2023

Using Association Rules to Inform Academic Advising: The First Steps to an AI-Academic Advising System

Kari Sandouka

Using Association Rules to Inform Academic Advising: The First Steps to an AI-Academic Advising System

Kari Sandouka
Dordt University
kari.sandouka@dordt.edu

ABSTRACT

The test-optional movement in higher education generates the need for creative solutions in academic advising. As a first step in developing an artificial intelligence-based academic advising system, the results of a longitudinal study are used to measure the learning gain of four essential skills (writing, reading, mathematics, and critical thinking). Association rules are generated to a) find the sequence of general education courses that are taken in sequence to generate learning gains and b) identify other factors that may be used to replace or supplement standardized test scores in course placement decisions. The practical implications of the results indicate a sequence of general education courses that support essential skills and provide the first steps in building an AI-academic advising system.

Keywords

Learning gain, association rules, pattern analysis, longitudinal study.

INTRODUCTION

The test-optional movement in higher education has changed the landscape of institutional admissions. Test-optional is an admission policy that does not require prospective students to submit standardized test scores [3]. The move to optional submission is a response directly related to the covid pandemic. The cascading effect of this change is now being identified in other areas of higher education, such as academic advising. Academic advising is critical to a student's academic success at universities. Academic advising creates a relationship to engage, support and guide students through their educational plans. The process relies on tacit and explicit knowledge of the university, the program, and the student's academic life [1]. The test-optional movement has introduced a more significant challenge to academic advising for placement into courses that fit a university's general education program. Previously, standardized test scores have informed academic advisors of placement options in Mathematics, English, and Literature. Academic advising has become more difficult with the implementation of test-option advising. Advisors do not have sufficient knowledge of how well-prepared a student may be.

The ETS Proficiency Profile (ETS) is a college-level assessment for student learning objectives. The primary purpose is to evaluate learning gain at various points throughout a student's college career [2]. The ETS measures core skills such as college-level reading, writing, critical thinking, and mathematics. The advantage of the Proficiency Profile is the direct evidence of student learning and score for specific essential skills. Leveraging proficiency scores may provide a solution to identify placement suggestions for general education courses (e.g., English and Mathematics). Analyzing the proficiency scores of a student body may result in a better understanding of the student's essential academic skills.

This study aims to investigate how longitudinal data may alleviate the difficulty of academic advising as a response to the test-optional movement in higher education. The study explores using the ETS Proficiency Profile (EPP) in association with general education courses to identify patterns that can inform placement suggestions without using standardized test scores. It is the first step in developing an artificial intelligence advising system.

METHODOLOGY

Association Rules is a machine learning algorithm to extract interesting relationships using correlations, frequent patterns, and association measures [4, 7]. This study uses the Apriori algorithm to generate association rules between EPP scores for the essential skills and general education classes taken. A longitudinal study design is used to investigate making informed decisions for academic advising without using standardized test scores [5, 6].

The sample for this study includes 1,451 EPP records from a small rural private university in the Midwest region of the United States. The dataset consists of records collected over eight years with matched pair data. Each academic year, the institution designates two days for EPP assessment. The first date targets incoming students during their first semester (e.g., first-year

students with less than 30 earned credits). The second date targets students in the spring semester of their third year (e.g., junior-level students with at least 75 earned credits). Consistent with the longitudinal study design [5,6], the cohorts were initially tested in their first year's fall semester and tested again in the spring semester of their third year. The sample for this study was not planned as a longitudinal cohort. The institution administers the EPP as an ongoing assessment of learning gains.

ANALYSIS AND DISCUSSION

The Proficiency Profile identifies four essential skills: critical thinking, reading, writing, and mathematics. The goal of the ETS is to identify learning gains. Following the longitudinal study of Roohr, Liu, and Liu, learning gain is calculated as the difference between the first and second test scores (overall and for each essential skill) [6]. The first step of the investigation was to identify learning gains for the essential skills. Students receive a score between 100-130 for each skill. The match-pair data shows changes in the average score for each essential skill (Figure 1). The percentage of students that increased their scores between the two assessments provides evidence of learning gains across all essential skills.

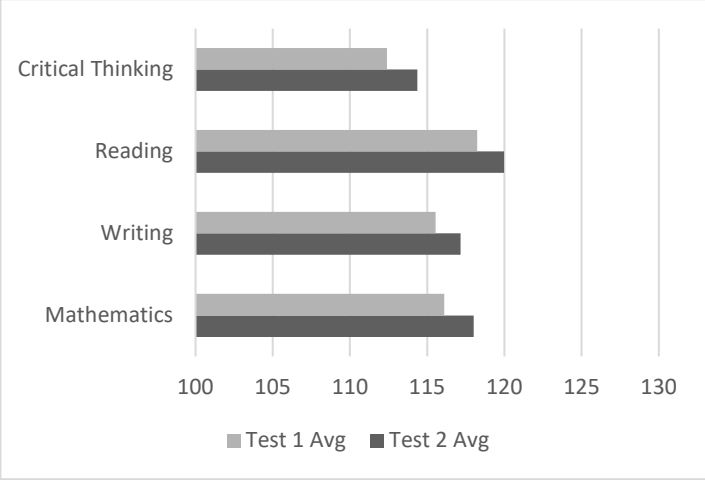

Figure 1. Group Average Scores for Each Assessment. ETS PP scale for the listed categories is 100-130.

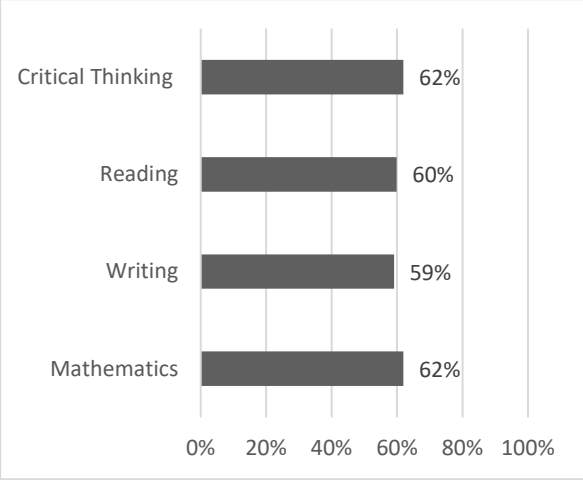

Figure 2. Percent of Students that Increased Scores on their second test.

Results from statistical tests show apparent differences between the first and second tests. We tested the paired scores with a hypothezied mean difference of zero (no difference between the first test as an incoming student and the second test as a third-year student). The paired sample t-test results are provided below. For each of the core skills, the p-values were below the alpha value of (0.05). Secondly, the t-critical value for the two-tailed test was less than the absolute value of the t-statistic. We conclude that there is a statistical difference between scores on the first test and scores on the second test, indicating learning gain in each of the four core skills.

The t-test results provide evidence of statistical significance between the two tests, leading us to question which of the classes in the general education program are contributing most to learning gains. The institution's general education program requires all students to take English Composition (CORE-120) or equivalent, Philosophy (CORE-200), Literature (CORE-180), Biblical Foundations (CORE-150), two history courses (CORE-140 and CORE-145), and a Mathematics course. Standardized test scores have helped to place students in the appropriate English and Mathematics courses. To identify patterns between student characteristics and general education courses, we focus on the learning gains in the essential skills.

For students that increased scores in the one of the core skill areas, we built a data set that included their general education courses and demographic attributes. The factor included: gender, miles from home, predicted first semester GPA (Pred GPA), first semester GPA (FS GPA), and high school GPA. Miles from home were numerically coded with values starting at zero and exceeding 1,000. The predicted GPA is an internal institutional measure. All GPA attributes were coded on the A-B-C-D-F letter grade scale prior to being used in the association rule model. The mlxtend Python library was used to implement the apriori algorithm. The model looked for rules that had a minimum lift of six and a minimum confidence of eighty percent. The most interesting rules generated from this data set are shown in Figure 3.

		Test 1	Test 2	Pearson's Correlation	T-Stat
Critical Thinking	*Mean*	112.40	114.36	.652	13.83*
	Variance	39.81	44.50		
Reading	*Mean*	118.24	119.98	.619	11.10*
	Variance	46.00	47.77		
Writing	*Mean*	115.54	117.16	.618	13.97*
	Variance	25.81	25.32		
Mathematics	*Mean*	116.12	118.02	.704	17.19*
	Variance	29.21	31.36		

Table 1. T-test Results for Essential Skills

The Proficiency Profile critical thinking questions measure students' abilities to distinguish between rhetoric, recognize assumptions, infer and interpret relationships, and draw valid conclusions [2]. Results from the t-Tests and Association Rules show that the general education program provides substantial opportunities for learning gains specific to critical thinking skills. More specifically, History, Literature, and Philosophy courses are protagonists in the learning gain. Historically, standardized test scores are not a factor in placing students in history, philosophy, or literature courses. The association rules identify the internal predicted GPA as a factor in determining how the core critical thinking skill increases through the general education sequence. For instance, students earning a B in high school, a predicted GPA of C, and taking two history, Bible, or Philosophy courses tend to do better in Literature and make at least a C in their first semester of college.

The writing questions measure students' abilities to recognize proper grammar, units of language, and figurative language [2]. The results from this investigation show evidence of learning gains through the general education courses, most specifically from the English, Literature, and History courses. Historically, the institution has used standardized testing to place students in an English class (remedial, English composition, or advanced). The association rules identify English Composition as a protaganist to earning a least a C in the first semester of college and increasing the writing core skill.

The reading questions measure students' abilities to interpret meanings, recognize explicitly presented information, and make inferences [2]. Along with the evidence of learning gains, the results show that the Biblical Foundations and Literature courses are foundational to creating learning gains for student academic success. The overall picture identifies that while the general education course set is required for all students, the association rules only identify one history course at a time, never the two together. For academic advising purposes, this insight is useful for advisors but does not help in answering the question of how to place students without standardized test scores.

The mathematics questions measure students' abilities to interpret mathematical terms, interpret tables and graphs, evaluate formulas, read scientific measurement instructions, and recognize formulas or expressions [2]. One of the most popular mathematics courses at the institution is Introduction to Statistics (STAT-131). There are several other courses available for students to meet their general education requirement, but this course is the only one that appeared in the association rules.

The alternative factors (miles from home, gender) did not have a significant impact on proficiency as they relate to the general education courses. The gender factor only appeared in the Mathematics association rules. The t-test scores for the mathematics essential skill show a difference between the two assessments. The test1 and test2 math scores are significantly different, but the percentage of students that increased their mathematics proficiency is very similar to those that did not.

Critical Thinking 'Interesting' Rules

{History Courses, HS-GPA(B), Pred GPA (C)} → {Literature, FS GPA (C)}

{Bible, HS GPA (B), Pred GPA (C)} → {Literature, FS GPA (C)}

{Philisophy, HS GPA (B), Pred GPA (C)} → {Literature, FS GPA (C)}

Writing 'Interesting' Rules

{English, Pred GPA (C)} → {FS GPA (C)}

{History Courses, Pred GPA (C)} → {FS GPA (C)}

{Literature, Pred GPA (C)} → {FS GPA (C)}

{Bible OR Philosophy, Pred GPA (C)} → {FS GPA (C)}

Reading 'Interesting' Rules

{Bible, History 2, Pred GPA (C)} → {Bible, Philosophy, FS GPA (C)}

{Bible, History 2, Pred GPA (C)} → {Literature, Philosophy, FS GPA (C)}

Mathematics 'Interesting' Rules

{HS GPA (B), FS GPA (C)} → {Pred GPA (C)}

{HS GPA (A), Statistics, Female, FS GPA (A)} → {Pred GPA (A)}

{Female, FS GPA (A)} → {Pred GPA (A)}

{HS GPA (A), 150-Miles, FS GPA (A)} → {Pred GPA (A)}

Figure 3. Assication Rules

CONCLUSION

Academic advising is a complicated process that greatly relies on explicit and tacit knowledge. As institutions adjust to changes introduced by the covid pandemic, student data becomes a treasure trove waiting to be discovered. The investigation discussed in this paper provides the first steps in exploring explicit information that can be used to supplement the lack of standardized test scores. The results are interesting and provide insight into institutional change in admissions and advising steps. However, they do not answer the initial question of how to place students in courses without standardized test scores. The longitudinal study provides evidence that learning gain is present and students are gaining academic success through the essential skills. The results of the pattern analyses is only a glimpse of what occurs throughout a college career. More analysis or data mining features are needed to answer the original question.

REFERENCES

1. 1. Bliquise, G. and Shaalan, K. (2022) AI-based Academic Advising Framework: A Knowledge Management Perspective *International Journal of Advanced Computer Science and Applications*, 13, 8, 193-203.
2. Education Testing Services (ETS) (2022), *ETS Proficiency Profile*. https://www.ets.org/proficiency-profile/about/test-content.html
3. Elias, Jacquelyn (2022) Test-Optional Policies Now Dominate Higher Ed *The Chronicle of Higher Education* November 16, 2022. https://www.chronicle.com/article/test-optional-policies-now-dominate-higher-ed
4. Kotsiantis, S., & Kanellopoulos, D. (2006). Association rules mining: A recent overview. GESTS International Transactions on Computer Science and Engineering, 32(1), 71-82.
5. Mathers, C.E, Finnery, S.J, and Hathcoat, J.D. (2018) Student Learning in higher education; A longitudinal analysis and faculty discussion. *Assessment & Evaluation in Higher Education*, 42, 8, 1211-1227.

6. Roohr, K.C., Liu, H., and Liu, O.L. (2017) Investigating student learning gains in collect: A longitudinal study. *Studies in Higher Education*, 42, 12, 2284-2300.
7. Shatnawi, R., Althebyan, Q., Ghalib, B., & Al-Maolegi, M. (2014). Building a smart academic advising system using association rule mining. arXiv preprint arXiv:1407.1807.