Analysis of College Athletic Success using Machine Learning

Daniel Pate, Gerard Tetegan, Jacob Lara, and Lane Whitehead
Dr. M. Jahangir Alam, Department of Economics, Texas A&M University, Spring 2024

Abstract

This study improves upon the causal estimation workflow proposed by Anderson (2017) by incorporating machine learning (ML) algorithms. Five ML models were evaluated, and gradient boosting (GB) yielded the lowest standard errors for the estimated average treatment effects. GB effectively handles both linear and non-linear relationships in the data, enabling the algorithm to objectively identify trends and enhance the reliability and consistency of the outcomes.

Introduction

- Anderson (2017) used propensity score weighting to address endogeneity by estimating the likelihood of winning for each team and applying these scores to weight observations when estimating treatment effects.
- While Anderson employed logistic regression to predict propensity scores, machine learning algorithms have now been implemented for this purpose.
- Anderson (2017) explicitly implemented non-linearity in the logistic regression model. In contrast, we allow the machine learning algorithms to train on the data, identify trends, and detect non-linearity autonomously, making the process less subjective.

Literature Review

- Tu (2019) compared the effectiveness of multinomial logistic regression (MLR), bagging (BAG), random forest (RF), and gradient boosting (GB) algorithms in estimating generalized propensity scores (GPS) to achieve unbiased average treatment effects (ATEs) in observational studies.
- A Monte Carlo simulation was conducted to evaluate the performance of these algorithms using their GPS estimates. The assessment was based on bias and mean squared error (MSE) across various sample sizes and association scenarios.
- Results indicated that the gradient boosting (GB) algorithm provided the most accurate and reliable ATE estimates, exhibiting the smallest bias and MSE across all scenarios.

Reference Outcome (Python)

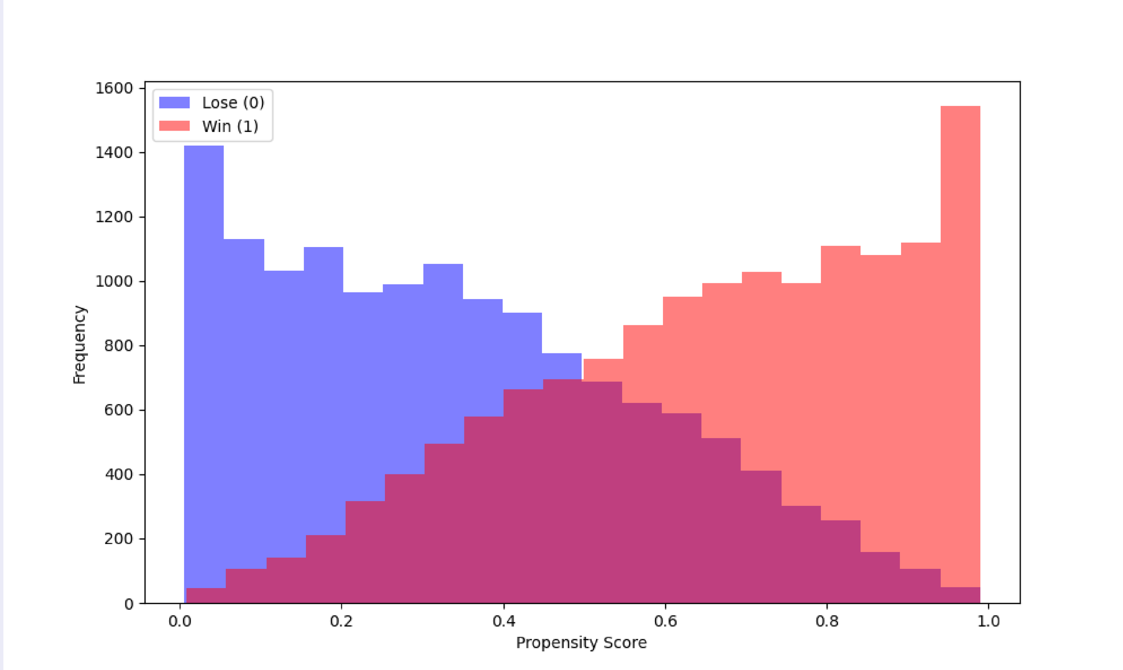
Anderson's (2017) evaluation was carried out in STATA and replicated in Python as shown below.

Outcome	Python Replication (2024)	
	Coefficient	Std Error
Alumni Athletic Operating Donations	191.2	65.0
Alumni Nonathletic Operating Donations	-137.4	96.1
Total Alumni Donations	267.4	267.1
Alumni Giving Rate	2.0e-4	7.0e-4
Academic Reputation	3.4e-6	1.6e-3
Applicants*	78.0	57.0
Acceptance Rate	-3.3e-4	1.6e-4
First-Time Out-of-State Enrollment	1.6	5.0
First-Time In-State Enrollment	12.6	6.4
25th Percentile SAT*	0.9	0.7

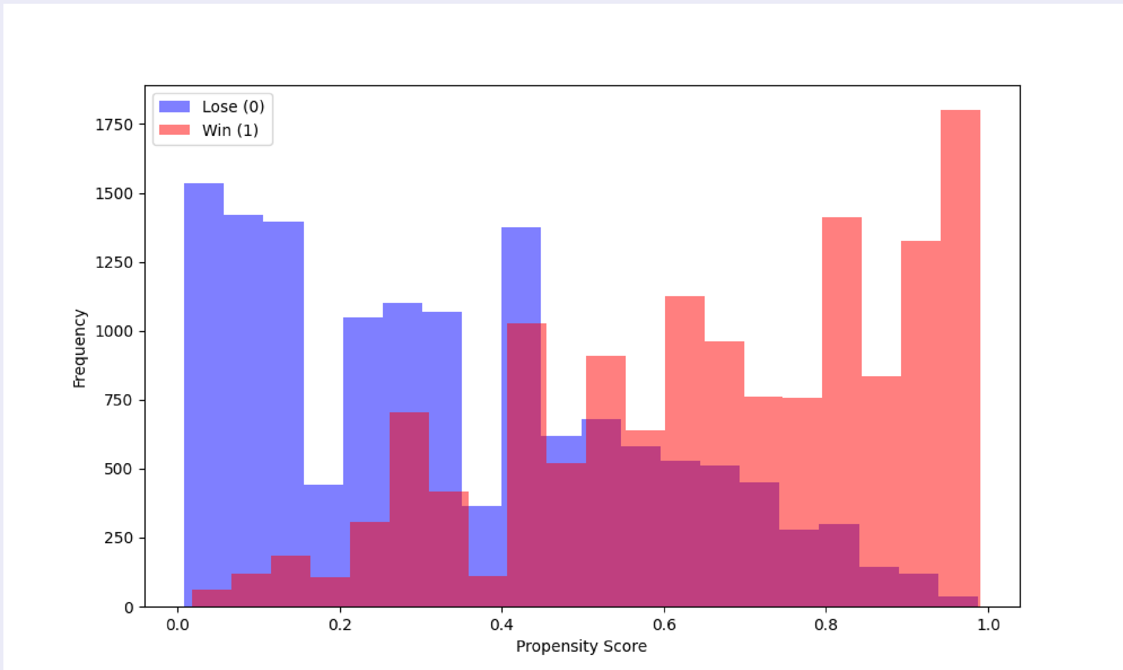
* coefficients not exactly matched with Anderson's outcome.

Propensity Score Distribution (Logit vs GB)

Logistic regression shows significant overlap between groups, indicating poor differentiation. In contrast, gradient boosting provides clearer separation, especially at higher scores, demonstrating its superior ability to handle complex, non-linear relationships and more accurately distinguish between winners and losers.



(a) Logistic Regression



(b) GB

Average Treatment Effects (GB Implementation)

Outcome	Coef.	SE	95% CI
Alumni Athletic Operating Donations	91.1	38.8	[15.1, 167.2]
Alumni Nonathletic Operating Donations	-40.8	107.8	[-252.1, 170.4]
Total Alumni Donations	170.5	253.3	[-326.1, 667.1]
Alumni Giving Rate	2.20E-04	6.50E-04	[-1.05E-03, 1.49E-03]
Academic Reputation	4.02E-03	1.56E-03	[9.60E-04, 7.96E-03]
Applicants	52.4	47.1	[-40.0, 144.7]
Acceptance Rate	-3.74E-04	1.65E-04	[-6.98E-04, -4.99E-05]
First-Time Out-of-State Enrollment	5.2	4.6	[-3.8, 14.2]
First-Time In-State Enrollment	14.5	6.4	[2.0, 27.0]
25th Percentile SAT	1.3	0.7	[-0.2, 2.7]

Conclusion

- Gradient Boosting yields more accurate propensity score estimates as evidenced by lower standard errors from the weighted regression.

References

- Anderson, M. (2017). The Benefits of College Athletic Success: An Application of the Propensity Score Design with Instrumental Variables. Review of Economics and Statistics, 99.
- Tu, C. (2019). Comparison of various machine learning algorithms for estimating generalized propensity score. Journal of Statistical Computation and Simulation, 89(4), 708-719.

Appendix