Chapter 11

Analyzing the Engagement of CAPT Program Users With Data Mining Methods:
High Scorers Are Not Always the Best Learners

John-Michael L. Nix
National Taitung University, Taiwan

ABSTRACT

Educators gauge learning with CALL applications via achievement metrics such as point scores or level advancement. Overreliance on such metrics may limit validity of learner engagement measures. This chapter used learning analytic approaches (e.g., cluster and regression analyses) to investigate user engagement with a web-based CAPT program. Cluster analysis identified four types of users with effort-based attributes best distinguishing among types. Regression analysis found that lines recorded has the strongest association with point scores. Follow-up retrospective time-series analysis of cluster members showed distinct trends in learning behavior that indicate possible goal orientations per group. These results imply that one must deconstruct and identify the aspects of engagement that are actually being measured by application metrics. Additionally, significant differences in engagement patterns exist within high and low scoring groups that are opaque to analysis across the whole sample. Finally, activity logs provide data that suggest variability in motivation types.

BACKGROUND

This paper is positioned at the nexus of learning analytics (LA) and CALL, and demonstrates that data mining techniques, specifically cluster analysis, can be effectively utilized by individual course instructors to understand students' trends of learning engagement. The CALL platform under investigation is English Central[1] (EC), originally marketed as a computer-assisted pronunciation training (CAPT) platform, which provides pedagogically enhanced videos sourced from the Internet and partner media providers

as material for pronunciation practice, listening practice and vocabulary learning. EC's online CAPT program provides text-dependent, contextualized speaking practice with native speaker input, some of which is authentic media and some of which is pedagogically designed material.

CAPT programs are a readily available option for language learners and instructors in the 21st century, and they are promoted by developers because of potential advantages for foreign language/ second language (L2) learning (Neri, Cucchiarini, Strik & Boves, 2002; Neri, Cucchiarini Strik, 2006). CAPT programs provide language learners an option of independent study with corrective feedback from an automated system, thereby enhancing the learning experience due to mitigation of negative affect (Neri, A., Cucchiarini, C., Strik, H., & Boves, L., 2002) such as loss of face (Chiu, Liou & Yeh, 2007) which occurs in live classrooms.

To date few researchers have evaluated EC to gauge its effectiveness in yielding English learning gains, as well as to probe levels of satisfaction, perceived effectiveness, and attitudes among users. Doubtless, effectiveness of the program is a primary concern for EFL educators and EC designers. Yet, scrutiny must also be directed towards user characteristics to probe critical factors impacting program efficacy. Of particular concern is learner engagement over the long run. Therefore, the present study departs from the paradigm of variable analytic studies of effectiveness and turns towards exploration of user characteristics to identify variables of interest mined from activity log data, which can be useful for future material evaluation studies of EC and other language learning platforms. This investigative approach is adapted from the fields of educational data mining and LA, which emphasize analysis of learner engagement with large-scale, technology-enhanced learning apparatuses such as massively open online courses (MOOCs), learning management systems (LMSs) or online platforms in order to identify action plans for course modification or learner interventions. These research aims mirror the aims of CALL material evaluation studies which seek to examine the various facets of program effectiveness. Thus, adoption of LA data exploration methodologies may identify user types, and retrospective analyses of their respective learning engagement patterns will yield insight that proves informative when designing program efficacy studies.

LITERATURE REVIEW

Insight From Learning Analytics

As CALL applications have increasingly evolved to online, cloud-based applications, the amount of learner data available to educators has increased as well. This data holds promise as an avenue to identify and design enhancements to learning platforms (Kizilcec, Piech & Schneider, 2013). However, the sheer enormity of available data may become an impediment to use (Ferguson, 2012), hence the rapid development of the field of LA and the statistical tools utilized therein (Ferguson, 2012; Perrota & Williamson, 2016; Seimens & Baker, 2012). The methodologies adopted by LA seek to find interpretable patterns in data sets of tens of thousands of individuals.

A primary tool of LA research is cluster analysis, which groups individuals together by finding a balance between discriminancy and convergence among a set of attribute variables, thereby identifying the particular types of individuals within a sample. Kizilcec, Piech and Scheider (2013) applied this technique to examine the engagement patterns of MOOC users over the respective durations of three courses. They consistently identified four types of users in the MOOCs: completing, auditing, disen-

Related Content

Analysis of Speeches by the Former President of the US, Barack Hussein Obama, Regarding the Middle East and Northern Africa
Alelign Aschale Wudie (2020). *International Journal of Translation, Interpretation, and Applied Linguistics (pp. 20-32).*
www.irma-international.org/article/analysis-of-speeches-by-the-former-president-of-the-us-barack-hussein-obama-regarding-the-middle-east-and-northern-africa/245798

Book Review: Understanding Metaphor Through Corpora: A Case Study of Metaphors in Nineteenth Century Writing
Yanwei Wang (2021). *International Journal of Translation, Interpretation, and Applied Linguistics (pp. 51-56).*
www.irma-international.org/article/book-review/281673

Creating a Micro-Immersion Environment Through Telecollaboration
Tasha N. Lewis (2019). *Computer-Assisted Language Learning: Concepts, Methodologies, Tools, and Applications (pp. 704-729).*
www.irma-international.org/chapter/creating-a-micro-immersion-environment-through-telecollaboration/219696

Computer-Aided Deductive Critical Discourse Analysis of a Case Study from Mauritius with ATLAS-ti 6.2
Komalsingh Rambaree (2014). *Innovative Methods and Technologies for Electronic Discourse Analysis (pp. 346-368).*
www.irma-international.org/chapter/computer-aided-deductive-critical-discourse/76998

Second Language Expressive Writing in Times of Global Crisis: Poetry as a Humanistic Practice
Jawad Golzarand Mir Abdullah Miri (2022). *Research Anthology on Applied Linguistics and Language Practices (pp. 785-804).*
www.irma-international.org/chapter/second-language-expressive-writing-in-times-of-global-crisis/303049