Using Data Acquired from Learning Management Systems to Predict Examination Results

Francis B. Lavoie\textsuperscript{1}, Pierre Proulx\textsuperscript{1}, Ryan Gosselin\textsuperscript{1}
\textsuperscript{1}Département de génie chimique et de génie biotechnologique, Université de Sherbrooke
Francis.B.Lavoie@USherbrooke.ca

Abstract – This work presents a novel learning management system (LMS), named Catalyseur, which allows the instructors to easily visualize which lessons and exercises allowed the students to better perform at an examination. This LMS feature is based on a regression methodology calculating easy-to-analyze models and being able to fit dynamic relationships. These models are calculated automatically and only require as human input to upload the student results at an examination.

1. INTRODUCTION

In a typical engineering course, several lectures and exercises are proposed to the students. It is however difficult to measure the efficiency of each pedagogic material in student success even if all the information is available. Indeed, learning management systems (LMS), such as Moodle, generate a large amount of data about student browsing and material access. It should then theoretically be possible to link these data to student examination results and retrieve the relevant provided materials leading to student success. However, creating such correlations can be cumbersome and may require several hours of work. Consequently, studies are generally performed by only correlating one or a few formative assessments with examination results \cite{1}. Only a small amount of all available information is used in such analysis.

To resolve this issue, we have developed an algorithm being able to automatically reveal pedagogic materials having contributed to student success as well as being able to retrieve the maximum time, before an examination, the completion of an exercise can be beneficial. This algorithm is in fact an adaptation of an algorithm initially developed for continuous industrial process analyses in which we modified a few constraints.

2. REQUIRED DATA

The algorithm requires data about students’ completion of lessons and exercises. These data are generated by the Catalyseur LMS, developed by the Department of Chemical and Biotechnological Engineering of the Université de Sherbrooke, in which each lesson and exercise has its own page. In each of these pages, the students are invited to click on a status button in order to indicate that the lesson is understood or that an exercise is completed. The students are encouraged to click on these buttons as the Catalyseur LMS

![Figure 1: (a) Principal course page in the Catalyseur LMS in which a polar area graphic representing a student completion through the different course topics is automatically generated. (b) Doughnut graphic representing a student completion of exercises](image-url)
automatically creates them completion graphics and statistics (Figure 1).
The algorithm requires a binary 3D $X$ matrix automatically generated by the LMS (Figure 2). This binary matrix contains information about the completion of exercises and lessons at different times before the examination, for each student. A value of “1” indicates that the related exercise or lesson was completed for the corresponding student and for the corresponding time before the examination, while a value of “0” indicates the opposite. The algorithm also requires a $y$ vector containing student evaluation results (Figure 2). This vector must be uploaded to the LMS by the instructor.

Once these data have been provided to the algorithm, no additional manipulations are required. The algorithm automatically searches which exercises and lessons completions (data in $X$) can predict the examination results (data in $y$). The algorithm also searches the maximum time before an evaluation at which the completion of an exercise or a lesson can improve student results at the examination.

3. CALCULATED MODELS

The results provided are staggering simple to analyze. As depicted in Figure 3a, some exercises are outlined by the algorithm, meaning that their completion had an important and positive impact in examination results. With these algorithm outputs, the instructors have additional information while making changes in course improvements. Indeed, non-relevant exercises and lessons can be removed, and the emphasis can be put on pedagogic materials made significant by the algorithm. The outlined exercises are also linked to a time dynamic (Figure 3b) which falls around 4 days before the examination. This dynamic means that if the outlined exercises were completed less than 4 days before the evaluation, the completion of the outlined exercises is considerably less beneficial for the examination success. This dynamic can be confirmed by similar studies stipulating that last-minute exercise completions largely increase anxiety [2], this mood being the first factor of performance decrements as it leads to the inability to suppress competing thoughts during the evaluation [3].

It is also possible to create models by associating the 3D $X$ matrix with a $y$ vector containing only the results for a specific part of the examination. From this, it is possible to analyze which exercises and lessons were relevant to perform adequately on a specific course topic evaluated during the examination.

![Figure 2: Required dataset for the algorithm](image)

![Figure 3:](image)

4. CONCLUSION

The algorithm was applied to several examinations in different engineering courses. An interesting trend is related to the extracted time dynamics which were similar between the different analyses. This then creates a pedagogic and psychologic study tool to eventually improve the performances of the students in their academic cursus. The developed algorithm is then a data analyzer helping the instructors to improve their courses as well as a tool for pedagogic studies.
References

