Academic Coaching: Creating Optimal Matches Between Students and Coaches

Tara Jawahar

OVERVIEW

An optimal matching between Academic Coaches and students can benefit both the student, whose academics, stress, and perception of self can improve significantly, and the Academic Coach, in terms of personal development and subject matter information gain. As of the 2019-2020 academic year, Michael Poljak, Assistant Director, and TC Eley, Coordinator, match students to coaches based on a variety of factors including the student's and the coach's program of study, gender, and school year. However, identifying the factors in this matching that most contribute to student success can help create better pairings. There may be underlying factors, such as personality types, race, family background, etc., that allow the coach to better connect with and help the student achieve their goals.

DATA SOURCES

Pre-and post-tests AY18-19 and AY19-20

- Subjective Measures
    - Goals
    - Implicit Theories of Intelligence (Mindset)
    - Sense of Social and Academic Fit (Belongingness)
    - Academic Self-efficacy (Self-efficacy)
    - Life Orientation Test (Optimisim)
- Objective Measures
    - Year at University
    - School (CIT, MCS, DC, SCS, CFA, TSB, Heinz)
    - Gender
    - Citizenship Status
    - Race
    - College Athlete
    - First Generation College Student
- Coach Feedback (only in post-tests)

Initial Consultations

- Year at University
- School
- Major

- New or Returning Student
- Classes
- Goals

Roster of Academic Coaches

- Year at University
- School
- Major
- Languages
- Countries

Note that only students with both pre- and post-test data available were included in this study. This dataset was 128 students (58% female, 42% male). The racial demographic is as follows: 52% Asian, 19% White, 13% Hispanic, and 9% Black, with the remaining 7% either identifying as more than one race or not reporting race. Approximately 27% of the students are international students. Of the 46 coaches that had at least one student, 52% are female and 48% are male.

FEATURES

In this study, we will be comparing the background information available about students and coaches. The predictor variables were extracted from the given pre- and post-test surveys, initial consultations, and coach information. These included the subjective measures of the student (by the student), subjective measures of the coach (by the student), student goals, and demographics for both the student and the coach. Features in the final dataset were binary variables that encoded whether the student and the coach were of the same gender, major, class, school, or race, as well as whether the student's citizenship status (0 = International, 1 = Domestic) and if they were a student athlete or first generation student.

Student Demographics

	Mean
Citizenship Status	0.729
Same Gender	0.698
Same Major	0.039
Same Class	0.031
Same School	0.286
Same Race	0.411

From these means, we see that the citizenship status is skewed in favor of domestic students, and that most student-coach pairs in this dataset have the same gender and are of the same race, while most do not have the same major or are the same class rank.

OUTCOME VARIABLES OF INTEREST

The subjective measures of the student were split into four different categories of mindset, belongingness, self-efficacy, and optimism. Pre- and post-test results were compared for each student by subtracting the post-test value and the pre-test value to get the net change in the measure and then averaging across all questions in that category (i.e. mindset, belongingness, self-efficacy, and optimism). The following confidence intervals indicate that we are 90% confident that the true net change (for all students) in each subjective measure is within that interval.

Net Change in Subjective Measures of Self

	90% Confidence Interval
Mindset	(-0.471, 0.045)
Belongingness	(0.010, 0.372)
Self-efficacy	(0.113, 0.336)
Optimism	(0.069, 0.323)

When looking at these measures, it appears that, on average, students in this dataset did not improve with respect to mindset after coaching. This is because the confidence interval includes 0 and we cannot assume that there is a positive or negative change in mindset after the sessions. Since the other three measures do not contain 0, we can assume that the population value (the net change for all students, including those outside the sample dataset) is positive.

METHODOLOGY

I applied regression techniques to this dataset to predict the student success metric (in mindset, belongingness, self-efficacy, and optimism) based on the features of the student-coach pair. For the final model, I used an ensemble learning method, in which multiple models are generated and then combined to create a stronger predictive model. I used gradient-boosted regression for its better performance with small datasets as compared to other types of regression. This model involves optimizing a loss function, using a weak learning method to make predictions, and incorporating an additive model to combine the weak learning methods and create a strong predictive model. Given that this dataset consists of 128 student-coach pairs, gradient-boosted regression can help account for this while also determining which features are important in the model. The training dataset – the dataset with which the model selects the optimized parameters – is on 90% of the dataset, or 115 students. The test dataset – the portion of the dataset that the model has not seen to test its accuracy – is 13 students.

RESULTS

Mean-Squared Error for Each Category

	MSE
Mindset	4.45
Belongingness	0.55
Self-Efficacy	0.44
Optimism	1.74

Since we would like MSE to be as close to 0 as possible, we see that the model can accurately predict belongingness and self-efficacy and is not as good at predicting mindset and optimism. For each success metric, the most important features are as follows:

	Most Important Features
Mindset	Same Gender, Same School
Belongingness	Same School, Same Race
Self-Efficacy	Citizenship Status, Gender
Optimism	Citizenship Status, Same Race

From these results, we see that different features are important based on the outcome variable. For self-efficacy and optimism, the student and coach having the same background is more important, whereas for mindset and belongingness background and academic discipline are necessary.

CONCLUSION

We can use these results to create better pairings between students and coaches that result in a higher level of student success. This analysis gives us the data regarding which subjective measures have more predictive power and the factors to look for in determining matches.

On average, the model is more accurate in predicting belongingness and self-efficacy, presumably because these two factors are directly related to the Academic Coaching program. However, the model does not accurately predict mindset in students, and it appears that in this dataset, mindset scores in the post-test are lower than in the pre-test. This may be due to a lack of long-term data that is required to show significant changes in mindset. There appears to be the greatest increase in self-efficacy when comparing the pre- and post-tests because of the emphasis on self-efficacy in coaching sessions.

My recommendations are as follows:

- Incorporate material about mindset when training coaches and address changes in mindset in sessions to emphasize its importance in student success and to see improvement in this area.
- Use pre-test results to determine the student's areas of weakness [mindset, belongingness, self-efficacy, optimism]. Prioritize matching students to (1) coaches with the same background (race, gender, citizenship status) if the student struggles with self-efficacy and optimism or (2) coaches in the same academic discipline (school/department) if the student is weaker in mindset and belongingness.

FUTURE WORK

This study illuminated underlying factors that should be prioritized when selecting coaches to work with students. However, a larger dataset would yield a more predictive model and would be useful in identifying other important factors. Since this dataset had a relatively small number of students, some factors may be over- or under-represented, and including more students would allow the model to better understand the relationship between the features and the outcome variables. Additionally, having additional information such as citizenship status for coaches and objective outcome variables (mid-semester, final, and cumulative QPAs) would allow us to explore other relevant features and create a more robust student success metric.