

Behavior of Online Learner in MOOC Platform Leading to Course Drop-out

Project Structure, Features and Model

Sitdhibong Laokok

April 19, 2021

Srinakharinwirot University



Dataset

This dataset has information collected from learner in MOOC platform with course information, enrollment, learner's course participations, and enrollment labelling for indicates drop-out status. The dataset separated in to 4 files to contain an information above.

Dataset Description

Each line contain the timespan of each course in logged information. The timespan of each course for calculating dropouts is 10 days after the last day of the course, i.e., course C if from 01/04/2014 to 30/04/2014 in the given data, a user enrolled the corse C will be treated as a drop-out if he/she leaves no record from 01/05/2015 to 10/05/2014.

Table 1: Course's timespan

Field name	Description
course_id	Course id
from	The first day of the course in log
to	The last day of the course in log

each line in this file describes a module in a course with its category, children objects and release time. Those modules represent different online materials of the courses, e.g., chapters, videos, problem sets and etc. The modules are organized as a tree, i.e., each course contains several chapters; each chapter contains several sections; and each section contains several objects (videos, problem sets, and etc).

Table 2: Course's object

Field name	Description
course_id	Course id
module_id	The ID of a course module
category	The category of the course module
children	The children modules of the course module
start	The time that the module was released to students

Each line is a course enrollment record with an enrollment id, a username U and a course id C, indicating that U enrolled in course C.

Table 3: Course Enrollment Information

Field name	Description
enrollment_id	Enrollment ID
username	Student ID
course_id	Course ID

Course Action Logs

Each line is a behavior record. Each event contains the following information: enrollment_id, username, course_id, time, source (server or browser), event, and object.

Table 4: Action Logs

Field name	Description
enrollment_id	Enrollment ID
username	Time of the event
source	Event source (Server or Browser)
event	In terms of event type, there are 7 different event types
object	The object the student access or navigate to

From Table 4, There are 7 type fo event collected from learner's participation in course. All possible values are shown in Table 5

Table 5: Events

Event Id	Description
1	Problem: Working on course assignments
2	Video: Watching course videos
3	Access: Accessing other courseobject except videos and assignment
4	Wiki: Accessing the course wiki
5	Discussion: Access the course forum
6	Navigate: Navigating to the part of the course
7	Page close: Closing the web page

Each line contains information about the ground truth of enrollments in the training set.

Table 6: Enrollment Labelling

Field name	Description
enrollment_id	Enrollment ID
status	Study status 1 for dropped-out, 2 for otherwise

Data Modeling

Data Modeling

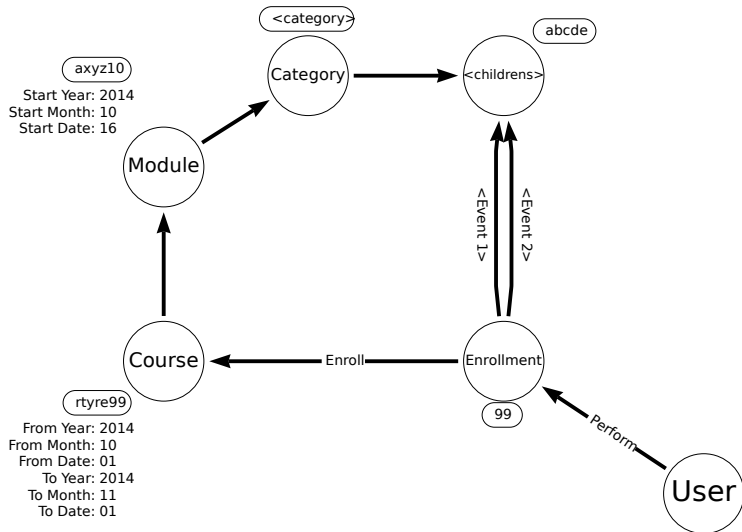


Figure 1: Data model