THE RESEARCH AGENDA

Evaluating the Impact of Entrance Awards on Yield Rates Using Machine Learning

By Alexander A. Ondrus and Ashmita De

Entrance awards are financial incentives offered to admitted students at or near the time of admission. These awards are a frequently used tool to enhance yield rates among post-secondary institutions, sometimes among specifically-targeted populations. The impacts of these awards have been studied in many different contexts, using a variety of methods (Bartik, Hershbein and Lachowska 2021, Cornwell, Mustard and Sridhar 2006, Crowne 2022, Gurantz and Odle 2022). The most effective way to quantify the effect of entrance awards is to assign them randomly among qualified applicants, and then compare the outcomes with the outcome of the group that was not assigned any awards. This method was used in (Firoozi 2022) but is not advisable in the authors' regulatory environment. Therefore, this research compared the observed yield rate among students to the yield rate that was predicted using a machine learning model that was trained on data collected before the introduction of entrance awards. The predicted yield rates can be thought of as the behaviour of a "synthetic control group" that behaves as would be expected if no entrance awards were offered, allowing the researchers to measure the impact of the awards.

NAIT Context

The Northern Alberta Institute of Technology (NAIT) is a publicly funded polytechnic institute located in Edmonton, Alberta, Canada which offers a combination of apprenticeship, degree, diploma, certificate, and non-credit programs to approximately 35,000 students per year. NAIT offers a wide variety of programming focused on the trades, applied sciences, health, and business.

The province of Alberta has 26 publicly funded postsecondary institutions (Government of Alberta 2024), six of which are in the Edmonton area. Several of these institutes offer business programming, resulting in competition for the same pool of applicants. NAIT introduced the JR Shaw School of Business (JRSSB) Entrance Award for the Winter 2024 term to encourage more new domestic student enrollments. This award is available to all domestic applicants to the Bachelor of Business Administration (BBA, a 4-year degree) or Business Administration (BUS, a 2-year diploma) programs who have a high school entrance average of 70 percent or greater. The award is $1000 (CAD), and eligible students are offered the award shortly after their admission offer letter.

Separate from concerns about the ability to compete for new students, NAIT recognizes its role in overcoming the legacy of colonialism and its impact on Indigenous populations in Canada. One of those impacts is lower postsecondary participation rates among Indigenous people. According to Colleges & Institutes Canada, Indigenous youth are just over half as likely (37 percent) to have completed or recently attended a postsecondary program as non-Indigenous youth (72 percent) (Colleges & Institutes Canada 2024). The Truth and Reconciliation Commission of Canada (2012) recognized this issue in its Calls to Action, specifically call #11:

> *We call upon the federal government to provide adequate funding to end the backlog of First Nations students seeking a post-secondary education.*

In alignment with this call, NAIT introduced the Four Directions Entrance Award in Winter 2024. This $1000 award is for Aboriginal (First Nation [Status and Non-Status], Metis, and Inuit) students. Like the JRSSB Entrance Award, the Four Directions Entrance Award is offered shortly after the admissions offer letter.

Related Work

Postsecondary institutes expect financial aid to have a positive impact on enrollment. The effect of financial aid on enrollment has been of interest to institutions for some time. A study of the HOPE program in the state of Georgia, which included a merit-based scholarship for students attending in-state colleges and a grant for those entering technical schools (Cornwell, Mustard and Sridhar 2006), showed an increase in total freshmen enrollment in four-year public and private institutions across Georgia colleges. Similarly, a study (Bartik, Hershbein and Lachowska 2021) on the effect of Kalamazoo Promise, the first major "place-based" scholarship program in the United States, also showed increase in college enrollment and credential attainment with stronger effects in women. Another case study of a merit-based scholarship program of a U.S. university's Japan-based branch showed a higher yield rate for scholarship recipients compared to non-recipients. In this study, it was also concluded that scholarships were a likely factor in an increased yield for the most academically talented students (Crowne 2022). A study has also investigated the impact of post-admission merit-based scholarships offered by selective colleges to attract academically talented students before enrollment deadlines. It examines the effect on enrollment and graduation outcomes across students with varying socioeconomic backgrounds. The finding showed that the scholarships increase enrollment among disadvantaged students but have minor negative impacts on the enrollment decisions and degree attainment of advantaged students. The study thus suggested limiting eligibility for such scholarships by socioeconomic status (Firoozi 2022). The effects of Florida's Bright Futures (BF) merit aid program on college enrollment and degree completion have also been studied previously (Gurantz and Odle 2022). Even though this study did not find any significant effect of Florida's BF program on college enrollment, out-migration, or degree attainment, it suggested investments in need-based aid highlighting prior research that found that the need-based Florida Student Access Grant did indeed improve student outcomes (Castleman and Long 2016).

Predicting enrollment using machine learning (ML) models is increasingly becoming a streamlined approach. There have been publications on using ML based methods for predicting enrollment (Shao, *et al*. 2022; Soltys, *et al*. 2021) and matriculation (Hansen 2020).

Similarly, NAIT uses a ML based method to predict the probability of enrollment for each individual student.

In literature, there have been several studies of the impact of financial aid on enrollment, investigating causal relationships and the differential effects on specific demographic groups. As enrollment prediction with ML models is becoming more frequent among post-secondary institutions, this paper compares the observed yield rate among admitted applicants to the predicted yield rate from a ML model that had no entrance award information. Thus, the difference between the prediction and the actual observation highlights the impact of the entrance awards on NAIT's enrollment yield rate.

Data

NAIT's JR Shaw School of Business (JRSSB) offers two business programs, the Bachelor of Business Administration (BBA, a 4-year degree) and the Business Administration (BUS, a 2-year diploma). Each of these programs take new students in the fall, winter, and spring terms (in decreasing order of size). Both programs began offering both the JRSSB Entrance Award and the Four Directions Entrance Award in the Winter 2024 term. To understand the benefits of these new Entrance Awards, the researchers compared the observed yield rate among students to the yield rate that was predicted using a machine learning model. To make comparisons between the predicted probabilities of enrollment and the observed yield rate, students who were admitted after the machine learning probabilities were generated were excluded. To account for potential bias in the machine learning algorithm (details in the Analysis section below), the study also included students admitted to the Winter 2023 term (again, only those for whom probabilities had been generated). This gave the sample sizes shown in Table 1.

The researchers gathered 27 variables for each of these students to ensure that the analysis was not influenced by factors outside of the entrance awards. These variables consisted of demographic information (gender, age at term start, immigration status, minority status, etc.), admissions information (whether applicants had previously enrolled in NAIT programming, their high school entrance average, etc.), enrollment information (enrolled or not, full-time vs. part-time status, etc.), and the estimated probability of enrollment (the predicted yield rate).

TABLE 1 ➤ Sample Sizes

Term	Entrance Award Received	Sample Size
Winter 2024	Four Directions	15
	JRSSB	119
	No Award	135
Winter 2023	N/A	243

Analysis

This section discusses how the researchers estimated the probability of applicants yielding using a machine learning model, identified and compensated for bias in that model, and used the adjusted probabilities to estimate the impact of entrance awards on yield rates.

Machine Learning Model

Since the Fall 2023 admissions term, NAIT has implemented machine-learning-based predictions for the number of new students each fall and winter term. This is done by using a machine-learning model to predict the probability that each applicant will enroll in any of the programs for which they have applied that term. Much of the motivation and ideas are similar (Hansen 2020), while the algorithms employed are different. The researchers trained models on 27 predictors, which consisted of demographic information (gender, citizenship status, residency location, minority status, etc.) as well as information about the student's application(s) (number of applications submitted, timing of application, number of interactions with admissions, number of programs applied for, etc.).

When selecting an algorithm for modelling the data, the researchers had three primary considerations. First, as a publicly funded institution the researchers work in an environment with a strong focus on accountability. This increases the importance of model interpretability if they are going to be using these predictions to direct

activities. Second, the data structure is complex, and there are a wide variety of variable types. There are some (but few) numeric variables, but most of the data is qualitative with several variables having many levels. These considerations make tree or rule-based models a good fit (Kuhn and Johnson 2013, 369). To combat the concern of instability in these models the researchers selected three different tree-based models, trained and tested each on historical data, and selected the one that performed best on the most recent past term. The three models considered were random forest as implemented in the randomForest package (Liaw and Wiener 2002), C5.0 as implemented in the C50 package (Kuhn and Quinlan 2023), and extreme gradient boosting as implemented in the xgboost package (Chen and Guestrin 2016).

When generating the new student enrollment predictions for the Winter 2024 term, the researchers first created a test/train split using data from all winter terms 2014–2022, performed cross-validation parameter selection, and generated predictions for all Winter 2023 applicants, which were validated against the actual results from the Winter 2023 term. The most accurate model was C5.0, with a prediction accuracy of 87 percent. Another C5.0 model was then trained using a test/train split on all 2014–2023 winter terms and used to generate probabilities of enrollment for all Winter 2024 applicants.

Bias Adjustment

Individual classification trees are known to suffer from high variance (Hastie, Tibshirani and Friedman 2017, 312), making them prone to over-fitting. All the algorithms listed above used some form of averaging (as in the case of the random forest) or boosting (as in the case of C5.0 and extreme gradient boosting). For a detailed treatment of these methods see (Kuhn and Johnson, Applied Predictive Modeling 2013, Ch. 13). The trade-off of these methods to reduce the variance is that they increase the bias. For an excellent introduction to bias, variance, and model complexity see (Hastie, Tibshirani and Friedman 2017, 219).

The bias of this model means that probability predictions for a sample that is not representative of the population will have a consistent error. The sample that researchers were interested in was:

- Students admitted to the BUS or BBA programs in Winter 2024
  - → These are only two of the nearly 80 programs that NAIT offers
- Domestic students
  - → The model was trained for both domestic and international students
- Students who were offered entrance awards
  - → For the JRSSB Entrance Awardees from high school, these students will have high school grades higher than average (*see* Figure 1, on page 13)

The researchers also tested to see whether the students in the data differed from the general population in any of the following ways:

- Age at term start
- Concentrations (specializations) applied for
- Portion of local applicants (*i.e.* applied from within the City of Edmonton)
- Portion of in-province applicants (*i.e.* applied from within the Province of Alberta)
- Gender

The researchers did not find any statistically significant differences between the sample they chose and the general population for any of these variables. Therefore, when estimating bias for each of the three groups (Winter 2024 applicants with no entrance award, with the JRSSB Entrance Award, and with the Four Directions Entrance Award), the researchers needed only control for the program of application, the national status (*i.e.* domestic only), and the entering high school grade average.

The challenge was how to estimate the bias of the model for each group without contaminating the probability estimates for the Winter 2024 predictions, *i.e.* how to separate the error due to bias from the impact of the entrance awards. To do this the researchers used a bootstrap procedure (Pruim 2018, 301) applied to the Winter 2023 applicant probabilities.

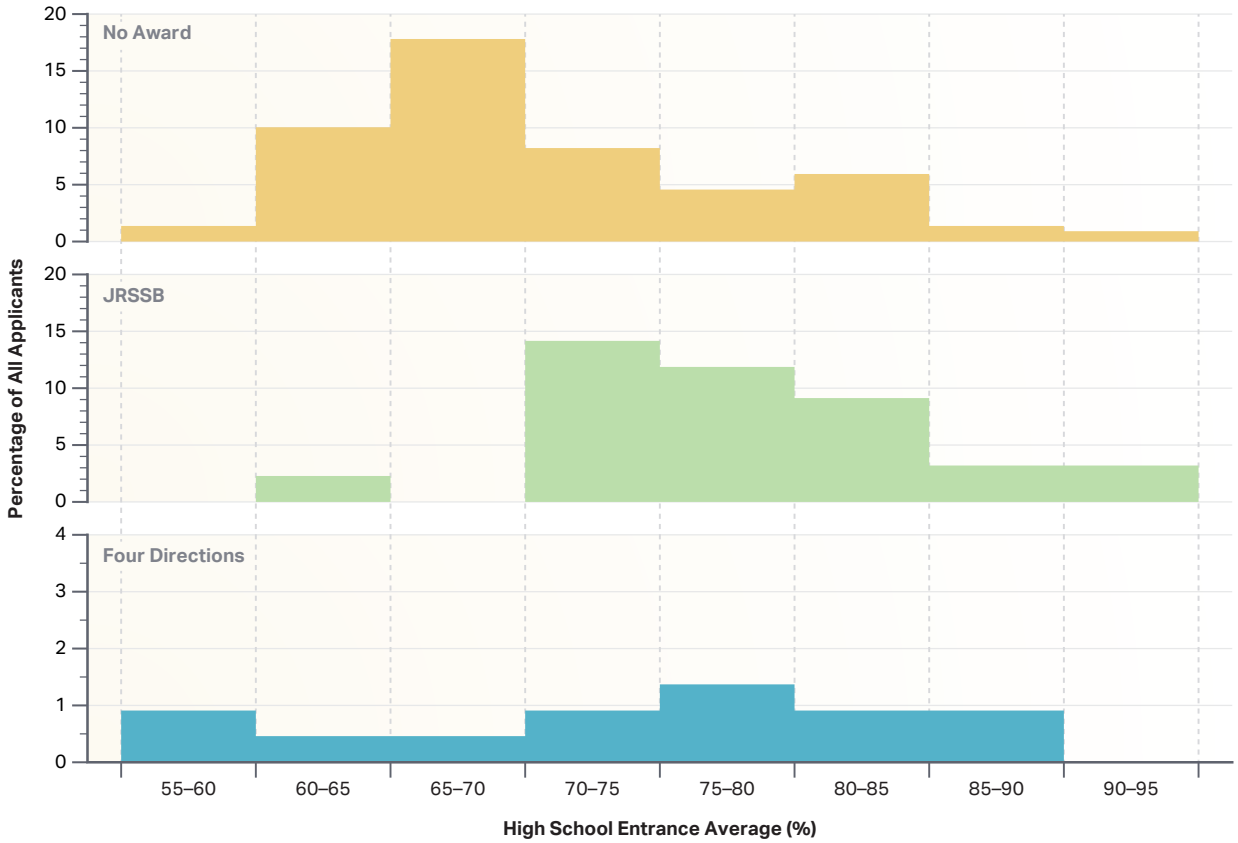

FIGURE 1 ➤ Distribution of High School Entrance Averages by Entrance Award Status

- For each group, draw a sample (with replacement) of domestic BUS and BBA applicants from the Winter 2023 data with the same sample size and distribution of high school grades.
- Calculate the difference between the average predicted probability and the yield rate for each group.

TABLE 2 ➤ Summary Statistics for Bias Estimates by Entrance Award Status

Entrance Award	Bias Estimate (%)	95% CI (%)
No Award	13	7, 19
JRSSB	6	-1, 12
Four Directions	7	-10, 24

The researchers repeated this procedure 1,000 times and examined the mean and distribution of the errors from step two. They saw that the distribution of estimates was normal for each group and calculated 95 percent confidence intervals for the bias estimates.

When comparing the predicted yield rates to the observed yield rates for each group, the researchers added these bias estimates (with confidence intervals) and used the *adjusted* probability estimates to determine the impact of entrance awards.

Synthetic Control Group

Once researchers know the appropriate bias estimate to use to adjust yield probabilities, the predicted probabilities function as a "synthetic control group." Recall that the machine learning model was trained on historical data that did not include entrance awards, so the adjusted probability predictions form the best estimate

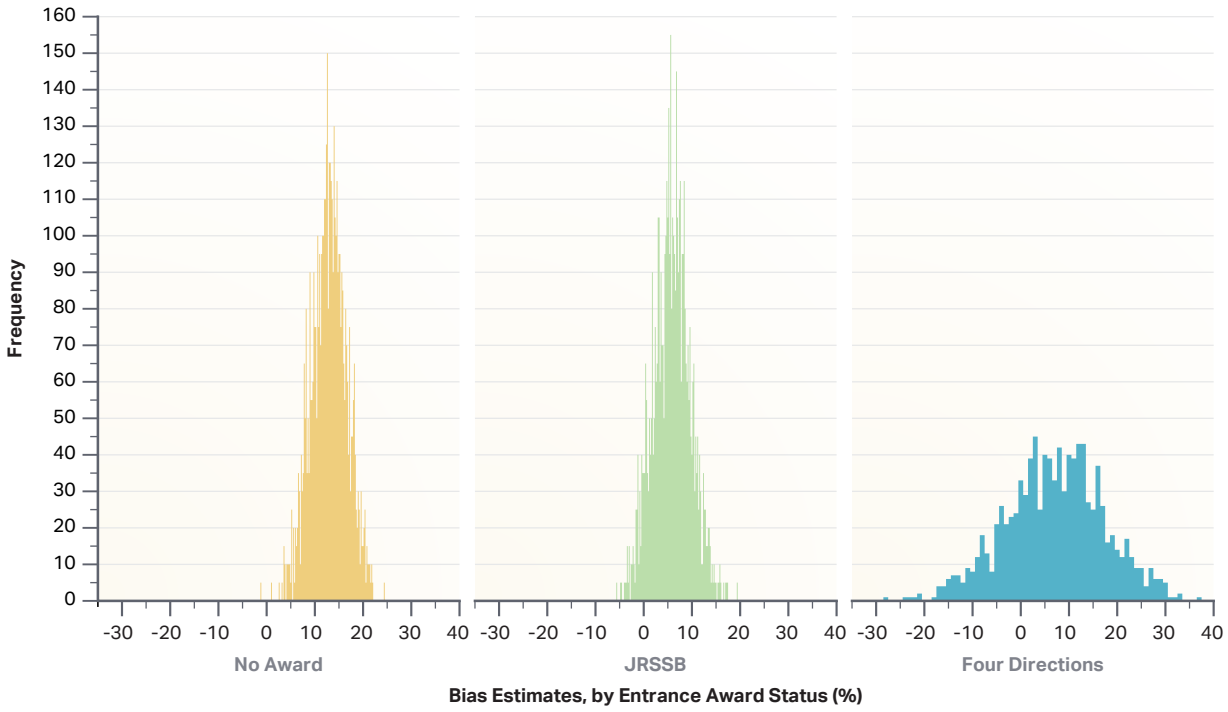

FIGURE 2 ➤ Distribution of Bias Estimates by Entrance Award Status

for how those students would have yielded without the awards. Thus, to estimate the impact of the awards researchers compared the adjusted predicted yield rate to the observed rate for the Winter 2024 term for each group. They did this by first calculating the mean prediction, then adding the bias estimate for that group (as well as the 95% CI), and then comparing. The results of this calculation are shown in Table 3.

Results

The students who were awarded a JRSSB Entrance Award yielded at a significantly higher rate than predicted, while those students who either did not receive an entrance award or received the Four Directions Entrance Award yielded at a rate within the predicted range. The impact of the JRSSB Entrance Award was dramatic, with the observed yield rate 14 percent higher than the upper bound of the prediction range.

To make the results more tangible, the researchers translated them into headcounts by applying the predicted yield rates to the number of admitted students in each category (those not offered awards, those offered the JRSSB Entrance Award, and those offered the Four Directions Award) (*see* Figure 3).

TABLE 3 ➤ Predicted and Observed Yield Rates by Entrance Award Status

Entrance Award	Predicted Yield (%)			Observed Yield (%)
	Unadjusted	Adjusted	95% CI	
No Award	46	59	53, 65	56
JRSSB	65	71	65, 78	92
Four Directions	63	70	54, 87	73

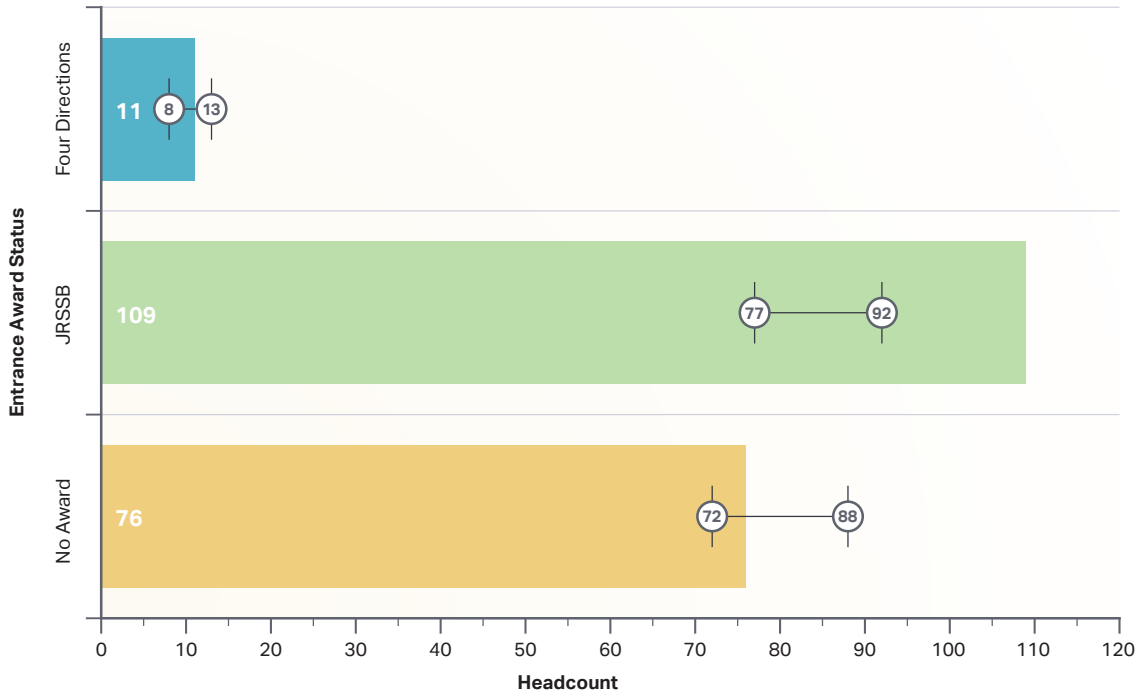

FIGURE 3 ➤ Predicted and Observed Head Counts by Entrance Award Status

Note: Error bars indicate predicted headcount for each status

This representation makes it clear that the Four Directions Award did not have sufficient sample size to make a conclusive statement about its impact. The number of students who did not receive an entrance award and yielded into the BUS or BBA programs was within the predicted range, which is as expected. Among students who received a JRSSB Entrance Award, this analysis shows that the impact of the award was that an additional seventeen to 32 students enrolled who otherwise would not have. For context, this means that the JRSSB Entrance Award increased the total domestic new intake for the Winter 2024 term by between 5 percent and 9 percent.

Conclusions and Discussion

This study aimed to examine the impact of entrance awards on yield rates in NAIT's business programs in the Winter 2024 term. The researchers found that, after adjusting for model bias, the observed yield among students who were offered a merit-based JRSSB Entrance Award significantly exceeded predictions. They were further able to quantify the impact the entrance awards had on yield rate by comparing observed and predicted headcount, estimating that the award induced an additional seventeen to 32 students to enroll in the business diploma and degree programs. The researchers were not able to observe a significant increase in the yield rates of applicants who were offered a Four Directions Entrance Award (awarded to all Indigenous applicants), due to a small sample size. They also saw that the yield rate for students who did not receive an entrance award fell within the predicted range. These results are of great interest to the institution as it considers the future use of entrance awards.

One historical limitation of such studies is the lack of a control group to make comparisons against. For example, Crowne (2022, 455) states that:

> *The analysis of yield of merit-based scholarship recipients and non-recipients alone does not account for a student's likelihood, or propensity, to enrol at a particular university. This can bias the results of the impact of*

merit-based scholarships on enrollment and potential overstate their effectiveness.

This work overcomes that obstacle by making comparisons of entrance awardees against hypothetical versions of themselves rather than against non-awardees, avoiding the bias identified by Crowne (2022). By measuring the impact of entrance awards in this way the researchers can provide valuable insight to administration on the effectiveness of these awards, which is particularly important in the current constrained funding environment.

Future avenues of potential research that was not pursued in this study relate to the differential impact of entrance awards on marginalized groups. The researchers were unable to complete the analysis of this type due to issues of sample size for the Four Directions Entrance Award case, but the techniques described above could be used at a larger institution to make progress in this direction.

Similar methodologies could be applied in the continuing student context in the future. NAIT is developing a machine learning-based early alert system aimed at identifying students at risk of failing to continue in their program. The primary purpose of this system is to trigger targeted intervention to increase the likelihood of retention, but the difference in predicted versus observed student behaviour will provide another means of evaluating the impact of the interventions themselves.

ACKNOWLEDGEMENTS

The authors would like to thank the following members of the Education Insights, Data and Research department at NAIT: Douglas Akhimienmhonan, Director, for his editing suggestions, and Toni Tessari and Imran Kalair, Data and Insight Specialists, for their assistance with gathering and interpreting the data. We would also like to thank Fred Shultz, Director of Enrollment Management, for his advice throughout the project and editing suggestions.

References

Bartik, T. J., B. Hershbein, and M. Lachowska. 2021. The effects of the Kalamazoo Promise Scholarship on college enrollment and completion. *The Journal of Human Resources*. 56(1): 269–310. Available at: <jhr.uwpress.org/content/56/1/269>.

Castleman, B. L., and B. T. Long. 2016. Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*. October 2016: 1023–1073. Available at: <jstor.org/stable/26553233>.

Chen, T., and C. Guestrin. 2016. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco: ACM. Available at: <doi.org/10.1145/2939672.2939785>.

Colleges & Institutes Canada. 2024. *By the Numbers: Indigenous Post-Secondary Education in Canada*. June 17. Ottawa, Ontario, Canada: Author. Available at: <collegesinstitutes.ca/by-the-numbers-indigenous-post-secondary-education-in-canada/>.

Cornwell, C., D. B. Mustard, and D. J. Sridhar. 2006. The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Program. *Journal of Labor Economics*. 24(4): 761–786. Available at: <jstor.org/stable/10.1086/506485>.

Crowne, N. 2022. The impact of merit-based scholarships on enrollment yield: A branch campus case study. *Journal of Higher Education Policy and Management*. 44(5): 443–458. Available at: <doi.org/10.1080/1360080X.2022.2109547>.

Firoozi, D. 2022. The impact of post-admission merit scholarships on enrollment decisions. *Economics of Education Review*. 86(Article 102221). Available at: <sciencedirect.com/science/article/pii/S0272775721001321>.

Government of Alberta. 2024. *Publicly Funded Post Secondary Institutions*. Edmonton, Alberta, Canada: Author. Available at: <alberta.ca/publicly-funded-post-secondary-institutions>.

Gurantz, O., and T. K. Odle. 2022. The impact of merit aid on college choice and degree attainment: Reexamining Florida's Bright Futures Program. *Educational Evaluation and Policy Analysis*. 44(1): 79–104. Available at: <doi.org/10.3102/01623737211030489>.

Hansen, D. M. 2020. Using artificial neural networks to predict matriculation of university

prospects. *Strategic Enrollment Management Quarterly*. 8(1): 21–32.
Hastie, T., R. Tibshirani, and J. Friedman. 2017. *The Elements of Statistical Learning* (2nd edition). New York: Springer.
Kuhn, M., and K. Johnson. 2013. *Applied Predictive Modeling*. New York: Springer.
Kuhn, M., and R. Quinlan. 2023. *C50: C5.0 Decision Trees and Rule-Based Models*. Available at: <topepo.github.io/C5.0/>.
Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. *R News*. 2(3): 18–22. Available at: <cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf>.
Pruim, R. 2018. *Foundations and Applications of Statistics*. Providence, RI: American Mathematical Society.
Shao, L., M. Leong, R. A. Levine, J. Stronach, and J. Fan. 2022. Machine learning methods for course enrollment prediction. *Strategic Enrollment Management Quarterly*. 10(2): 11–29.
Soltys, M., H. Dang, G. R. Reilly, and K. Soltys. 2021. Enrollment predictions with machine learning. *Strategic Enrollment Management Quarterly*. 9(2): 11–18. Available at: <aacrao.org/docs/default-source/sem/semq-0902-soltys.pdf>.
Truth and Reconciliation Commission of Canada. 2012. *Truth and Reconciliation Commission of Canada: Calls to Action*. Monograph. Winnipeg, Manitoba, Canada: Author.

About the Authors

Ashmita De

Ashmita De, Ph.D., is a Research Analyst in the Education Insights, Data and Research department at the Northern Alberta Institute of Technology (NAIT). She is part of the Enrollment Planning and Reporting unit, specializing in enrollment projections through advanced statistical and machine learning-based methods. De received her Ph.D. in Biomedical Engineering from the University of Alberta.

Alexander A. Ondrus

Alexander A. Ondrus, Ph.D., is a Senior Research Specialist in the Education Insights, Data and Research department at the Northern Alberta Institute of Technology (NAIT). He is responsible for the Enrollment Planning and Reporting unit in NAIT's Education Insights department. His primary interest is the application of statistical techniques, including machine learning, to better understand the student lifecycle. Ondrus received his Ph.D. in Mathematics from the University of Alberta.