

## **Project proposal: Detecting dropouts in MOOCs using machine learning techniques**

In recent years, MOOCs have gained extreme popularity. Many MOOC platforms have been developed, some targeting global audiences and some targeting specific niches. Coursera as one of the pioneers shows that 61% of users who completed at least one MOOC reported educational benefits and high 72% of people reported career benefits. Knowing these numbers, it comes as a surprise to hear that the normal dropout rate in MOOCs on different platforms is 95%. These 2 statistics are pretty contradictory. One would expect that the high number of people with educational and career benefits would be motivation for more people to complete courses, but that certainly is not the case.

The topic here worth examining is the reason behind the high dropout rate. Is there a way for the course organizer to predict who the dropouts may be? Or at least how many to expect? Is there a way to decrease this number?

By analyzing learning activity data provided by platforms offering MOOCs, we can group the users study behaviour into a couple of categories. We can then analyze whether dropout rates are high due to some course characteristics, or is it the student's behaviour and personality. Based on the findings we would be able to train a model to predict possible dropouts early in the course so that some "nurturing activity" can be used for such students to prevent the dropout.

The analysis and the machine learning techniques that will be used for training the model will be decided upon deciding which dataset will be used. Below are a couple of links where different MOOC datasets are used and also a couple of previous similar research.

### Datasets:

1. <https://snap.stanford.edu/data/act-mooc.html>
2. <https://pslcdatashop.web.cmu.edu/>
3. <https://www.kaggle.com/kanikararang94/mooc-dataset>
4. <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZZH3UB>
5. <https://opendata.stackexchange.com/questions/1367/has-any-mooc-coursera-edx-udacity-or-others-publicly-released-some-of-their>

### Related work:

1. <http://moocdata.cn/publications>
2. <http://ifs.aminer.cn/misc/moocdata/publications/AAAI19-Feng-dropout-moocs.pdf>
3. <https://ieeexplore.ieee.org/abstract/document/8363340>