-Supervised Learning, GBDT: Gradient Boosting Decision Tree cal and ML techniques (Logistic regression [31–38], SVM [32, 33, 39, 40] etc.) have been applied to MOOC data to create prediction models. Other works involve Recurrent Neural Networks (RNN) with Long Short Term Memory (LSTM) [41], Convolution Neural Network (CNN) and Recurrent Neural Network (RNN) in combination [42] for successful predictions. Among ML algorithms, some works also focus on decision tree [43–45], sentiment-based artificial neural network [46], deep neural network [47], and natural language processing statistical models [48–50]. Li et al. [51] proposed a semi-supervised learning model for the task of dropout prediction. Table 2 details the state-of-the-art research on the dropout prediction in MOOCs.

Although most prior studies consider click-stream features as engagement patterns [4], few use social networks or grades to make predictions [77, 78]. Other works, e.g. [32], use a combination of user demographics, forum activities, and learning behavior to predict learners' learning effectiveness.

Many predictive models with reasonable accuracy measures have been proposed to identify at-risk students. Apart from predictions, these models may initiate an intervention before learners quit [34]. Low completion rates in MOOCs have motivated researchers in the automatic detection and prevention of student dropouts. Whitehill et al. [34] conducted a "dynamic survey intervention" on students who recently dropped out of a MOOC to learn the reason. They concluded that the mere act of inquiring propelled students to rejoin the MOOC. Some researchers have deployed their predictive models as AI Plugins on browsers to make real-time predictions. One such example is a recent work by Feng et al. [66], where authors deployed their algorithm to XiaoMu, an AI-based learning assistant system on XuetangX to predict dropouts as the MOOC proceeds. Upon identifying an "at-risk student", XiaoMu sends the learner an intervention message to encourage them. Results show that their strategy improves users' engagement significantly.

This study, however, uses activity features and the impact of friends as engagement patterns in MOOCs. We use the demographic similarity between learners to discover their friend relationships. Students with a similarity score above a threshold are considered as friends. Prior research to predict at-risk students generally applies supervised learning models. These models require a huge amount of labelled data, which due to the high time-cost of labelling, are often limited [51]. We, however, employ a semi-supervised learning-based model - *self-training* to predict dropouts. We use a deep learning model as a classifier in the self-training technique. To study the efficiency of the deep model to deal with the dropout prediction problem, we also use it separately to make predictions. Furthermore, we evaluated the contribution of different features in predicting dropouts in MOOCs.

3 Methods and Methodology

Few studies have applied user demographics to find users' similarity measures for other tasks, like friend matching and item recommendations. A well-established finding in sociology is "friends tend to be similar" [79, 80]. In this paper, user demographic information (gender, age, education) has been applied to measure the similarity between them. We compute users' similarities using *cosine similarity*, which has been used in several prior studies to measure proximity among members [81, 82]. Mazhari et al. [83] propose a combination of Levenshtein distance, Dice's coefficient, Cosine similarity, and Jaccard index to measure demographics-based similarity among users for friend matching. We consider people with high similarity scores as friends.

Users' online study behavior is said to influence one another. The probability that a learner completes the course and obtains a certificate increases three times when he/she has one or more "certificate friends" [32]. Feng et al. [66] inferred that students' probability of dropout increases from 0.33 to 0.87 when the count of his dropout friends ranges from 1 to 10, indicating the strong influence of friends' attrition behavior.

3.1 Dataset description

The dataset used in this paper comes from XuetangX, one of the largest MOOC platforms in China. The dataset is publicly available and has previously been used by Feng et al. [66]. They proposed a Content-aware Feature Interaction Network (CFIN) model that incorporates context information (user's demographic and course information) and learning activity to make predictions. However, we use the demographic information of the students to measure similarity (and thereby, friendship) among them. Our model uses a combination of students' learning activity and influence from friends to predict dropouts. This dataset contains 685,387 registered students. The courses in the dataset based on their learning mode are classified into Instructor-paced Mode (IPM) and Self-paced Mode (SPM). There are 698 IPM and 515 SPM courses in the dataset. Table 3 provides the descriptive statistics of XuetangX dataset.

Table 3: Descriptive Statistics of XuetangX dataset (Adapted from [66]).

Enrollments	Count
Total	685,387
Dropouts	578,076
Completions	107,311
Courses	1,213

The information available in the dataset includes information about the course (ID, start and end date, course_type, category), student demographic information (user_id, gender, education, birth), their activity records, and their final result (dropout or non-dropout). Table 4 represents attributes of different files. The dataset contains rich information with multiple types of students' activities: video watching, web page clicking, creating threads, and many more.

3.2 Preprocessing of Input

The records in the dataset are raw, which means they cannot be input to the deep neural network, and require preprocessing. The preprocessing of the input data can be divided into two parts:

3.2.1 Making friends based on demographics

As shown in Table 4, the user profile information file has the user's demographic information - user_id (the user's id), gender, education, and birth (the birth year). These records in the dataset are in text format, which cannot be used to calculate the similarity between users. For this, we convert them to vectors that could be processed further. In this paper, we compute friendship between users who have enrolled in the same course.

The "age" attribute of a user is calculated using the user's *birth* year and course's *year*. The users are then classified into three categories based on their age (≤25:young, >50:old, $(25, 50]$:middle). Each attribute may take several different values. We convert each value (of gender and age_group) to their corresponding one-hot vector. One-hot encoding scheme converts a variable to a binary variable with the number of bits equal to the number of unique values the variable can have. The education attribute is label encoded based on pre-defined labels. Label encoding transforms a variable to a number. The dataset contains a rich set of values for the education column, however we define labels for values available in the courses' data used in this study. An example of converting the attributes of a raw record to its corresponding vector representation is shown in Figure 1.

Table 4: Attributes in various files of XuetangX dataset (Adapted from [66, 84]).

User Profile Information	
Attribute	**Meaning**
user_id	the id of user
gender	the gender of user
education	user's education level
birth	user's birth year
Course Information	
Attribute	**Meaning**
id	the id (number) of course
course_id	the id (string) of course
start	course start time
end	course end time
course_type	course mode (0: instructor-paced course, 1: self-paced course)
category	the category of course
Log Activities Files	
Attribute	**Meaning**
enroll_id	denoting the user and course id pair
username	the id of the user
course_id	denoting the course's id
session_id	the id of session
action	the type of activity record
object	corresponding object of the action
time	the time when each action occurred
Truth Files	
Attribute	**Meaning**
enroll_id	denoting the user and course id pair
truth	the label of user's dropout (1: dropout, 0: non-dropout)

The vectors thus obtained are used to compute the similarity between users. Users with high *cosine similarity* score i.e., greater than or equal to 0.8 are considered as friends. Algorithm 1 outlines the implementation of making friends based on their demographic information.

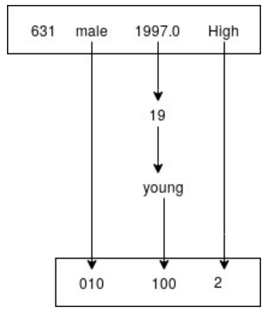

Figure 1: Example of converting raw data to their corresponding vector form (Adapted from [42]).

3.2.2 Preprocessing activity-log

In this section, we elaborate the preprocessing of log activities (Table 4). Firstly, we one-hot encode the action attribute, converting each value of the attribute to a vector with a size value of 18. Since the volume of these vectors is huge, they need to be combined. For this, we add up all these vectors of the same date corresponding to each enroll_id in a bitwise fashion. This method of converting and combining one-hot vectors has been used in prior studies, e.g., [42].

Two additional attributes - *drop_friends* and *non_drop_friends* are concatenated to the added one-hot vectors. Any friend with the last action before the current date is regarded as a "dropout friend", thus incrementing the *drop_friends* for the corresponding enroll_id. Similarly, friends with actions on or after the current date are called "non-dropout friends", and they add to *non_drop_friends* value. This results in 63 vectors, each of size 20.

Lastly, we concatenate all vectors (for unique enroll_id) into a matrix, with each row to be the vector for each date. Algorithm 2 provides a detailed implementation insight of the preprocessing stage.

The dropout prediction problem, considered in this paper, is a classification problem with a huge difference in the number of instances of each class (Class Imbalance Problem). Several sampling techniques (Undersampling, Oversampling, and Hybrid methods) are used for handling the Class Imbalance Problem in various domains. In this paper, we use random undersampling and random oversampling methods. We did a train-validation split of 80%-20%, and sampling techniques were applied to the training set alone to get an equal proportion of dropouts and non-dropouts.

Algorithm 1: Making friends based on the user's demographics

```
Procedure GetFriends(learner)
    Get users enrolled in learner's course in students
    Fill missing values of students.gender and
      students.education with "noInfo_gender" and
      "noInfo_edu" respectively
    foreach s in students do
        Set s.age ← course_year − s.birth
    end
    Fill missing values of students.age with arithmetic mean of
      age attribute
    foreach s in students do
        if s.age ≤ LOW then                        // LOW: 25
            s.age_group ← young
        else if s.age ∈ (LOW, MID] then            // MID: 50
            s.age_group ← middle
        else
            s.age_group ← old
        end
    end
    foreach s in students do
        s.gender ← one_hot(s.gender)
        s.age_group ← one_hot(s.age_group)
        s.education ← label_encode(s.education, LABELS)
        /* LABELS: [noInfo_edu:0, Middle:1, High:2,
           Associate:3, Bachelor's:4, Master's:5,
           Doctorate:6]                                  */
    end
    foreach co_learner in students do          /* co_learner ≠
      learner */
        Set
          similarity ← cosine_similarity(learner, co_learner)

        if similarity ≥ THRESHOLD then      /* THRESHOLD: 0.8
          */
            friends ← friends.add(co_learner)
        end
    end
    return friends
```

Algorithm 2: Preprocessing the data

```
Input  :Users enrolled in same course as students, their logs as
        action, last date of log action as last_action_date,
        course_start and course_end date
Output :63X20 sized matrices for each user
foreach s in students do
    s.action ← one_hot(s.action)    /* One-hot vector size:
      18 */
end
foreach s in students do      /* students grouped by enroll_id
  and date */
    foreach date ∈ [course_start, course_end] do  /* ∀s runs
      63 times */
        Add one-hot vectors corresponding (s, date)
        Set s.drop_friends & s.non_drop_friends to 0 for date
          // Resultant vector size: 20
        foreach friend in GetFriends(s) do
            if friend.last_action_date ≥ date then
                /* last_action_date for non-dropouts set
                to a date after course ends */
                s.non_drop_friends ←
                  s.non_drop_friends + 1
            else
                s.drop_friends ← s.drop_friends + 1
            end
        end
    end
    Sort vectors corresponding s based on date
    Make matrix corresponding s     /* Resultant matrix size:
      63 × 20 */
end
```

3.3 Deep Model

The deep learning model, inspired by the prior work of Wang et al. [42], contains an input layer, an output layer, and six hidden layers. The input to the prediction model is the matrix representation of action logs and count of friends, and output is a binary class, where 1 represents the student dropped out of the course, and 0 represents the student did not drop out. As illustrated in Figure 2, the first and third layers perform convolution operation to extract features. The second and fourth layers in the deep model are pooling layers; here, we use max pooling. The fifth is the fully connected or dense layer, which combines features extracted from previous layers succeeded by the recurrent layer. The last layer of the model is the dense layer.

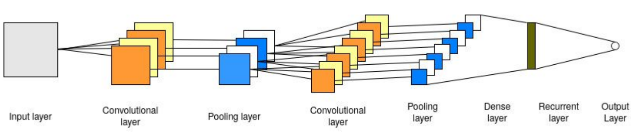

Figure 2: Illustration of different layers of the Deep Neural Network (Adapted from [42]).

3.4 Self-Training

Self-training is a method for semi-supervised learning [85]. As described in Figure 3, the learning process of a self-training algorithm starts with training a classifier with labelled data, then repeatedly bootstrapping the classifier with additional labelled data obtained from its own predictions with a high confidence score. This process continues until some pre-defined termination condition is reached. Self-training has previously been applied to several problems ranging from fine-grained classification [86] to parsing [87]. However, its capability has not been explored in the context of MOOCs analytics. The classifier we use for self-training is the Deep Model described in section 3.3. To measure the confidence of predictions, we train five neural networks with different initial random weights. A count of networks that agree on the label is regarded as the confidence score for that prediction. Data with a confidence score above a threshold is appended to the training dataset. The retraining of the classifier stops when either the unlabelled dataset becomes empty, or the number of maximum iterations are reached. A non-science course (course_id: *"course-v1:TsinghuaX+JRFX01+2016_TS"*) is used to train the initial model while two different non-science courses (course_id: *"course-v1:TsinghuaX+00510663X+2016_TS" and "course-v1:MITx+15_390x_2016T1+2016_TS"*) are used to retrain the model repeatedly. The labels in the two non-science courses are known, but we ignore them for retraining purposes.

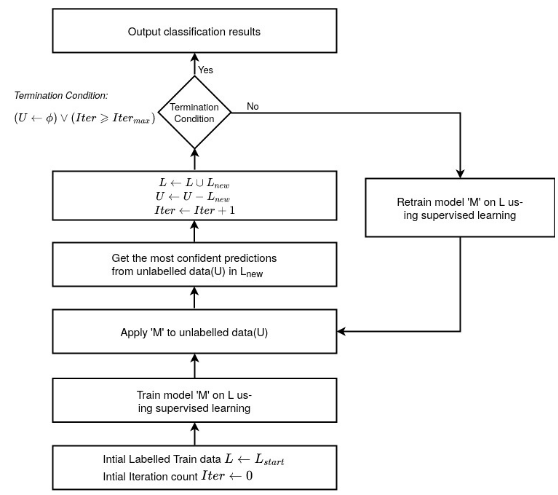

Figure 3: Procedural flow of the Self-Training Semi-Supervised learning technique.

4 Results

As mentioned before, the dropout prediction problem suffers from class imbalance in the dataset. In such cases, accuracy is not considered an appropriate measure for model evaluation. Therefore we use F1 Score, Precision and Recall as evaluation metrics. Table 5 lists the details of parameters in Deep Model. The batch size was set to 20. The model was trained for 50 epochs with early stopping patience set to 5. We implemented the model in *Keras* [88]. We use sigmoid activation function with binary cross-entropy (as the loss) function to train the model. In *self-training*, the number of maximum iterations were fixed to 10. We selected pseudo-labelled data with confidence score greater than 3 in majority voting to re-train the model.

Table 6 presents the optimizer-wise results of an ensemble of deep models trained on three non-Science IPM courses. The ensemble was a collection of thirty deep models having the same configuration with initial random weights set to the default values of each layer from *Keras* [88]. They were trained on the same courses' dataset, and predictions were made. The actual prediction was calculated as the mode of all the predictions. Our results

Table 5: Hyperparameters used in implementing the Deep Model

Model Hyperparameters (Adapted from [42])			
Layer	**Name**	**Activation**	**Hyperparameters**
0	Input	-	-
1	Convolution	relu	Filters: 20, Filter Size: 5 × 5
2	Max Pooling	-	Pool Size: 2 × 2
3	Convolution	relu	Filters: 50, Filter Size: 5 × 5
4	Max Pooling	-	Pool Size: 2 × 2
5	Fully Connected	relu	Units: 20
6	Simple RNN	relu	Units: 50

Optimization Hyperparameters	
Optimizer	**Learning Rate**
SGD	0.0005
RMSProp	0.0005
Adagrad	0.0001
Adadelta	0.05
Adam	0.005
Adamax	0.0001
Nadam	0.0001

Table 6: Results on Science IPM course under Different Metrics

Optimizer	F1(%)	Precision(%)	Recall(%)
SGD	90.36	95.70	85.58
RMSProp	90.91	95.74	86.54
Adagrad	92.00	95.84	88.46
Adadelta	89.80	95.66	84.62
Adam	90.82	96.74	85.58
Adamax	89.23	95.60	83.65
Nadam	88.54	96.59	81.73

show that Adagrad outperforms all other optimizers with the highest F1 score, and its Precision on Science course achieves 95.84%. Compared to Adadelta and Adamax, SGD achieves 0.96% and 1.93% Recall improvement in Science course, respectively. The model trained using RMSProp optimizer achieves F1 score 90.91%, Precision 95.74%, and Recall 86.54%. Adam reports 90.82% F1 score, 96.74% Precision and 85.58% Recall.

In order to identify the importance of different features, we use the permutation feature importance method (introduced by Breiman [89]). In this, we measure the contribution of a feature by calculating the variety in the model's prediction after permuting the feature. Particularly, we shuffle every type of activity feature one by one and then evaluate the performance of the model on the modified dataset. The activity features used are grouped into four major categories - video activity, problem-related activity, influence-from-friends, and remaining miscellaneous activities. The importance of various activity features are shown in Table 7. To evaluate the contribution of each feature more accurately, we calculate differences between the performances of deep models on the original dataset and permuted dataset on the F1 score. These values are illustrated in Table 8. Results show that all features are useful for predictions. The miscellaneous click-stream features play the most important role, followed by influence from friends.

Table 7: Contribution analysis for different features (F1 score)

Optimizer	Video(%)	Problem(%)	Friends(%)	Others(%)
SGD	80.45	72.73	61.44	59.21
RMSProp	71.86	68.71	49.30	49.30
Adagrad	86.32	79.56	74.42	73.56
Adadelta	80.68	72.39	50.36	45.59
Adam	79.78	77.19	72.73	68.75
Adamax	78.65	71.52	43.28	39.69
Nadam	76.30	67.08	43.28	39.39

Table 8: Difference between performance of deep model on original and modified dataset (F1 score)

Optimizer	Video(%)	Problem(%)	Friends(%)	Others(%)
SGD	9.91	17.63	28.92	31.15
RMSProp	19.05	22.2	41.61	41.61
Adagrad	5.68	12.44	17.58	18.44
Adadelta	9.12	17.41	39.44	44.21
Adam	11.04	13.63	18.09	22.07
Adamax	10.58	17.71	45.95	49.54
Nadam	12.24	21.46	45.26	49.15

Experimental results of self-training on a Science based IPM course (course_id: *"course-v1: TsinghuaX+400182X+2016_TS"*) are shown in Table 9. SGD proves to be top performer with highest F1 score followed by Adadelta (F1 score: 91.63%). The model trained using Adamax optimizer achieves F1 score 89.45%, Precision 93.68% and Recall 85.58%. Nadam, on the other hand, performs a little better with F1 score 91.09%, Precision 93.88% and Recall 88.46%. RMSProp reports 82.42% F1 score, 96.15% Precision and 72.12% Recall. Adagrad and Adam achieve F1 score 90.64% and 90.36% respectively.

Table 9: Results of Self-Training on Science IPM course under Different Metrics

Optimizer	F1(%)	Precision(%)	Recall(%)
SGD	94.29	93.40	95.19
RMSProp	82.42	96.15	72.12
Adagrad	90.64	92.93	88.46
Adadelta	91.63	93.94	89.42
Adam	90.36	95.70	85.58
Adamax	89.45	93.68	85.58
Nadam	91.09	93.88	88.46

Table 10 presents the contribution of various features in predicting MOOC dropouts by self-training technique. Similar to supervised learning results, we observe that all features play essential roles in dropout prediction, but "Others" features followed by influence from friends are the most prominent ones in all the cases.

Table 10: Contribution analysis for different features (F1 score)

Optimizer	Video(%)	Problem(%)	Friends(%)	Others(%)
SGD	85.26	80.00	56.16	52.11
RMSProp	60.00	44.77	30.89	27.87
Adagrad	78.21	69.88	66.25	63.75
Adadelta	73.68	69.94	59.60	52.78
Adam	89.69	82.42	81.56	78.89
Adamax	70.30	67.09	54.79	44.93
Nadam	77.53	78.16	71.85	69.88

5 Discussion

In this study, we explored the influence of friends in the prediction of dropouts in MOOCs. We used two techniques - supervised and semi-supervised based models for this task. To study the efficiency of the models, the training and testing were done across different courses. The deep model was trained on three non-Science courses (course_id:*"course-v1:TsinghuaX+00510663X+2016_TS"*, *"course-v1:TsinghuaX+JRFX01+2016_TS"* and *"course-v1:MITx+15_390x_2016T1+2016_TS"*) after class balancing. To maintain consistency with the other two courses' data, the additional action attribute available in the Business course was omitted for the experiment. The deep learning model, thus trained, was used to make predictions on a Science (course_id:*"course-v1:TsinghuaX+400182X+2016_TS"*).
In the case of self-training, the initial model was trained on an Economics MOOC (course_id:*"course-v1:TsinghuaX+JRFX01+2016_TS"*). The model was then repeatedly trained on two non-Science courses (course_id:*"course-v1:TsinghuaX+00510663X+ 2016_TS" and "course-v1:MITx+15_390x_2016T1+2016 _TS"*). Finally, the semi-supervised trained model was used to make predictions on a Science course (course_id:*"course-v1:TsinghuaX+400182X+2016_TS"*). Results show both the techniques achieve high performance in predicting MOOC dropouts.

6 Conclusion

MOOCs have grown rapidly in previous years and are continually growing. Their easy availability and flexibility have attracted millions of students. Despite the growing popularity of MOOCs, they are faced with a plethora of challenges. One such challenge that has attracted the attention of the researcher's community is addressing a high dropout rate. In this study, we adopt a semi-supervised learning model to identify potential dropouts. The model, apart from click-stream activity pattern, uses impact from friends as engagement patterns. We employed demographic similarity between users to determine friendship amongst them. Experimental results show that our approach achieves high performance in terms of F1 score on a public dataset extracted from XuetangX. The *self-training* technique achieves comparable results, if not better than state-of-the-art supervised learning methods. Results on the permuted dataset present the significant role of the learner's friends' attrition in predicting her behavior in the MOOC. Although this study concludes *self-training* as an effective approach to predicting MOOC dropouts, this work may be extended by using more customized classifiers that could predict not only MOOC specific dropouts but also other aspects of the learner's behavior. Model deployment across various MOOC platforms and generating real-time predictions on running MOOCs may also be considered as the future scope of this work.

Acknowledgement: We thank Rishabh Narang, MS Computer Science, Columbia University and Shristi Mudgal, MS Computer Science, Technical University of Munich for useful discussions. Their remarks and suggestions were extremely helpful to carry out the research further.

References

[1] Masters K., A brief guide to understanding MOOCs, The Internet Journal of Medical Education, 1(2), 2011, 2
[2] Hew K.F., Cheung W.S., Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges, Educational research review, 12, 2014, 45–58
[3] McAuley A., Stewart B., Siemens G., Cormier D., The MOOC model for digital practice, 2010
[4] Dalipi F., Imran A.S., Kastrati Z., MOOC dropout prediction using machine learning techniques: Review and research challenges, in 2018 IEEE Global Engineering Education Conference (EDUCON), IEEE, 2018, 1007–1014
[5] Shah D., Online Degrees Slowdown: A Review of MOOC Stats and Trends in 2019 - Class Central, 2020
[6] Shah D., Year of MOOC-based Degrees: A Review of MOOC Stats and Trends in 2018 - Class Central, 2019
[7] Jewitt K., The MOOC Revolution–massive open online courses: the answer to problems facing education or an experiment that could destroy centuries of tradition., Compass: Journal of Learning and Teaching, 10(1), 2017
[8] Hood N., Littlejohn A., Quality in MOOCs: Surveying the terrain, 2016
[9] Clark D., Donald Clark Plan B: MOOCs: taxonomy of 8 types of MOOC, http://donaldclarkplanb.blogspot.com/2013/04/moocs-taxonomy-of-8-types-of-mooc.html, 2013, (Accessed on 02/16/2020)
[10] Swenson P., Taylor N.A., Online teaching in the digital age, Sage Publications, 2012
[11] Jordan K., Massive open online course completion rates revisited: Assessment, length and attrition, The International Review of Research in Open and Distributed Learning, 16(3), 2015, 10.19173/irrodl.v16i3.2112
[12] Catropa D., Big (MOOC) Data: Inside Higher Ed, 2013
[13] Khalil H., Ebner M., MOOCs completion rates and possible methods to improve retention-A literature review, in EdMedia+ Innovate Learning, Association for the Advancement of Computing in Education (AACE), 2014, 1305–1313
[14] Yuan L., Powell S., MOOCs and open education: Implications for higher education, 2013
[15] Belanger Y., Thornton J., Barr R.C., Bioelectricity: A quantitative approach–Duke University's first MOOC, EducationXPress, 2013(2), 2013, 1–1
[16] Conole G.G., MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs, Revista de Educación a Distancia, (39), 2013
[17] Onah D.F., Sinclair J., Boyatt R., Dropout rates of massive open online courses: behavioural patterns, EDULEARN14 proceedings, 1, 2014, 5825–5834
[18] Peltier J.W., Drago W., Schibrowsky J.A., Virtual communities and the assessment of online marketing education, Journal of Marketing Education, 25(3), 2003, 260–276
[19] Hone K.S., El Said G.R., Exploring the factors affecting MOOC retention: A survey study, Computers & Education, 98, 2016, 157–168
[20] Peltier J.W., Schibrowsky J.A., Drago W., The interdependence of the factors influencing the perceived quality of the online learning experience: A causal model, Journal of Marketing Education, 29(2), 2007, 140–153
[21] O'Brien B., Online student retention: can it be done?, Association for the Advancement of Computing in Education (AACE), 2002
[22] Open Culture, The Big Problem for MOOCs Visualized, http://www.openculture.com/2013/04/the_big_problem_for_moocs_visualized.html, 2013, (Accessed on 01/30/2020)
[23] Kolowich S., Coursera Takes a Nuanced View of MOOC Dropout Rates, 2013
[24] Grover S., Franz P., Schneider E., Pea R., The MOOC as Distributed Intelligence: Dimensions of a Framework & Evaluation of MOOCs., in CSCL (2), 2013, 42–45
[25] Parr C., Mooc completion rates 'below 7%', 2013
[26] Toven-Lindsey B., Rhoads R.A., Lozano J.B., Virtually unlimited classrooms: Pedagogical practices in massive open online courses, The internet and higher education, 24, 2015, 1–12
[27] Margaryan A., Bianco M., Littlejohn A., Instructional quality of massive open online courses (MOOCs), Computers & Education, 80, 2015, 77–83
[28] Parker A., Interaction in distance education: The critical conversation, AACE Journal, 1(12), 1999, 13–17
[29] Sunar A.S., White S., Abdullah N.A., Davis H.C., How learners' interactions sustain engagement: a MOOC case study, IEEE Transactions on Learning Technologies, 10(4), 2016, 475–487
[30] Alario-Hoyos C., Pérez-Sanagustín M., Delgado-Kloos C., Muñoz-Organero M., Rodríguez-de-las Heras A., et al., Analysing the impact of built-in and external social tools in a MOOC on educational technologies, in European Conference on Technology Enhanced Learning, Springer, 2013, 5–18
[31] Nagrecha S., Dillon J.Z., Chawla N.V., MOOC dropout prediction: lessons learned from making pipelines interpretable, in Proceedings of the 26th International Conference on World Wide Web Companion, International World Wide Web Conferences Steering Committee, 2017, 351–359
[32] Qiu J., Tang J., Liu T.X., Gong J., Zhang C., Zhang Q., Xue Y., Modeling and predicting learning behavior in MOOCs, in Proceedings of the ninth ACM international conference on web search and data mining, ACM, 2016, 93–102
[33] Liang J., Li C., Zheng L., Machine learning application in MOOCs: Dropout prediction, in 2016 11th International Conference on Computer Science & Education (ICCSE), IEEE, 2016, 52–57
[34] Whitehill J., Williams J., Lopez G., Coleman C., Reich J., Beyond prediction: First steps toward automatic intervention in MOOC student stopout, Available at SSRN 2611750, 2015
[35] Boyer S., Veeramachaneni K., Transfer learning for predictive models in massive open online courses, in International conference on artificial intelligence in education, Springer, 2015, 54–63
[36] Kizilcec R.F., Halawa S., Attrition and achievement gaps in online learning, in Proceedings of the second (2015) ACM conference on learning@ scale, ACM, 2015, 57–66
[37] He J., Bailey J., Rubinstein B.I., Zhang R., Identifying at-risk students in massive open online courses, in Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015
[38] Taylor C., Veeramachaneni K., O'Reilly U.M., Likely to stop? predicting stopout in massive open online courses, arXiv preprint arXiv:1408.3382, 2014
[39] Kloft M., Stiehler F., Zheng Z., Pinkwart N., Predicting MOOC dropout over weeks using machine learning methods, in Proceedings of the EMNLP 2014 workshop on analysis of large scale social interaction in MOOCs, 2014, 60–65

[40] Amnueypornsakul B., Bhat S., Chinprutthiwong P., Predicting attrition along the way: The UIUC model, in Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs, 2014, 55–59
[41] Fei M., Yeung D.Y., Temporal models for predicting student dropout in massive open online courses, in 2015 IEEE International Conference on Data Mining Workshop (ICDMW), IEEE, 2015, 256–263
[42] Wang W., Yu H., Miao C., Deep model for dropout prediction in MOOCs, in Proceedings of the 2nd International Conference on Crowd Science and Engineering, ACM, 2017, 26–32
[43] Xing W., Chen X., Stein J., Marcinkowski M., Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization, Computers in human behavior, 58, 2016, 119–129
[44] Al-Shabandar R., Hussain A., Laws A., Keight R., Lunn J., Radi N., Machine learning approaches to predict learning outcomes in Massive open online courses, in 2017 International Joint Conference on Neural Networks (IJCNN), IEEE, 2017, 713–720
[45] Al-Shabandar R., Hussain A., Laws A., Keight R., Lunn J., Towards the differentiation of initial and final retention in massive open online courses, in International Conference on Intelligent Computing, Springer, 2017, 26–36
[46] Chaplot D.S., Rhim E., Kim J., Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks, in AIED Workshops, volume 53, 2015, 54–57
[47] Whitehill J., Mohan K., Seaton D., Rosen Y., Tingley D., Delving deeper into MOOC student dropout prediction, arXiv preprint arXiv:1702.06404, 2017
[48] Crossley S., Paquette L., Dascalu M., McNamara D.S., Baker R.S., Combining click-stream data with NLP tools to better understand MOOC completion, in Proceedings of the sixth international conference on learning analytics & knowledge, ACM, 2016, 6–14
[49] Robinson C., Yeomans M., Reich J., Hulleman C., Gehlbach H., Forecasting student achievement in MOOCs with natural language processing, in Proceedings of the sixth international conference on learning analytics & knowledge, ACM, 2016, 383–387
[50] Coleman C.A., Seaton D.T., Chuang I., Probabilistic use cases: Discovering behavioral patterns for predicting certification, in Proceedings of the Second (2015) ACM Conference on Learning@ Scale, ACM, 2015, 141–148
[51] Li W., Gao M., Li H., Xiong Q., Wen J., Wu Z., Dropout prediction in MOOCs using behavior features and multi-view semi-supervised learning, in 2016 international joint conference on neural networks (IJCNN), IEEE, 2016, 3130–3137
[52] Gardner J., Brooks C., Student success prediction in MOOCs, User Modeling and User-Adapted Interaction, 28(2), 2018, 127–203
[53] Moreno-Marcos P.M., Muñoz-Merino P.J., Maldonado-Mahauad J., Pérez-Sanagustín M., Alario-Hoyos C., Kloos C.D., Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs, Computers & Education, 145, 2020, 103728
[54] Xing W., Du D., Dropout prediction in MOOCs: Using deep learning for personalized intervention, Journal of Educational Computing Research, 57(3), 2019, 547–570
[55] Liu T.y., Li X., Finding out reasons for low completion in MOOC environment: an explicable approach using hybrid data mining methods, DEStech Transactions on Social Science, Education and Human Science, (meit), 2017
[56] Chen Y., Zhang M., MOOC student dropout: pattern and prevention, in Proceedings of the ACM Turing 50th Celebration Conference-China, 2017, 1–6
[57] Mourdi Y., Sadgal M., Berrada Fathi W., El Kabtane H., A machine learning based approach to enhance MOOC users' classification., Turkish Online Journal of Distance Education (TOJDE), 21(2), 2020
[58] Mubarak A.A., Cao H., Zhang W., Prediction of students' early dropout based on their interaction logs in online learning environment, Interactive Learning Environments, 2020, 1–20
[59] Chen C., Sonnert G., Sadler P.M., Sasselov D.D., Fredericks C., Malan D.J., Going over the cliff: MOOC dropout behavior at chapter transition, Distance Education, 41(1), 2020, 6–25
[60] Sun D., Mao Y., Du J., Xu P., Zheng Q., Sun H., Deep Learning for Dropout Prediction in MOOCs, in 2019 Eighth International Conference on Educational Innovation through Technology (EITT), IEEE, 2019, 87–90
[61] Chen J., Feng J., Sun X., Wu N., Yang Z., Chen S., MOOC dropout prediction using a hybrid algorithm based on decision tree and extreme learning machine, Mathematical Problems in Engineering, 2019, 2019
[62] Liao J., Tang J., Zhao X., Course drop-out prediction on MOOC platform via clustering and tensor completion, Tsinghua Science and Technology, 24(4), 2019, 412–422
[63] Alamri A., Alshehri M., Cristea A., Pereira F.D., Oliveira E., Shi L., Stewart C., Predicting MOOCs dropout using only two easily obtainable features from the first week's activities, in International Conference on Intelligent Tutoring Systems, Springer, 2019, 163–173
[64] Hassan S.U., Waheed H., Aljohani N.R., Ali M., Ventura S., Herrera F., Virtual learning environment to predict withdrawal by leveraging deep learning, International Journal of Intelligent Systems, 34(8), 2019, 1935–1952
[65] Wen Y., Tian Y., Wen B., Zhou Q., Cai G., Liu S., Consideration of the local correlation of learning behaviors to predict dropouts from MOOCs, Tsinghua Science and Technology, 25(3), 2019, 336–347
[66] Feng W., Tang J., Liu T.X., Understanding dropouts in MOOCs, Association for the Advancement of Artificial Intelligence, 2019
[67] Cristea A.I., Alamri A., Kayama M., Stewart C., Alshehri M., Shi L., Earliest predictor of dropout in moocs: a longitudinal study of futurelearn courses, 2018
[68] Haiyang L., Wang Z., Benachour P., Tubman P., A time series classification method for behaviour-based dropout prediction, in 2018 IEEE 18th international conference on advanced learning technologies (ICALT), IEEE, 2018, 191–195
[69] Qiu L., Liu Y., Liu Y., An integrated framework with feature selection for dropout prediction in massive open online courses, IEEE Access, 6, 2018, 71474–71484
[70] Ardchir S., Talhaoui M.A., Jihal H., Azzouazi M., Predicting MOOC Dropout Based on Learner's Activity, International Journal of Engineering & Technology, 7(4.32), 2018, 124–126
[71] Vitiello M., Walk S., Chang V., Hernandez R., Helic D., Guetl C., MOOC dropouts: A multi-system classifier, in European Conference on Technology Enhanced Learning, Springer, 2017, 300–314
[72] Cobos R., Wilde A., Zaluska E., Predicting attrition from massive open online courses in FutureLearn and edX, in Proceedings of the 7th International Learning Analytics and Knowledge Conference, Simon Fraser University, Vancouver, BC, Canada, 2017, 13–17

[73] Wang F., Chen L., A Nonlinear State Space Model for Identifying At-Risk Students in Open Online Courses, International Educational Data Mining Society, 2016
[74] Vitiello M., Walk S., Hernández R., Helic D., Gütl C., Classifying students to improve MOOC dropout rates, Research Track, 2016, 501
[75] Tang J.K., Xie H., Wong T.L., A big data framework for early identification of dropout students in MOOC, in International Conference on Technology in Education, Springer, 2015, 127–132
[76] Yang D., Wen M., Howley I., Kraut R., Rose C., Exploring the effect of confusion in discussion forums of massive open online courses, in Proceedings of the second (2015) ACM conference on learning@ scale, 2015, 121–130
[77] Jiang S., Williams A., Schenke K., Warschauer M., O'dowd D., Predicting MOOC performance with week 1 behavior, in Educational data mining 2014, 2014
[78] Rosé C.P., Carlson R., Yang D., Wen M., Resnick L., Goldman P., Sherer J., Social factors that contribute to attrition in MOOCs, in Proceedings of the first ACM conference on Learning@ scale conference, ACM, 2014, 197–198
[79] Feld S.L., The focused organization of social ties, American journal of sociology, 86(5), 1981, 1015–1035
[80] Bahns A.J., Pickett K.M., Crandall C.S., Social ecology of similarity: Big schools, small schools and social relationships, Group Processes & Intergroup Relations, 15(1), 2012, 119–131
[81] Chen T., He L., Collaborative filtering based on demographic attribute vector, in 2009 ETP International Conference on Future Computer and Communication, IEEE, 2009, 225–229
[82] Vozalis M.G., Margaritis K.G., Using SVD and demographic data for the enhancement of generalized collaborative filtering, Information Sciences, 177(15), 2007, 3017–3037
[83] Mazhari S., Fakhrahmad S.M., Sadeghbeygi H., A user-profile-based friendship recommendation solution in social networks, Journal of Information Science, 41(3), 2015, 284–295
[84] MoocData, http://moocdata.cn/data/user-activity, (Accessed on 05/29/2020)
[85] Li M., Zhou Z.H., SETRED: Self-training with editing, in Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2005, 611–621
[86] Nartey O.T., Yang G., Wu J., Asare S.K., Semi-Supervised Learning for Fine-Grained Classification with Self-Training, IEEE Access, 2019
[87] McClosky D., Charniak E., 0001 M.J., Effective Self-Training for Parsing, in R.C. Moore, J.A. Bilmes, J. Chu-Carroll, M. Sanderson, eds., Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 4-9, 2006, New York, New York, USA, The Association for Computational Linguistics, 2006
[88] Chollet F., et al., Keras, https://keras.io, 2015
[89] Breiman L., Random forests, Machine learning, 45(1), 2001, 5–32