To define

COMPUTER-HUMAN INTERACTION IN LEARNING AND INSTRUCTION CHILI

École Polytechnique Fédérale de Lausanne

Farah Bouassida,[*]

Supervisor: Lukasz Kidzinski[†]

Supervisor: Pierre Dillenbourg[‡]

December 27, 2015

[*]Communication systems' Masters farah.bouassida@epfl.ch

[†]lukasz.kidzinski@epfl.ch

[‡]pierre.dillenbourg@epfl.ch

Abstract

The goal of this project was to implement a web platform that teaches a set of concepts through different kind of material and to gather different information about the learner. The different learning paths are a video on a first concept than a video on the second concept, or a video than a text or a text than a video or two consecutive texts on both teacher concepts. The project's purpose is to demonstrate a possible correlation between the learner available information and his learning gain, in order to build afterwards, an adaptive learning platform.

Contents

1 Introduction

This project is trying to establish a relationship between a learner profile and the learning gain he achieves depending on the different style of teaching activities he takes. Later, this could be used to construct an adaptive learning platform based on the psychological features of the learner. For this project we tested the difference between two equivalent explanations of the same concept one presented with a text in an encyclopedia fashion and the other presented with a video selected from a MOOC. Previous research have proven that we can classify learners based on common sense, such as high-aptitude versus low-aptitude, highly motivated versus poorly motivated. Each category of learner is then proposed a different learning strategy and psychological tests are available to identify the different styles of learners. But the goal would be to find automatically the best teaching material corresponding to each style and to widen the understanding of learners' profile. In fact, thanks to machine learning we can classify learners based on wider criteria as we collect information on different aspects of a person and we test the correlation of aa combination of them or a subset of them to predict the learning gain he would achieve with a given learning path.

2 Learning platform

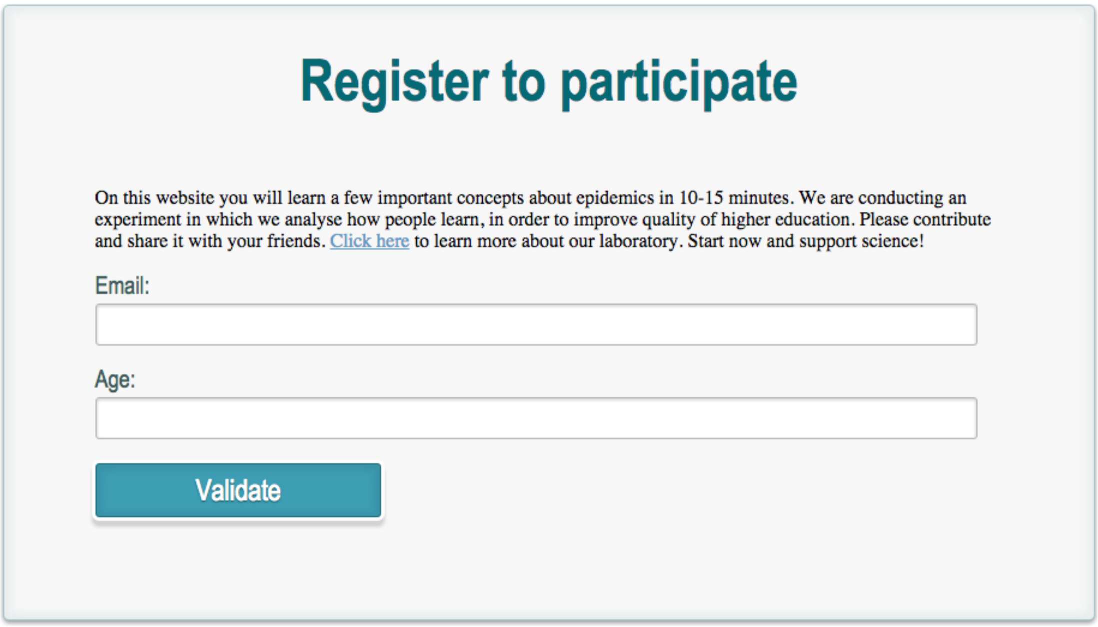


Figure 1: Learning interface

2.1 Psychological test interface

In order to gather the data for our analysis, we had to develop a learning web platform[1] coded in Django. In the following sections we would call "user" a person that participates to the experiment.

To that purpose, we tried a set of different questions that could be related to different psychological aspects of a person. We list them below together with the corresponding rating system:

- Two questions derived from IQ tests that would be grouped together in a variable named "IQ". -1 was attributed to any wrong answer and +1 was attributed to the good answer.

- Ten personality questions from the International Personality Item Pool [2]. Five questions are related to the conscientiousness of a person and five questions related to the openness of a person. In each case, three items were positively marked and two negatively marked. For this part of the test, the answer varied from "I strongly disagree" to "I strongly agree" and the points ranged from 1 to 5 or -1 to -5 in the case we gave negative marks. The result for the consciousness was stored in the column named "conscious" whereas the openness' results was stored in a variable named "open".

- Five questions related to the logic and the imagination capacity of a person. These five questions were taken from the paper folding test [3]. +1 was attributed to to the right answer while -1 was attributed to a false one. The results were stored as a variable named "paperFolding".

  Besides some informations about the person is gathered with the following questions:

- One question about the academic level of the test taker. The possibilities to pick from were: "Haven't graduated high school", "High school graduate", "Apprenticeship", "College student", "Bachelors", "Masters", "Doctorate". The answers in this order were assigned a number from 1 to 7. The results were saved as a variable named "level" and treated as a categorical variable in the data mining process.

- One question about the gender of the person. We asked if the person was a Male or a Female during the first test the user had and the answers were stored in the column named "gender".

- One question asking in which country the test taker is living.

- One question about the age of the person saved as the variable "age".

[1] https://github.com/chili-epfl/learning-platform
[2] http://ipip.ori.org The Big-Five Factor Structure
[3] http://steinhardtapps.es.its.nyu.edu/create/assessment/vz2/start.html

One other implied feature when taking the test, could be the time a person takes to accomplish a test or an activity or the whole experiment.

The experiment is a website that guides you through different steps in a precise order explained later in section 2: "Learning platform". The goal for this website is to teach a new concept through different paths assigned randomly and in a uniform way to each visitor of the page.

At the end of the experiment, we also ask questions about how the test went and how people felt about it to have additional information that could infer relevant features.

The user of the page, is asked the same set of questions about the theme of epidemics before and after he learns about it so that we could calculate the learning gain of a person as the difference between the scores he gets at the two tests.

2.2 Website Hierarchy

1. The registration form: It is the first step of the experiment where the user needs to register entering his email and his age. The email is an assertion that the user has a unique identifier. A user could take only once the test in order to measure his learning gain when he goes only once though the teaching activities. Note that the user is free to have the learning activities more than one time during this one time experiment meaning that he could watch a video more than once while in the video page, pause it, move forward etc...
   Besides, doing the experiment only one time permits us to measure the time it takes a person to go through all the steps and is also used to get variable for the data mining.

2. The psychological test: The questions of this test are enumerated in the introduction and the answer of the user to them constitues the features that would help us to search for patterns among people and to imply the best path for them to improve their learning experience.

3. The pretest: It is the same as the post test and is composed of six questions with different degrees of difficulty about "Epidemics", the theme treated in the learning activities of this experiment.

4. The activities: There is two successive activities each one aims to teach one concept. The first concept is the epidemic's definition and causes and the second one is the epidemic's reproductive number. Each concept could be explained with a video or a text. In that way, we had four different teaching paths: text-text, text-video, video-text, text-text.

5. The post test: see pretest.

6. The assessment form: It is composed of some of the NASA-TLX [4] questions and gathering the opinion of the user about different aspects of the experi-

[4] http://humansystems.arc.nasa.gov/groups/tlx/

ment taken. This questionnaire is known as an assessment tool that rates the perceived workload in order to assess a task and the answers could also infer useful features to choose the best learning path of the user.

3 Data Description

We gathered a total of 167 participants but only 77 of them accomplished both pre and post tests and went through all the experiment steps (from the registration to the assessment form). As our goal is related to the learning gain, we work with these 77 users. Among them, we have 25 female user and 52 male user. We have also the following distribution for the possible learning paths:

- 27 users for text-text.
- 20 users for text-video.
- 12 users for video-text.
- 18 users for video-video.

4 Data visualization

As we have few observations, we try to limit ourselves to a reasonable number of variables to avoid building bad models of regression due to the curse of dimensionality. We choose 15 variables among them 13 are continuous and two are categorical. We start by visualizing the box plot of the learning gain as a function of the learning path showed in figure 2. We observe that we cannot see a significant change in the learning gain that could support the hypothesis that one path is always the best for all types of learners.

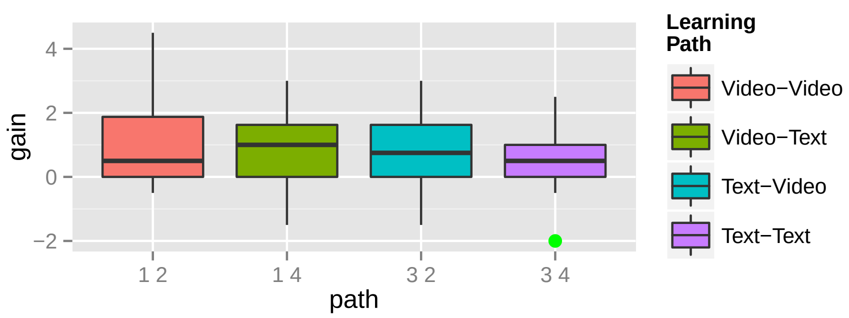


Figure 2: Learning gain according to the learning path

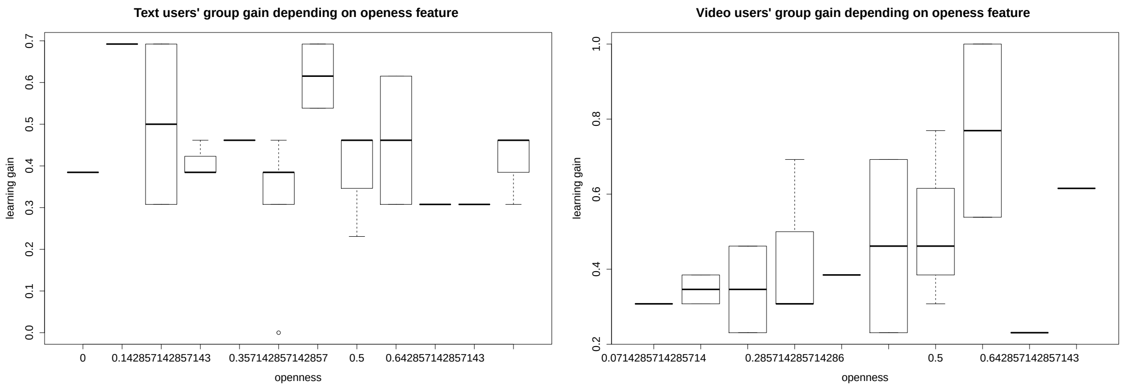
We try now to see if the learning gain is correlated to some feature or some subset of features and if this also depends on the path proposed to the learner. For this experiment, we didn't try to use predefined categories of learners. Our approach was to collect as much data on the user's learning behavior as possible. To select these features we used different human-defined semantics for psychological aspects but there without assumed correlation between them and semantics and the demographic factors of the learner. We then tried plotting each feature against the learning gain achieved by the user and compare it for the for possible paths that a user could get. Figure 3 show the box plot of the learning gain depending on the gender. The results cannot be fully trusted to make direct comparison between genders because of the high proportion of males regarding to the number of female participants. But we can see that, within our observations, the video-video path and the text-video path worked better for females whereas video-text path was more successful around males.

Figure 3: Learning gain according to the learning path

Thereafter, we tried to focus only on the video-video path and the text-text path to make the comparions and find feature correlations with the learning gain. In some cases we can observe a potential relation between the variable and the obtained learning gain. For instance, we can see that in figure 4a representing the case of the text path, the learning gain seems not to be impacted by the openness of the person whereas for the video path showed in figure 4b we get better results the more the user is open.

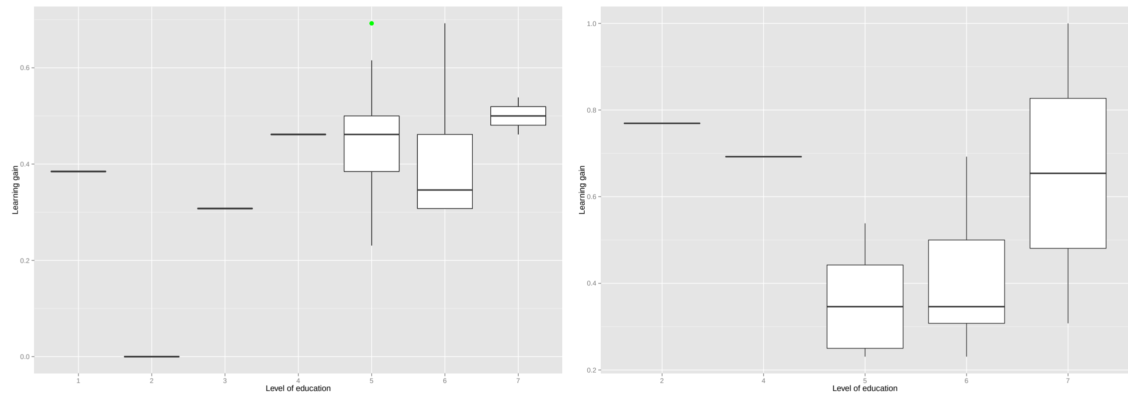
We can also observe in figures 5a and 5b that the highest the educational level of the user is the better his learning gain is, and in particular, in the case of video only learning material people with a doctorate perform much better than the other categories. Similarly, for the users that are assigned only video material, we have a correlation of 0.4986 between the time they spend watching the first video and their learning gain. This result is quite logical as people that doesn't watch the whole video won't achieve good results in the post test.

In contrast, other variables as the consciousness of the user doesn't seem to affect the learning gain. Thus, we could think that some psychological features could indicate a difference in the learning experience of the user depending on the activities he undertake to learn the new concepts. Demographic factors as the gender or the level of studies of the user could also influence his learning experience.



(a) Users given a text path **(b)** Users given a video path

Figure 4: Learning gain depending on openness



(a) Users with text only material **(b)** Users with video only material

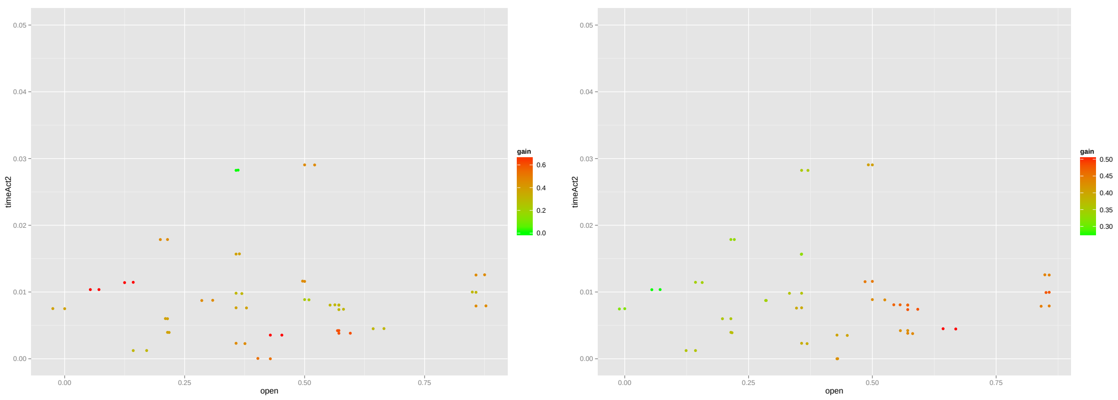
Figure 5: Learning gain depending on the achieved level of studies.

5 Building a regression model

In this step of the project, despite of the number of complete observations (77 complete experiment results), we try to build a regression model with machine learning algorithms to predict the learning gain of a user. If we achieve a good model, we could prove that, with the gathered user information of our experiment, we could compare the learning gain that a person would have with the different possible paths. Then, in a second step, we can map back the results to the features' space to see if there is some regions performing better for one path more than another.

Before applying machine learning algorithms, we scale the data with the minimum and maximum values to obtain, in the case of the continuous variables, values between 0 and 1. The level and gender variables are considered as factors in R.

We tried then the Support Vector Machines (SVM) regression technique with the package e1071 in R. SVM provides the usage of kernels what allow us to have a non-linear regression. The radial kernel had performed better for our regression model. After the tuning of the parameters, we obtain a Root Mean Squared Error (RMSE) of 0.14 for the text only path and of 0.2 for the video only path. We used a radial kernel and performed a 5-fold Cross Validation (CV) with 80% of the observations used for the train set and 20% of the observations for the test set. With dummy encoding of the categorical variable for SVM use we don't achieve better results. We try then to perform simple feature selection, taking only variables that have correlation above 0.1 with the learning gain taking the absolute values of the correlations. We again don't see an improvement in the predictions. We can think that the poor prediction results are due to the lack of data as we have only 27 observations for text-only path and 18 observations for video only path. The fact also that we base our prediction on human variables could also introduce some unpredictability to our regression. Nonetheless, we use the model built for video only path to predict the learning gain of the users that had the text only path and we can observe a change for certain users. In fact, when we plot the learning gain for a pair of variables as the variable "open" that indicates the openness of a person and the variable "timeAct2" for the time a person took for the second activity, certain points have different learning gain. In fact, figure 6b in comparison with figure 6a shows that people with high openness would have higher learning gain with video path while people with very low openness perform better with text material. This could confirm what we saw as correlation between the openness of the user and his learning gain and could suggest that we can adapt the activities proposed to the learner depending on this psychological feature.



(a) Users with text only material

(b) Users with video only material

Figure 6: Comparison of the gain of the users that had text material with their estimated gain with video material.

6 Conclusion