
Users behavioural inference with Markovian decision process and active learning

Firas Jarboui^{1,2}, Vincent Rocchisani², and Wilfried Kirchenmann²

¹ ENSTA, France and ENIT, Tunisia

² ANEO, Boulogne Billancourt, France

{fjarboui,vrocchisani,wkirschenmann}@aneo.fr

1 Introduction

Studies on Massive Open Online Courses (MOOCs) users discuss the existence of typical profiles and their impact on the learning process of students. One of the concerns when creating a new MOOC is knowing how the users behave when going through the contents. We can identify either quantitative methods that allow you to infer hardly interpretable groups of similar behaviour[1] or hardly context-transposable qualitative methods[2]. Our ambition is to find an efficient way to identify the behavioural pattern of interest to a given human expert. Within the *#MOOCLive* project³, we developed a mix-method to match the quantitative interpretation to the context needs.

2 Methodology

We tackled the following three problems in order to achieve our goal.

- The definition of a quantitative metric to compare behaviours
- The inference of qualitative sets behaviours from existing ones and test their reliability for describing the reality.
- The convergence between the quantitative-based clustering and the qualitative sets of behaviour to classify the users accordingly

In order to achieve our goal we define three main tasks. We start by quantifying the interest of users for the platform’s activity. This will allow us to define a distance between their behaviours. Then, we iteratively make hypothetical class definitions and test how well they fit the existing population. This is repeated until both the classifier and the classes suggested by the process are deemed satisfactory. This process’ breakdown is represented in *Fig. 1*.

1. **Quantitative modelling of the user:** We define the structure of the MOOC as a Markovian decision process framework. Let H be the history of actions the user performed on the platform. We define the gain function \widehat{G}_H of a user as the expected value of a categorical soft-max probability distribution over SGF , a Sample from the space of all possible Gain Functions.

³ *#MOOCLive* Virchow-Villermé ANR-15-IDFN-0003-04

The value associated to each element of this sample is the sum of rewards that the user's action would yield under the given gain function. This is thoroughly discussed in [3]. Each user is then characterised by the expected utility of each state with a discount factor γ .

$$\left\{ \begin{array}{l} U(G|H) = \sum_{a \in H} G(a) \\ \mathcal{P}(G|H) = \frac{e^{U(G|H)}}{\sum_{G' \in SGF} e^{U(G'|H)}} \end{array} \right\} \Rightarrow \widehat{G}_H = \sum_{G \in SGF} G \times \mathcal{P}(G|H)$$

2. **Qualitative class definition:** This step is purely human. The experts are asked to interfere and define the classes that will be used to build the quantitative classification. In this stage, the expert intervention is purely based on his *a priori*. If the expert's *a priori* is invalidated during the process, he will have to restart from here with an updated point of view.
3. **Fitting the classification:** To have well classified users a Gaussian kernel label propagation is used. This provides a probability distribution of membership to each pattern for each behaviour. An active learning process is used to iterate the propagation of the labels under the supervision of the human expert. After each fold, we sample the users randomly and test if the output probability distribution makes sense. The human expert either **agrees** with the results, **changes** them or tags them as **unsure**.

If the rate of changed results is high, we continue the active learning loop. As a result, the rate of bad labels will decay.

Once the classifier stabilizes, we consider the rate of behaviours that the expert tagged as unsure. *If this exceeds a threshold*, we roll back to the second step to challenge the *a priori* class definitions.

If the unsure tags rate is low enough, we can safely assume that the two models converged with respect to the expert.

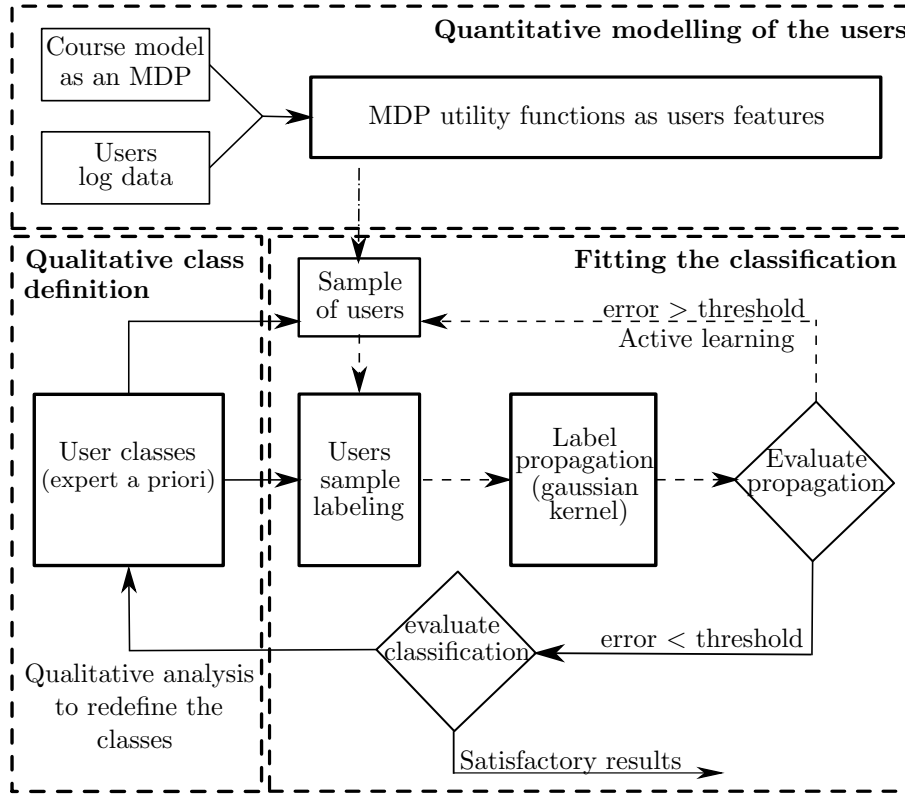
We applied this methodology on a MOOC⁴ with a sociologist. We started with an *a priori* of three user profiles. Up to this date, after three iteration of the methodology, we were able to identify seven profiles that fulfil the context needs and to classify the users accordingly.

3 Conclusion

Our method assists a human expert to find the optimal information about the studied population. Although this work is still in progress and only tested on MOOC log data, it should be applicable on other log data streams of information. Future tests will involve marketing related data. We are currently investigating the efficiency of this method as well as the best techniques to use for each step. This is part of a preliminary work for a thesis.

⁴ <https://www.fun-mooc.fr/courses/VirchowVillerme/06005/session01/about>

Fig. 1. Users behavioural inference process



References

1. Chase Geigle and Cheng Xiang Zhai: Modelling MOOC Student Behaviour With Two-Layer Hidden Markov Models. Learning at Scale (2017)
2. Paula de Barba Carleton Corin, Linda Corrin and Gregor Kennedy: Visualizing patterns of student engagement and performance in moocs. (2014)
3. Constantin A. Rothkopf and Christos Dimitrakakis: Preference Elicitation and Inverse Reinforcement Learning. cornell university library (2011)