Enhancing Student Success through Predictive Learning Analytics: Expanding the Learning Analytic Application at the University of Alberta

Ying Cui

Introduction

The latest Horizon Report Higher Education Edition (Johnson, Adams Becker, Cummins, Estrada, Freeman, & Hall, 2016) predicts that learning analytics will be increasingly adopted by higher education institutions across the globe in the near future to make use of student data gathered through online learning environment to improve, support and extend teaching and learning. The Horizon Report is an annual forecast of emerging technologies and their effects on education by the EDUCAUSE Learning Initiative (ELI) and the New Media Consortium (NMC). The 2016 Horizon report defines learning analytics as "an educational application of web analytics aimed at learner profiling, a process of gathering and analyzing details of individual student interactions in online learning activities". It can help to "build better pedagogies, empower active learning, target at-risk student populations, and assess factors affecting completion and student success".

University of Alberta is beginning to explore the use of learning analytics for improving student learning experiences (Shiri, 2016). A recent TLEF project (2014-2015), led by the co-PI of the proposed project, Dr. Ali Shiri, developed a learning analytics software application for the University of Alberta eClass learning management system to support the summary and visualization of student activities and interactions with eClass. This newly developed learning analytic application operates as an integral part of the University of Alberta eClass environment and is expected to become available to all instructors and students on campus Spring 2017. The application allows instructors and students to visualize and demonstrate individual and comparative views of student activities such as logins, submissions, resources accessed and frequency, and interaction with discussion forums, lessons, and quizzes. Demonstrations of the new learning analytic application are depicted in appendix A. The application provides basic descriptive analyses of student activities on eClass and serves as a starting point to making use of the large and diverse amount of student data collected through the University learning systems in supporting data-informed decision making and curricular improvement.

To fulfill the full potential of the learning analytics application, this two-year project proposes the next stage of the development in order to move from the focus on the summary of student data to more sophisticated and advanced predictive and statistical analyses in order to develop actionable information for the diagnosis and prediction of student outcome. As a result, the current proposal aims to build on, expand and enhance the functionality of the eClass learning analytic application by developing and building in a predictive student success algorithm. It will allow students to harness the power of learning analytics to monitor and self-regulate their own learning and allow instructors to identify students at risk as well as develop and implement effective learning interventions, and ultimately help students succeed.

Methodology

The goal of developing a predictive student success algorithm is to build statistical models that use multiple sources of information available within the University of Alberta systems to make predictions of how each student is doing and which students might be at risk and could benefit from additional help. The proposed algorithms may include the following components as the basis of the prediction: student performance as indicated by percentage of points earned to date, student activities and engagement in the course as reflected by student interaction with the eClass system, student demographic characteristics and past academic history. Based on these data types, statistical modelling techniques (e.g., cluster analysis, decision tree, logistic regression and support vector machine regression) will be utilized to categorize students into different risk levels. The use of multiple components provides a more holistic and complete picture of a student's efforts, learning experiences and background, which potentially improves the validity of predictions of student success. We have consulted with Diane Alguire, the director of the University Information and Privacy Office, and confirmed that the study can be conducted within the ethical and legal parameters that will balance privacy protection of students with information use in learning analytics.

The predictive algorithm will be built into the learning analytic application and be run on demand by students and instructors. For the purpose of privacy protection of student information, students and instructors will not have access to data on students' demographic background and past academic history. Rather, students and instructors will be provided with an estimate of the likelihood of student success. With this information, students can reflect on their learning and the learning process, and plan their future learning in a more proactive and structured manner. Instructors can use this information to identify students at risk. By further examining and comparing resource uses and activities of different risk groups, instructors can provide real-time feedback to the potential at risk students that a) outlines their current performance in the course, b) suggests students to change their behaviors (e.g., participate in group discussions or submit assignment on time) in order to increase their chance of success in the course, and c) encourages students to have face to face meetings with instructors or visit various student support centers on campus such as student success center or student accessibility services. This type of personalized learning advice is realized and supported by data-informed solutions and was previously not feasible in large class sizes. To achieve this, the project will focus on the following activities:

- An environmental scan of existing predictive algorithms of student success implemented in learning analytics applications in different higher education institutions
- Exploring the integration of data from various information systems within the university for more efficient use and analysis in learning analytics and evaluating different data de-identification techniques (Khalil & Ebner, 2016) for the purpose of privacy protection
- Design and development of a predictive algorithm based on available student data and examination of its accuracy, generalizability and predictive power through validation
- Exploring other useful features for students that could be included in the learning analytic application such as calendar reminders of assessment due dates and exams, the percentage of assessments the student has completed, and student ranking in the class
- Expanding the university eClass learning analytic application by adding the predictive algorithm and identified useful features

- Testing the expanded learning analytic application and evaluating its usability, usefulness and intuitiveness though the evaluation of student performance, focus group discussions, and student and instructor responses to questionnaires.

Evaluation

The participation and feedback of instructors and students at the University of Alberta will be essential in the design and modification of the expanded learning analytic application to make it a functional and easy-to-use tool for students and instructors. Two focus group discussion sessions that includes instructors and students across campus will be held. The first session, held at the early stage of the project, aims to provide a venue for brainstorming the desirable features and functions to be built in the expanded learning analytic application. The second focus group session will be conducted at the beginning of the second year of the project to identify potential issues and problems with the expanded learning analytic application in order for the project team to address them before the pilot implementation of the application on eClass.

The expanded learning analytic application will be piloted and evaluated with ten courses offered in the Fall 2018 that make use of eClass at the University of Alberta. The courses will be selected in such a way that they represent a wide range of Faculties and Programs, have large class sizes, and have been considered "challenging" by students in the past. The reason for these selection criteria lies in the fact that the predictive student success algorithm is believed to be most beneficial to large difficult classes where personalized feedback of student learning is most needed but previously not feasible. We expect that courses that implement the predictive algorithm will show an increase in satisfactory grades and a decrease in unsatisfactory grades and withdrawals. Online questionnaires regarding the usability, usefulness and intuitiveness of the expanded learning analytic application will be developed and administered to approximately 50 students and ten instructors at the end of the fall semester. Students will be invited to comment on whether the feedback from learning analytics is indeed helpful in facilitating and improving their learning in the course. This will serve as the validity evidence of the predictive algorithm. The usefulness and impact of the proposed learning analytic predictive model will rely on student experience of using the tool and their performance and success.

Innovation and Alignment with *For the Public Good*

Capturing, tracking and analyzing end-users activity data to predict users' future needs and provide personalized recommendations has long been adopted and utilized by contemporary online businesses such as Amazon, Netflix, and Google. However, student data captured within the higher education institution (e.g., learning management systems, student information system, and student services) have yet to be properly integrated and utilized to realize its full potential for providing valuable insight for students and instructors to facilitate and support learning. Learn analytics is still an emerging field in education (Avella, Kebritchi, Nunn, & Kanai, 2016). The adoption and application of learning analytics in higher education is still mostly small scaled and preliminary. The proposed project will advance the learning analytic application in higher educational institutions and promote holistic analysis and interpretation of the wealth of student data captured by educational and learning systems in generating actionable solutions that help personalize, support and improve student learning experiences.

By developing a useful tool that helps students monitor and self-regulate their learning and assists instructors in providing timely, optimized and constructive feedback, this project aligns nicely with the University of Alberta's *For the Public Good* mission in creating *a supportive learning environment* and meets the objectives of the University of Alberta's new Institutional plan to *"Inspire, model, and support excellence in teaching and learning"*.

Collaboration

The current eClass learning analytic application is developed by the co-PI Dr. Ali Shiri in close and successful collaboration with the Centre for Teaching and Learning (CTL) and Information Services and Technology (IST). Building on this positive and sustainable collaboration model, we will continue to work closely with CTL and IST in enhancing and updating the learning analytic application. To plan for the project, the PI and co-PI have consulted with Dr. Norma Norcente, the Associate Director of CTL (Educational Technology), and Jeff Rawlings, Director of IST (Relationship Management) on the feasibility and utility of the learning analytic expansion. Their comments and suggestions provided useful input in this proposal. The PI and co-PI will continue to work with the CTL and IST during the implementation of the project to ensure the proper development, code testing, compatibility and security analysis, and quality assurance of the expanded learning analytics application. This kind of collaboration and coordination is vital and will also ensure the sustainability of this project and its outcome.

In addition, collaboration with a wide range of instructors and students on campus is an integral component of the design, modification and evaluation of the expanded analytic application. The participation and feedback of the instructors and students will be essential in ensuring the functionality, usability, and intuitiveness of the expanded learning analytic tool. Student learning outcomes will be used as important indicators of the validity of the predictive student success algorithm and the overall effectiveness of the expanded learning analytic application.

Sustainability/Impact on Students

Our close collaboration with CTL and IST and the direct involvement of a large and diverse group of instructors and students across the Faculties and Programs on campus will ensure the sustainability of the expanded learning analytics application. We expect that all instructors and students at University of Alberta will be able to access, utilize and benefit from the expanded analytic application in September 2019. The goal is to support all students on campus to foster and enhance student success by empowering active student learning and provide instructors with a useful tool that helps them to provide valid and timely feedback. Given that Moodle, the basis of eClass, is open source, we will share our analytic tools with other institutions to promote the future use of learning analytics in higher education. In addition, this TLEF project will be leveraged to obtain additional external funding to further explore and evaluate the use of learning analytics in the higher education context in general to enhance student learning experience.

Dissemination

This expanded learning analytic application will be shared with all instructors on campus. In collaboration with the CTL, an introductory workshop will be held to introduce the application to the University of Alberta community at the end of the second year of the project. Additionally, we plan to present our application at one of the University of Alberta International Week events.

Video tutorials will be created for instructors and posted on the university website in order to provide easy access to the instructions on the use of the new expanding analytic application.

Furthermore, the results of this project will be shared outside the University of Alberta through presentations at conferences and published research papers in scholarly journals. The information to be disseminated will address such diverse areas as learning analytics, educational data mining and educational management in higher education. Conference proposals will be submitted to the International Conference on Learning Analytics and Knowledge (LAK) Conference. Journals we seek to publish include the International Journal of Educational Management, the Journal of Learning Analytics, and the Journal of Educational Data Mining.

References

Avella, J. T., Kebritchi, M., Nunn, S.G. & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review, *Online Learning*, *20* (2), 13-29.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., and Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Khalil, M. & Ebner, M. (2016). De-identification in learning analytics. *Journal of Learning Analytics*, *3*(1), 129-138.

Shiri, A. (2016) Learning Analytics: Supporting Teaching and Learning through Learner's Data Analytics and Visualization. *Proceedings of the International Conference on Data Analytics*, Venice, Italy, October 9-14, 2016.