

## A Linear Ensemble of Classification Models with Novel Backward Cumulative Features for MOOC Dropout Prediction

Chih-Ming Chen, Man-Kwan Shan, Ming-Feng Tsai, Yi-Hsuan Yang, Hsin-Ping Chen, Pei-Wen Yeh, and Sin-Ya Peng



Department of Computer Science, National Chengchi University



Research Center for Information Technology Innovation, Academia Sinica

## Background:

Students' high dropout rate on MOOC platforms has been heavily criticized, and predicting their likelihood of dropout would be useful for maintaining and encouraging students' learning activities. In KDD Cup 2015, we predict

## **Competition Description:**

The competition participants need to predict whether a user will drop a course within next 10 days based on his or her prior activities. If a user U leaves no records for course C in the log during the next 10 days, we define it



DeepWalk on MOOC Dataset:

alleviate the feature sparsity problem

**Treat a random walk as** 

Backward Cumulative Features:

10/20 10/21 10/22 10/23 10/24 10/25 10/26 10/27 Ui X 0



© Copyright 2015 CLIPLab@NCCU, MACLab@AS.CITI