The Agile Learning Model: Using big data to personalise the acquisition of accounting skills

Brent Gregory
UNE Business School
University of New England

Matthew Wysel
UNE Business School
University of New England

Sue Gregory
School of Education
University of New England

Big data mirrors the accounting process to the extent that it deals with how we capture, categorise, summarise and report information so that users can make informed decisions. By modelling this process, we can both demonstrate the future of accounting to our students, and build an agile learning environment that identifies for a student their 'next crucial action' in the learning process. Presented in this paper is a pilot study.

Keywords: Agile learning, education big data, personalised learning, automated intervention

Introduction

Effective learning is predicated on feedback, as feedback enables the student to determine the effectiveness of their efforts in achieving the desired goal (Gregory, Uys & Gregory, 2014). This feedback can vary from: personalised to heuristic, continuous to discrete, adaptive to rigid, and is regarded as the key to self-regulated learning environments and the challenge in classical, didactic environments (Butler & Winne, 1995).

In more sophisticated and effective learning environments both the student and the teacher actively participate in a mutual process of feeding back and processing information regarding the efficacy of the learning process to one another, with the common goal of refining the process (Cantillon & Wood, 2011). Feedback, therefore, relies on the provision and recording of data, its categorisation and summarisation, and timely reporting to one or more parties within the learning process.

Historically, administrative overheads associated with evaluating the numbers of students typically found in school and university classes has enforced highly uniform, and therefore structured, evaluation processes with corresponding rigidity within tuition programs. In this case, institutions' operational requirements to process sufficient volumes of students drive an increase in curriculum rigidity, a reduction in the number of assessments – and therefore a widening of feedback loops, and consequentially force students to increase their reliance on heuristic techniques to guide them through their learning environment.

In order to develop our agile learning system a pilot study was undertaken through the selection of student activities, assessing learning and gathering feedback, we may need to step back from of our existing operational assumptions. This will enable us to adapt to realities such as the inputs to the system, (the students), are not homogeneous. A starting point may be to garner information about the student that does not come from the traditional assessment tasks. As well, we should develop resources that are outside the normal learning resources for a particular unit of study. One of our initial challenges will be to identify the questions that we need to ask.

This suite of challenges is precisely where big data is most powerful. The commercial utilisation of big data hinges on the ability to collect large quantities of discrete sample points, assemble them for analysis and assessment and respond to the aggregated information in a timely manner. Typically, the commercial collection of sample points within big data applications involves an attempt to garner high quality information about the true position, activities, or knowledge of the subject, while minimising the impact of external influences associated with the assessment. Moreover, the systems created to capture and respond to the collection of big data are designed to retain sufficient agility to deliver personalised content to individuals based on the statistical correlation of one subjects' current position to numbers of previous subjects' next successful actions. Big data provides schools and universities the framework to: leverage data created by large numbers of students in the past;

facilitate sufficient continuity in assessment of current students to minimise external factors introduced by infrequent and high-value assessments; and provide the means by which a curriculum can be tailored to the point of personalisation.

Big data is the existence of large quantities of discrete sample points, observed by a unified system, and assembled to yield the potential for analysis either between different sample points or across equivalent sample points at different time intervals. The U.S. National Institute of Standards and Technology summarises these same elements as volume, variety, velocity and variability (U.S. Department of Commerce, 2015) and they acknowledge that there are a range of working definitions, including the view of Drew "Big data, which started as a technological innovation in distributed computing, is now a cultural movement by which we continue to discover how humanity interacts with the world—and each other" (Drew as quoted in U.S. Department of Commerce, 2015, p. 11).

Background

AFM101 is the introductory accounting unit that is core to all business awards at the University of New England. As a result, it has a diverse student cohort and has been a unit that has experienced high attrition and failure. Typically, more than 50% of students that enroll in the unit do not successfully complete the unit. A range of data is used in order to match a particular student with the resources that will be most useful to them. Over recent years, a significant set of resources to solve particular learning problems has been developed. However, the general development of resources has been ad hoc rather than structured. This means benefits to students are piecemeal. To provide an example of the process, we will now peer through a narrow window of time and data collection points through a pilot study.

Big data mirrors the accounting process to the extent that it deals with how we capture, categorise, summarise and report information so that users can make informed decisions. We have attempted to model this process not only as a demonstration of accounting but also to help us develop an agile learning system. Data is collected from assessment tasks and from other sources as well, including student questionnaires. A variety of vehicles are used to process the data (Moodle [Learning Management System], Excel, Qualtrics) and the results of the processing are imported to a central database (Excel). Reporting occurs on a number of levels and the key reporting aims to identify the resources that are likely to be most effective for a particular student and then guide the student to their next crucial action (NCA). When that is completed, the process reports the NCA. The process currently requires substantial human intervention and would benefit both students and the teaching staff if the systems were sufficiently compatible to improve the level of automation.

Several other education systems have adapted to employ some version of an agile learning path (see, for example, http://www.lynda.com and http://www.khanacadamy.org), but noticeably traditional tertiary institutions have been slow to adopt this method. The growth in sophistication and adoption of distance-based learning has both undercut the ad hoc personalisation of education through socialisation that formally happened during tutorials or small-group meetings, and enabled a new breed of personalisation, centered around flexible learning patterns. Fundamentally, this personalisation of curriculums leverages the inherent agility facilitated by the application of big data to large quantities of data.

Our progress to the agile learning system has been a gradual process. In this paper we highlight using components of the big data model to personalise the student learning experience. While students produce torrents of data as a by-product of their ordinary operations, and each of us is now a walking data generator (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012), the key discipline is still collecting the data. In traditional data collection models, we collect and store data primarily around final results in assessment tasks. However, the increased opportunity for data collection, analysis, reporting and responding that online computer systems have created now enables us to collect data before, during and after the trimester. Using these additional data points, we are able to refine both our understanding of student's needs and facilitate the management of these needs. In this paper we will focus on data associated with students' basic math skills and their learning style as categorized by a representative systems bias test and the conclusions and potential interventions that can be automated as a result of this granular knowledge. Following we will describe the recent history and the consequences of adding these extra data points.

Preliminary results

Methodology

A pilot study was undertaken over two trimesters to see if interventions implemented in the second iteration of the teaching were successful, through the comparison of two cohorts of students, in 2014 and again in 2015, through a total of 332 students (155 in 2014 and 217 in 2015). All students were studying in off-campus mode (i.e., online, from a distance).

2014 Trimester 2 (115 off campus students)
At the commencement of each trimester, students were provided with the opportunity to participate in two diagnostic tests: The Diagnostic Math Test (DMT) and the Representational Systems Biases Test (RSBT). The results of these tests are quantiatively assessed and used to personalise the learning of particular categories of students. Of particular interest are those students whose DMT and RSBT results align with the results of previous generations of students who went on to fail or under-perform in the unit. As more data collection points were added, there was the capacity to increase the level of personalisation and insight. This was especially true as we combined data points from different aspects and contexts within the same sample set.

The DMT comprises of 10 reasonably straight forward math questions and generally takes around five to seven minutes to complete. Each question is awarded 10 marks. The RSBT comprised of five questions which provide some indication about the person's preferred way to represent information (visual, auditory, kinaesthetic or auditory digital).

Based on the results achieved in these tests, students were then guided toward different resource sets. To this stage, the personalisation has been unsophisticated. A student with a low DMT score (less than 8/10) was directed to math resources and other learning resources were highlighted as available to students depending upon their RSBT results.

2015 Trimester 1 (217 Off Campus Students)
The same process was applied in the following trimester that this unit was taught, in 2015. However, based on outcomes from 2014 Trimester 2, two significant interventions were introduced:

1. The development of a Math Help Area on Moodle which was made up of a selection of basic math videos. All students that scored 70 or less in the DMT were directed to the Math Help area. Students who scored 80 were made aware of these resources in case they chose to utilise them; and,
2. A slightly more proactive approach identified to students at risk. This especially related to student non-participation in these activities, initially by automatically generated personalised emails and built to a phone call from an academic.

Results

2014 trimester 2 (115 off campus students)

The results of these diagnostic tests have strong predictive value in identifying students who could excel and those that may struggle. Table 1 indicates the key role math skills play in students succeeding in the unit. It provides an analysis of the DMT results for 115 off campus students in Trimester 2, 2014 (the full cohort that were enrolled). Students were divided into three categories based on their final grade in the unit. 'Not Succeed', includes all students that failed or did not complete the unit. 'Pass' includes all students who receive a grade of pass or credit and students who receive a distinction or better made up the 'Excel' group. The table reports student categories based on their score in the DMT. For instance, 84% of students that did not complete the DMT did not pass the unit. Whereas, for the students that scored 100% in the DMT, 39% did not complete the unit and 32% excelled.

Table 1: Analysis of result in Diagnostic Math Test
Unit result organised by score in DMT

2014, Trimester 2, Off Campus (n=115)			
Score	Excel	Pass	Not Succeed
#NA	0%	16%	84%
<70	0%	33%	67%
70	0%	60%	40%
80	25%	25%	50%
90	17%	46%	38%
100	32%	29%	39%

2015, Trimester 1, Off Campus (n=217)			
Score	Excel	Pass	Not Succeed
#NA	9%	22%	69%
<70	0%	67%	33%
70	38%	38%	25%
80	13%	63%	25%
90	23%	39%	39%
100	35%	31%	34%

When the results are overlaid for the RSBT on the 'Not Succeed' group, it was found that 75% of failures were made up of students that did not complete the RSBT. At the other end of the spectrum, students who scored full marks in the DMT were identified as 'auditory digital' in the RSBT. Only 18% of this group did not pass the unit and 43% appeared in the 'excel' category. We can further improve the predictive value when we overlay other data points (repeat student, award enrolled, phone number available). These results also highlight that our ability to predict outcomes was superior to our ability to bring positive change, which in turn implied a need to review the existing intervention strategies.

2015 Trimester 1 (217 off campus students)

A similar pattern of results occurred in Trimester 1, 2015 as displayed in Table 1 (2015). However there are some differences. When compared to 2014, we notice a general improvement in grades, with a substantial improvement in results for students with scores of 70 or below in the DMT. This group were directed to the Math Help resources. During the trimester, six videos were viewed a total of 130 times for a total of 194 minutes. However, no data was recorded on the extent to which individual students took up these resources. While it is possible that the math help resources made a contribution, the extent of that contribution may not be significant. For instance, all students who scored 70 in the DMT and still achieved a HD also identified as 'auditory digital' in the RSBT. So other factors could also at play.

Limitations of the study

The small sample size used means the value of this study is more as a pilot program and preparing for the future than in reporting outcomes that we can confidently act upon. Reporting the outcomes were also challenging because of the multitude of ways that the data can be presented.

Further research

In the immediate short term, we will extend the size of the database to include all students that have enrolled in the unit in the last four years and include all of the data points that have been collected. As well, we will investigate the utilisation of commercially available, large, unstructured database platforms to host and analyse all data which will support a more rigorous pattern analysis.

Findings and conclusion

Using principles from big data, we can develop an agile learning system that will enable a more personalised learning path for individuals. The appropriate use of data can signpost students to their next crucial action. Key to our future success will be how we better identify, collect and index the most useful data. This data must extend beyond results of assessment tasks. While the rich combinations of the data provides a challenge in global reporting, it does support precision in guiding the students' study path. There is also a challenge in either developing or identifying the resource warehouse needed to support students and then in mapping the various paths through the resources.

Asking the right questions will be key to developing an agile learning system. To date, the identification of these questions has been a cyclic process. For instance, we find a pattern in the data (e.g. students that were directed to the math help resources did better than expected), but when we attempted to find what caused that pattern we found that we had not collected the data that would most help us identify the causes (to what extent did they use the resources? did those resources help?). In the words of Google's director of research, Peter Norvig "We don't have better algorithms. We just have more data" (McAffee et al., 2012, p. 63). We need not only more data, but we need the appropriate data. And, we need a unified system that will enable us to work with that data. Finally, when the data has identified what action is needed, we need the resources that will support the specific needs of each student. This requires the development, warehousing and indexing of the resources. Success requires cooperation – particularly between those with access to the data and those with access to the analytical tools and platforms.

References

Butler, D., & Winne, P. (1995), Feedback and self-regulated learning: A theoretical synthesis, *Review of Educational Research, 64*(3), 245–281. https://doi.org/10.3102/00346543065003245

Cantillon, P., & Wood, D. (2011). *ABC of Learning and Teaching in Medicine*: Wiley.

Gregory, B., Uys, P., & Gregory, S. (2014). The role of instant feedback in improving student understanding of basic accounting concepts. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology*. Proceedings ascilite Dunedin 2014 (pp. 634–637). https://app.box.com/s/016cdyv8dq1pp0yhp1vw/3/2704865194/23032568124/1

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. The management revolution. *Harvard Bus Rev, 90*(10), 61–67.

U.S. Department of Commerce. (2015). *Draft NIST Big Data Interoperability Framework: Volume 1, Definitions* (pp. 1–33). Gaithersburg, MD. Retrieved from http://dx.doi.org/10.6028/nist.sp.1500-1

Gregory, B., Wysel, M., & Gregory, S. (2015). The Agile Learning Model: Using big data to personalise the acquisition of accounting skills. In T. Reiners, B.R. von Konsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Globally connected, digitally enabled*. Proceedings ascilite 2015 in Perth (pp. 445-449). https://doi.org/10.14742/apubs.2015.940

Note: All published papers are refereed, having undergone a double-blind peer-review process.