Dual Enrollment, First-Year Retention, and Graduation: Analyzing the Impact of Dual Enrollment on Student Success Outcomes at a Public Research University

Michael Tucker
University of South Carolina

Jonathan Poon
University of South Carolina

Sophie Greenwell
University of South Carolina

As more higher education institutions participate in offering college-level courses to high school students (often referred to as dual enrollment) it becomes increasingly important to evaluate the impact of participation in these courses on subsequent higher education student success outcomes, such as first-year retention and graduation rates. In recent years, there has been an increase in the body of literature devoted to this topic, however, much of this literature is aimed at analyzing the impact of dual enrollment courses on performance in community colleges. This study will examine the relationship between high school students taking dual enrollment courses and their later performance at a public, R1 university. Additionally, this study will employ Classification and Regression Trees (CART), a type of supervised statistical learning for identifying the success factors for dually enrolled students based on academic, demographic, and socio-economic features. This study intends to offer results that provide generalizable knowledge for institutional decision-making pertaining to dual enrollment and success in higher education.

Keywords: dual enrollment, first-year retention, graduation rate, machine learning, classification and regression trees, decision trees

INTRODUCTION

There is a present and urgent interest in understanding the impact of dual enrollment (DE) programs on college student success across the country, driven principally by recent governance from the National Center for Education Statistics (NCES) for higher education institutions to begin tracking additional information regarding these types of students as they matriculate at college (Institute of Education Sciences [IES], 2022). Institutional researchers are now challenged with offering analyses of complex higher education datasets, which are interpretable by those who make decisions at the executive level. Further emphasizing the current need to better understand the relationship between DE and student success, the Integrated Postsecondary Education Data System (IPEDS) Technical Review Panel (TRP) met in August of the previous year to discuss how to better collect additional data on DE through their web-based data collection system (IPEDS TRP, 2021).

Relatedly, in recent years, there has been significant scrutiny of the personal economic value of an undergraduate college degree for prospective students. A consistent, noteworthy finding has been that time to degree is one of the most impactful variables related to the economic value of an undergraduate college degree (Lobo & Burke-Smalley, 2017; Wright & Ross, 2021; Heckman & Letkiewicz, 2020). These authors have found that the longer it takes to obtain the intended undergraduate degree, the less economically sound it becomes for a student to have earned that degree.

For this study, the definition of DE used was provided by the NCES that "[dual enrollment is] an organized system with special guidelines that allows high school students to take college level courses" (NCES, 2009). One purported benefit of taking DE courses is that obtaining these college credits in high school can potentially expedite obtaining an undergraduate degree. Other benefits of DE include offering exposure to the rigor of university-level courses, increasing the range of academic content for high school students, and easing the transition from a high school to a university environment (Dare & Nowicki, 2015).

DE was initially devised in the 1950s at the University of Connecticut to keep talented high school seniors motivated to continue excelling in education and avoid academic boredom (Grant, 2019). In the following decades, DE would slowly expand to be offered in other areas in the United States but would remain a decentralized practice, governed by individual universities in partnership with their resident states. The first formal efforts to nationally standardize DE policy and practices occurred in 1997, which would eventually lead to the establishment of the National Alliance of Concurrent Enrollment Partnerships (NACEP) (*NACEP History*, n.d.).

Since that time, efforts to encourage high school student participation in DE programs across the United States have exponentially increased with a reported growth rate in DE of over seven percent annually (*NACEP History*, n.d.). There has also been considerable effort to expand access to these programs across the country, with more recent reports showing 34% of all high school students participating in some form of DE courses (NCES, 2019). Additionally, in 2017-18, 82% of all public high schools offered DE opportunities for students (NCES, 2020).

In the state of South Carolina, high school students have had the option to participate in DE since the early 90s. The Commission for Higher Education has stated that "the purpose [of dual enrollment courses] is to provide an avenue through which highly talented high school youth can earn college credit while simultaneously meeting high school graduation requirements…" (South Carolina Commission on Higher Education [SC CHE], 2004). In this policy document, DE was also purported to aid participants in obtaining their college degree in a timelier manner (SC CHE, 2004).

Much of the research related to DE at this time has been aimed at understanding its effectiveness in preparing high school students for success in the context of a community college environment (Andrews, 2004; D-Amico et al., 2013; Hunter & Wilson, 2018; Ganzert, 2014; Lawrence & King, 2018; Jones, 2017). This may be simply due to the comparative prevalence of community colleges to four-year institutions throughout the United States, however, this is not directly stated in any of the aforementioned articles. These studies consistently found that high school student participation in DE was positively associated with numerous student success metrics such as degree attainment and first-year GPA. Methodologically, these papers focused primarily on establishing the bivariate relationship between DE and various success metrics.

A consistent, albeit small, body of research has been devoted to better understanding the relationship of DE with four-year institution student success over the past decade. These studies have found similarly that participation in DE is positively correlated with student success metrics such as first-year GPA, graduation rate, and persistence (Bowers & Foley, 2018; Myers & Myers, 2017; Jones, 2014; Phelps & Chan, 2016). Among the literature, there remains a need to explore the nuance of this relationship as well as provide interpretable results to improve institutional effectiveness.

The primary goal of this study was to examine the relationship between student demographic characteristics of DE and non-DE participants and subsequent student success at a large, R1 four-year institution. It was theorized that there would be a statistically significant difference in first-year retention and six-year graduation rate between students who enrolled with DE participation, and those that did not.

The study offers methodological sophistication over traditional regression modeling by employing classification and regression trees (CART), an application of supervised machine learning, to increase the validity and interpretability of the findings (Brieman et al., 1984). More specifically, CART offers the advantage of visualization to accompany the analysis which can improve understanding and therefore, actionability of the results (Hastie et al., 2009). The results of the CART model development are intended to be employed by advisors as early as the beginning of the spring term of a student's first year of enrollment. The models developed for this study can help inform advisors of students who are both at risk of not returning in the fall and as those who are not on track to graduate within six years. Although CART has been used extensively across multiple fields since its initial development in the early 80s, it has seen little use in the field of higher education and institutional research. At the time of writing this paper, a quick search of the Education Resource Information Center (ERIC) using the following search criteria yields no pertinent results related to CART decision trees and DE:

> *("Classification and Regression Trees" OR "CART") AND ("Dual Enrollment" OR "Concurrent Enrollment")*

Decision trees, as they were used in this study, specifically refer to those of a statistical classification and regression nature. While the practical and organizational nature of displaying information in a tree fashion may date much further back in time, Wei-Yin (2014) explains that the first of these trees used for statistical analysis was in 1963 (Morgan & Sonquist) and was employed to inform conclusions regarding survey data on income and level of education. Wei-Yin (2014) goes on to explain that several statistical sophistications occurred, namely due to increases in computing power, from that time until the seminal work "Classification and Regression Trees" by Breiman et al. in 1984. While there have been continuous refinements of CART, as well as the development of other algorithms, CART remains a highly utilized methodology for developing decision trees that can handle mixed, dynamic datasets.

Although CART is a simplistic concept where the intent is to partition explanatory spaces into optimal rectangular sections, it offers several advantages over traditional regression methods (Hastie et al., 2009). It is flexible enough to handle a mixture of categorical and continuous data, both for explanatory and response variables. CART is also not affected by outliers, collinearities, or heteroscedasticity (Mubayi, 2017; Pittendrigh, 2016). It does not vary due to any monotone transformation of the explanatory variables in the tree, and, due to the growing and pruning process, can handle many explanatory variables while still producing interpretable results. In classification decision tree development, several algorithms work together to aid in the growing and pruning of a tree. That is, recursive algorithms that measure the complexity of the tree, the rate at which the tree misclassifies the data, and minimize the impurity of the tree's root nodes.

In classification decision tree development, several algorithms work together to prune a tree with a small misclassification error rate while also determining a level of complexity that can describe the pattern and nature of the data being presented. The idea of minimizing the misclassification error rate is analogous to the traditional confusion matrix used to measure the accuracy of a traditional simple or multiple logistics regression model. That is the number of observations misclassified within each of the tree's partition regions is divided by all observations presented to the decision tree. In conjunction with this effort, a method known as cost-complexity growing and pruning is an algorithmic process of cross-validating the observations within the explanatory space to determine what will optimally describe the data in a complex way while also reducing it in size for optimal interpretability. To grow the tree to optimal complexity, weakest link pruning is conducted to minimize the cost of how complex can the tree be to be able to describe the explanatory space that yields an outcome response.

METHODOLOGY

Data

This study used institutional-level, student academic data collected for the Fall 2013, Fall 2014, and Fall 2015 cohorts at the University of South Carolina Columbia Campus. A cohort was defined here in alignment with the Integrated Postsecondary Education Data System (IPEDS) as full-time, first-time, degree-seeking students (IPEDS, 2021). Additionally, the final dataset was trimmed to only include students who were considered residents of South Carolina. This ensured the integrity of the final dataset as complete DE academic records could only be reliably obtained for resident students. A student was considered full-time if they were reported as being enrolled in at least 12 hours of credits during the given term. Additionally, a student was considered first-time degree-seeking if they reported as having never attended another university or college and declared the intent to seek a degree upon admission. These cohorts were selected to ensure the ability to examine the six-year graduation rate, while still being reflective of current pedagogical practices. The resulting population included 7,420 students across the three cohorts.

Explanatory Variables

The primary explanatory variable of interest for this study was whether a student had ever participated in DE. Where possible, this was determined using the institutional data available for the University of South Carolina system. Within the University of South Carolina system, all eight campuses offer DE to local populations throughout the state. It is important to note that for a student to have participated in DE in the state of South Carolina, they must have met state-mandated requirements (SC CHE, 2004).

Additional variables considered included Gender, Race/Ethnicity Group, First-Generation Student, Pell Grant Recipient, High School Core GPA, and University GPA after the first term. Gender was defined in line with the IPEDS Glossary (Broyles, 1995) and was collected at the time of first enrollment. Race/Ethnicity Group was defined in alignment with recent institutionally adopted higher-education racial groupings piloted by the University of California (Lee, 2008). Following Lee's (2008) example, Asian students were excluded from the underrepresented minority grouping and instead constituted their demographic group in the study population. First-Generation Student was collected during student admissions and defined in alignment with the Federal TRIO Programs as a student whose parents did not complete a baccalaureate degree (Program authority; authorization of appropriations, 2011). Pell Grant Recipient referred to any students who received the undergraduate need-based Federal Pell Grant during their first academic year at the University of South Carolina (Federal Pell Grants: amount and determinations; applications, 2011). High School Core GPA was defined as the cumulative grade point average (GPA) calculated exclusively from courses required for high school graduation. University GPA was defined as the GPA calculated from courses taken at the enrolled institution during the starting fall term. All aforementioned variables acted as explanatory variables for subsequent statistical significance testing and CART modeling.

Response Variables

The response variables used in this study were First-Year Retention and Six-Year Graduation. First-Year Retention was defined in alignment with IPEDS as a student enrolled in a fall term cohort who then enrolled in the following fall term at the same university (IPEDS, 2021). Six-Year Graduation was also defined in alignment with IPEDS as to whether a student started and completed a bachelor's or equivalent at the same institution within six years (IPEDS, 2021). Within the data, successful re-enrollment or graduation was denoted as a 1, and unsuccessful re-enrollment or graduation was denoted as a 0, respectively; the average of these class values produced the rate of each success outcome and acted as response variables for significance testing and CART modeling.

Statistical Analysis

To help justify supervised machine learning model development, the DE and non-DE groups were first independently explored to ensure appropriate comparisons could be made between each cohort and the response variables in this study. This involved ensuring that the relative proportions of each demographic characteristic were consistent between DE and non-DE groups, due to the large difference in sample sizes between groups (Table 1).

The first primary goal of this study was to determine if there was a statistically significant difference in student success metrics at a public, R1 university between students who had participated in DE and those who did not. The proposed hypothesis was that the student sample across two populations, dual enrolled within the University of South Carolina System and with no record of DE, would have different success rates. That is,

$$H_o: \mu_{hw} = \mu_{ft}$$
$$H_a: \mu_{hw} \neq \mu_{ft}$$

where, μ_{hw} represents population success rate of Columbia campus new freshmen cohort who participated in DE in prior terms within University of South Carolina System; μ_{ft} represents population success rate of new freshmen cohort who had no record of participation in DE in prior terms within University of South Carolina System before attending the University of South Carolina Columbia campus; μ represents population success rate which will consist of First-Year Retention, and Six-Year Graduation.

Welch-Satterthwaite T-Test

To test these hypotheses and account for possible unequal variance among the sample groups, statistical inference tests were performed using SAS software version 9.4 on DE participation and each of the two-success metrics. The first step in this process was to determine if the sample variances were equal. Homogeneity of variance testing was conducted using the folded form F-test. If the variances between groups were equal, the Student's T-Test was employed to test the hypotheses. If sample variances were unequal, the Welch-Satterthwaite T-Test was employed to test the hypotheses while accounting for the unequal variance. The results of these tests were used to determine if there was a significant relationship between DE participation and each success metric. Once statistically significant differences were demonstrated between DE and non-DE groups for each of the two response variables, it was then appropriate to move to multivariate analysis. As such, the secondary goal of this study was to employ supervised machine learning to support the primary hypothesis and aid in the explanation of dually enrolled student success outcomes (Hastie et al., 2009).

Machine Learning – Decision Trees

As Hastie et al. (2009) suggest, the primary reasons why machine learning methods are advantageous are their predictive accuracy and interpretability. Machine learning methods attempt to decrease bias and variance over their traditional counterparts that may offer low bias but produce a large variance. Additionally, traditional regression models that incorporate many explanatory variables result in clunky, multi-dimensional equations as their output. These results can be difficult to interpret and are, consequently, less actionable in nature. The output of machine learning models provides highly interpretable visuals that allow for ease of use in decision making. Due to the highly statistically significant differences observed between DE and non-DE groups for the two response variables, the overall population of 7,420 was divided into 901 DE participants and 6,519 non-DE participants for supervised machine learning. A set of tree-based models were developed for each group independently. The intention of independently developing these models for each group was to compare the tree structure, splits, and terminal nodes (leaves) between the DE and non-DE group.

CART

CART is a binary recursive partitioning method for processing continuous and/or categorical data (Hastie et al., 2009). It was used in this study to identify the explanatory variables that most contributed to the variance observed for the two response variables, First-Year Retention and Six-Year Graduation. In practice, this involved the independent, sequential development of classification trees for each success metric, First-Year Retention followed by Six-Year Graduation. This process was conducted for both DE and non-DE subpopulations, which resulted in two sets of two classification trees (four trees total).

The following analyses were conducted on both DE and non-DE subpopulations. First, a CART model was developed incorporating all explanatory variables for each subpopulation; these models were built to learn the patterns of response variable First-Year Retention. A second CART model was then developed for each subpopulation on the same explanatory variables and response variable Six-Year Graduation. In the results section, these CART models were referred to as retention and graduation learners, respectively.

This study utilized the HPSPLIT procedure in SAS software 9.4 to produce CART decision trees for each response variable. Missing data for any cohort observation resulted in the deletion of that cohort observation, as per the default specification in the HPSPLIT procedure. This resulted in the deletion of 25 observations, due to an unknown High School Core GPA. A primary reason for the utility of HPSPLIT was its robustness in tree model diagnostic features. More specifically, these algorithms provide diagnostic results to determine whether the decision tree needs to grow in complexity or be pruned for optimal interpretability, and to avoid overfitting the data.

To ensure optimal pruning was reached, ten-fold cross-validation was employed. Ten-fold cross-validation is an algorithmic process that creates ten data sets with equally distributed random samples from the entire observation dataset and independent decision trees. Additional information on this optimization method can be found in the originally proposed theory by Breiman et al. in 1984. During tree growth, the impurities of each branch created on the tree were minimized. That is, splitting the data by an explanatory variable into a binary split that maximizes the number of observations that match the outcome response on the left or right side of the split. CART models were trained to predict unsuccessful re-enrollment or graduation (EVENT=0). High specificity, the reduction in instances of false negatives, was a primary diagnostic in assessing model validity. For each CART model, it was desired to achieve high specificity as this minimized the potential for advisors to erroneously intervene with a student who has intentions of returning the next fall terms. The method used for impurity reduction in this study was information gain, also known as cross-entropy (Hastie et al., 2009). Through these decision tree model diagnostics, this study produced a set of retention and graduation learners that describe which factors contribute to the success of DE and non-DE freshmen cohorts at the University of South Carolina.

RESULTS

Across the three cohorts, 12.14% participated in some form of DE before their first year enrolled at the university. The overall first-year retention rate for this group was 88.11%. The overall six-year graduation rate for the cohorts used in this study was 75.82%.

Descriptive statistics for explanatory and response variables, within the DE group and non-DE group, can be found in Tables 1 and 2. All proportions of the demographic variables were relatively similar between the groups of DE participants and non-DE participants. A simple visual comparison of the first-year retention rate and six-year graduation rate across DE and non-DE groups immediately provides some evidence for a noticeable difference (Figure 1). The results of subsequent Welch's T-Tests aided in providing significant evidence to support the alternative hypothesis. As mentioned in the methods, it was necessary to account for sample size differences between the DE and non-DE groups for the two response variables. The result of Welch's t-test for First-Year Retention indicated there was significant evidence to support the alternative hypothesis that retention rates between DE cohorts and non-DE cohorts are different ($df=1303.90$, $t = 4.62$, $p <.0001$). From exploratory observation, the first-year retention rate of DE cohorts being studied had a combined rate of approximately 92% while non-DE cohorts had a combined first-year retention rate of approximately 88%. The result of Welch's t-test of Six-Year Graduation indicated there was significant evidence to support the alternative hypothesis that graduation rates between DE cohorts and non-DE cohorts are different (df=1233.3, t = 4.78, p <.0001). From exploratory observation, the six-year graduation rates of DE cohorts being studied had a combined rate of approximately 82%, while non-DE cohorts had a combined first-year retention rate of approximately 75%. The significant results of Welch-Satterthwaite T-Tests on both response variables, seen in Table 3, justified to move to a more complex analysis of the data.

The CART retention and graduation learner development process occurred in two sequential stages for both DE and non-DE cohort groups: the retention learner stage, followed by the graduation learner stage. By way of the HPSPLIT function in SAS, the results of the DE and non-DE retention learners were determined first, followed by those of the graduation learners.

Through the aforementioned methods of growing and pruning, the optimal DE retention learner, seen in Figure 2 and Table 4, produced a misclassification rate of 0.0699, with a tree structure having three terminal nodes from a total of two levels of splits into the data. By observing the cost-complexity plot, seen in Figure 3, the misclassification rate was minimized at three leaves while also yielding the lowest cost-complexity value. A tree of this size also produced a specificity rate of 99.76%. As a result, the DE retention learner determined that a DE cohort's University GPA was the most important and the only variable that explained the cause for them to return or not return in the following fall term (Table 5). It is important to note the root node split was based on the criteria of the cohort University GPA being less than 2.428 (left split), or greater than or equal to 2.428 (right split). This right split was the first terminal node. By observation, approximately 94% of the DE cohorts in this right split returned the next fall term.

For the next group, the optimal non-DE retention learner produced a misclassification rate of 0.0909 with a tree structure having seven terminal nodes from a total of six levels of splits into the data (Figure 4 and Table 6). Through cost-complexity plot analysis, two additional, pruned trees with five and six leaves were observed to see whether optimal complexity and interpretability could be achieved (Figure 5). As a result, a tree of seven leaves was determined to be the most adequate. The final non-DE retention learner derived produced a specificity rate of 99.00%. The non-DE retention learner determined that a non-DE cohort's University GPA was the most important followed by their associated Race/Ethnicity Group, First-Generation Student, and High School Core GPA (Table 7). It is important to note the root node split was based on the criteria of the cohort University GPA being less than 2.12 (left split), or greater than or equal to 2.12 (right split). This right split was the first terminal node. By observation, approximately 92% of the non-DE cohorts in this right split returned the next fall term.

Through methods of growing and pruning previously mentioned, the optimal DE graduation learner produced a misclassification rate of 0.1698, with a tree structure having three terminal leaves from a total of two levels of splits into the data (Figure 6, 7, and Table 8). A tree of this size also produced a specificity rate of 99.86%. As a result, the DE graduation learner determined that a DE cohort's University GPA was the most important and the only variable that explained the cause for them to achieve graduation within six years (Table 9). It is important to note the root node split was based on the criteria of the cohort University GPA being less than 3.18 (left split), or greater than or equal to 3.18 (right split). This right split was the first terminal node. By observation, approximately 90% of the DE cohorts in this right split graduated within six years.

For the next group, the optimal non-DE graduation learner, seen in Figure 8, produced a misclassification rate of 0.1910 with a tree structure having 10 leaves from a total of nine levels of splits into the data. A tree of this size produced a specificity rate of 95.94%. As a result, the non-DE graduation learner determined that a non-DE cohort's University GPA was the most important followed by impactful importance from their associated High School Core GPA, First-Generation Student, and Gender (Table 11). It is important to note the root node split was based on the criteria of the cohort University GPA being less than 2.80 (left split), or greater than or equal to 2.80 (right split). This right split was the first terminal node. By observation, approximately 83% of the DE cohorts in this right split graduated within six years. Interestingly, by methods of cross-validation and cross-entropy seen in Figure 9, without human-supervised diagnostics the model yielded a non-DE graduation learner with 23 terminal leaves (22 splits). As discussed in the methods section, the primary goal of decision tree-based learning models is to achieve high complexity and optimal interpretability. It was determined through cost-complexity plot observation that a graduation learner describing the data with 10-15 leaves achieved almost near identical probabilities, misclassification rates and fit statistics as a graduation learner with 23 leaves. In a series of terminal node size comparison analyses, the reduction of leaves yielded similar results in graduation learner diagnostics. Based on these observations, 10 leaves appeared to be the most suitable for providing high complexity and optimal interpretability.

CONCLUSION

The mission of DE programs throughout the country has evolved to better suit the current needs of the education system. These programs began as individual, disjointed efforts to stifle boredom during senior year in groups of exceptional high school students, however many have now shifted focus to providing high school students the opportunity to begin earning college credit at a statewide level (Grant, 2019). Likely due to the increase in prevalence and use of these programs across the country, the NCES has started requesting additional information regarding this subpopulation of students (IES, 2022).

In South Carolina, dual enrollment has been offered for several years but little analysis of its impact on future student success metrics has been completed (SC CHE, 2004; D'Amico, 2013). While there have been a few somewhat analogous studies in neighboring states, it is critically important to better understand the specific successes and challenges that students face in the context of their learning environment (Partridge et al., 2021; Ganzert, 2014). This study has demonstrated that there is a significant relationship between DE participation during high school in South Carolina and subsequent success metrics at the state's flagship institution. Furthermore, implementing CART modeling as a sophistication of traditional regression techniques has provided new, highly interpretable insights into demographic and academic characteristics that impact a student's odds of returning after their first year in college, as well as their odds of earning an undergraduate degree within six years. When considering the primary classification variables that drive the node splits for each supervised decision tree between DE and non-DE students the most notable comparison occurs in the absence of any other demographic explanatory variables for the DE subpopulation of students. The DE retention learner tree showed that only University GPA was important for the development of the optimally pruned tree. The non-DE retention learner on the other hand included Race/Ethnicity Group, First-Generation Student, and High School Core GPA as important explanatory variables in the tree. One theoretical interpretation of the differences seen between these models is that DE acts as a "protective measure" against not re-enrolling after a student's first year. Previous studies have identified certain "risk factors" that can negatively influence a student's likelihood of being retained or graduating (An, 2012). This study would suggest that participation in DE before college helped to mitigate the influence of some of these risk factors causing students to not return after their first year. Similarly, the DE graduation learner found that, of the explanatory variables used, University GPA was the only variable of importance. For the non-DE graduation learner, the explanatory variables of High School Core GPA, First Generation Status, and Gender appear as variables of importance (Table 11). Once again, DE seems to guard against unfavorable results such as failure to return after the first year or failure to graduate within six years, in comparison to those without DE participation.

The CART models also serve the purpose of identifying at-risk students in future freshmen cohorts, during the beginning of the spring term. These predicted results would intend to help guide advisory efforts at the university and to ensure that no potentially at-risk students were omitted from academic advisor consideration. This methodology would ideally be replicated annually in the early spring to further refine the predictive capabilities of the CART models and provide information to academic advisors. The CART model procedure includes inherently advantageous outputs over traditional linear or logistic regression models. Unlike traditional regression method coefficients, the accompanying visuals from the CART procedure (Figures 2, 4, 6, and 8) aid in the explanation of the model's purpose and findings. Additionally, the rule sets generated (Tables 12-15) are more easily understandable to target audiences, such as executive-level leadership and academic advisors.

LIMITATIONS

The nature of any studies involving the comparison of DE and non-DE subpopulations will inherently lead to differing-sized datasets, oftentimes with differing variances of response value distribution. Additionally, the study was limited to resident students who participated in DE through the university system. Future studies would benefit from including data on DE participation for non-residents as well as resident students who participated in DE through other means outside of the university system. Several of the explanatory variables relied on self-reported data such as First-Generation Student, Race/Ethnicity Group, and Gender. It is also important to note that the individual student's experience cannot be fully explained by inferential statistics. It can only be used as guide to infer success of DE and non-DE students.

FUTURE ENDEAVORS

As this is the first paper to employ CART modeling to examine the relationship between DE and student success, further studies and collaborations can provide additional information integral to supporting students at institutions. It will be crucial to engage with constituents when making use of these supervised machine learners within the university's data warehouse. This step could allow for advisors and executives to track, predict, and report student success independently of institutional researchers. Additionally, expanding the explanatory variables to include the National Survey of Student Engagement (NSSE) or extracurricular data may also improve the learner's understanding of factors contributing to student success. Lastly, comparative analysis of different machine learning models, such as the Patient Rule Induction Method (PRIM), Random Forest, or Hierarchical Mixture of Experts (HME) can be conducted to determine which is most accurate in predicting student success.

ACKNOWLEDGMENTS

The authors would like to give a special thank you to Dr. Faxian Yang, Director of Institutional Research at the University of South Carolina. Without his steadfast and continuous guidance, we would not have been able to conduct this research.

REFERENCES

Andrews, H.A. (2004). Dual credit research outcomes for students. *Community College Journal of Research and Practice*, *28*(5), 415–422. https://doi.org/10.1080/10668920490444445

Bowers, D., & Foley, V. (2018). Advanced placement and dual enrollment as related to college readiness and retention at a Tennessee University. *Journal of Academic Administration in Higher Education*, *14*(1), 5–10.

Breiman, L., Friedman, J., Olshen, R.A., & Stone, C.J. (1984). *Classification and Regression Trees*. CRC Press LLC.

Broyles, S. (1995). *Integrated postsecondary education data system glossary*. National Center for Education Statistics.

D'Amico, M.M., Morgan, G.B., Robertson, S., & Rivers, H.E. (2013). Dual enrollment variables and college student persistence. *Community College Journal of Research and Practice*, *37*(10), 769–779. https://doi.org/10.1080/10668921003723334

Dare, L., & Nowicki, E. (2015). Conceptualizing concurrent enrollment. *Gifted Child Quarterly*, *59*(4), 249–264. https://doi.org/10.1177/0016986215597749

Ganzert, B. (2014). Dual enrollment credit and college readiness. *Community College Journal of Research and Practice*, *38*(9), 783–793. https://doi.org/10.1080/10668926.2012.719483

GovInfo. (2011). *Federal Pell Grants: Amount and determinations; applications, 20 U.S.C. § 107*. Retrieved from https://www.govinfo.gov/content/pkg/USCODE-2007-title20/pdf/USCODE-2007-title20-chap28-subchapIV-partA-subpart1-sec1070a.pdf

GovInfo. (2011). *Program authority; authorization of appropriations, 20 U.S.C. § 107a-11*. Retrieved from https://www.govinfo.gov/content/pkg/USCODE-2020-title20/pdf/USCODE-2020-title20-chap28-subchapIV-partA-subpart2-divsn1-sec1070a-11.pdf

Grant, K. (2019). *The development of concurrent enrollment at the University of Connecticut*. Retrieved from https://ece.uconn.edu/about/history/

Hastie, T., Friedman, J., & Tisbshirani, R. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Heckman, S.J., & Letkiewicz, J.C. (2020). Navigating risky higher education investments: Implications for practitioners and consumers. *Journal of Financial Counseling and Planning*, *32*(1), 131–145. https://doi.org/10.1891/jfcp-18-00002

Hunter, M.P., & Wilson, J.E. (2018). Dual enrollment and retention in Tennessee community colleges: Implications for practice. *Community College Journal of Research and Practice*, *43*(3), 232–236. https://doi.org/10.1080/10668926.2018.1428240

Institute of Education Sciences. (2022). *IPEDS coordinator workshop & state data conference.* [Unpublished conference proceedings].

Integrated Postsecondary Education Data System Technical Review Panel. (2021). *Report and Suggestions from IPEDS Technical Review Panel #63: Capturing and Clarifying Dual Enrollment Data (Part II)*. Retrieved from https://edsurveys.rti.org/IPEDS_TRP_DOCS/prod/documents/TRP63_Summary.pdf

Integrated Postsecondary Education Data System. (2021). *2020-21 Survey materials: Glossary*. OMB NO. 1850-0582 v.29.

Jones, S.J. (2014). Student participation in dual enrollment and college success. *Community College Journal of Research and Practice*, *38*(1). https://doi.org/10.1080/10668926.2010.532449

Jones, S.J. (2017). Supporting the mission through dual enrollment. *New Directions for Community Colleges*, (180), 75–83. https://doi.org/10.1002/cc.20283

Lawrence, T.B., & King, S.B. (2018). Dual enrollment participants' attainment of an associate's degree from the college in which they participated in dual enrollment as compared to nonparticipants' attainment. *Community College Journal of Research and Practice*, *43*(4), 307–310. https://doi.org/10.1080/10668926.2018.1463302

Lee, S.S. (2008). The de-minoritization of Asian Americans: A historical examination of the representations of Asian Americans in affirmative action admissions policies at the University of California. *Asian American Law Journal*, *15*(1), 129–152. https://doi.org/10.15779/Z388P3M

Lobo, B.J., & Burke-Smalley, L.A. (2017). An empirical investigation of the financial value of a college degree. *Education Economics*, *26*(1), 78–92. https://doi.org/10.1080/09645292.2017.1332167

Morgan, J., & Sonquist, J. (1963). Problems in the analysis of survey data, and a proposal. *Journal of American Statistical Association*, *58*(302), 415–434.

Mubayi, A. (2017). Computational modeling approaches linking health and social sciences: sensitivity of social determinants on the patterns of health risk behaviors and diseases. *Handbook of Statistics*, pp. 249–304. https://doi.org/10.1016/bs.host.2017.08.003

Myers, C.B., & Myers, S.M. (2017). Dual enrollment and undergraduate graduation rates in the United States: An institutional and cohort approach using the 2006-2014 IPEDS. *Research & Practice in Assessment*, *12*, 5–17.

NACEP. (n.d.). *NACEP's History*. Retrieved from https://www.nacep.org/about-nacep/history/

Partridge, M.A., Schaller, T.K., Berry, R.L., & Routon, P.W. (2020). The economic benefit from tuition savings for dual enrollment students in Georgia. *Journal of School Choice*, *15*(4), 655–667. https://doi.org/10.1080/15582159.2020.1865091

Phelps, L.A., & Chan, H.Y. (2017). Optimizing technical education pathways: Does dual-credit course completion predict students' college and labor market success? *Journal of Career and Technical Education*, *31*(1). http://doi.org/10.21061/jcte.v31i1.1496

Pittendrigh, A., Borkowski, J., Swinford, S., & Plumb, C. (2016). Knowledge and community: The effect of a first-year seminar on student persistence. *The Journal of General Education*, *65*(1), 48. https://doi.org/10.5325/jgeneeduc.65.1.0048

South Carolina Commission on Higher Education. (2004). *Statewide higher education policy for delivery and transferability of "dual enrollment" coursework offered in high schools*. Retrieved from https://www.che.sc.gov/CHE_Docs/academicaffairs/dualenrollment.pdf

The National Alliance of Concurrent Enrollment Partnerships. (2015). *Fast facts on dual and concurrent enrollment*. Retrieved from https://www.nacep.org/resource-center/fast-fact-on-dual-and-concurrent-enrollment/

U.S. Department of Education. National Center for Education Statistics. (2019). *Dual enrollment: Participation and characteristics*. (Report No. NCES 2019-176). National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubs2019/2019176.pdf

U.S. Department of Education. National Center for Education Statistics. (2020). *Dual or Concurrent Enrollment in Public Schools in the United States*. (Report No. NCES 2020-125). National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubs2020/2020125/index.asp

U.S. Department of Education. National Center for Education Statistics. (2009). *Public use data files and documentation (PEQIS 14): Dual enrollment programs and courses for high school students*. (Report No. NCES 2009-045). National Center for Education Statistics. Retrieved from https://nces.ed.gov/surveys/peqis/download/data/peqis14doc.zip

Wei Yin, L. (2014). Fifty years of classification and regression trees. *International Statistical Review*, *82*(3). 329–348. https://doi:10.1111/insr.12016

Wright, A.M., & Ross, M.M. (2021). Human capital investment for adolescents: Barriers and opportunities. *Journal of Financial Counseling and Planning*, *32*(3), 387–401. http://dx.doi.org/10.1891/JFCP-19-00088

APPENDIX

TABLE 1
EXPLORATORY VARIABLE DISTRIBUTION

Exploratory Variables	Dual Enrollment Participation (%)	No Dual Enrollment Participation (%)
Gender		
Female	56.6	53.37
Male	43.40	46.63
Race/Ethnicity Group		
Asian	2.77	4.97
Non-resident alien	0.33	0.09
Two or more	4.00	4.65
Unknown	0.00	0.57
Underrepresented minority	12.65	14.83
White	80.24	74.89
First Generation Student		
First generation	29.41	29.77
Non-first generation	70.59	70.23
Pell Grant Recipient		
Non-Pell Grant recipient	75.58	71.53
Pell Grant recipient	24.42	28.47

TABLE 2
STUDENT GROUP SUCCESS RATES

Response Variables	Dual Enrollment Participation (%)	No Dual Enrollment Participation (%)
First-Year Retention		
Retained	92.12	87.56
Not retained	7.88	12.44
Six-Year Graduation		
Graduated in six years	81.69	75.01
Did not graduate in six years	18.31	24.99

TABLE 3
T-TEST ANALYSIS SUMMARY

Response Variables	Folded F-Test Significance	Variances	T-Test Method	t Value	DF	T-Test Significance
First-Year Retention	<.0001	Unequal	Welch-Satterthwaite	4.62	1303.9	<.0001
Six-Year Graduation	<.0001	Unequal	Welch-Satterthwaite	4.78	1233.3	<.0001

TABLE 4
NODE INFORMATION FOR DE RETENTION LEARNER

Node	Path	Observations	Not Retained (%)	Retained (%)	Classification
0	Root Node	901	7.88	92.12	Retained
1	Root Node	901	7.88	92.12	
	University GPA < 2.49	55	40.00	60.00	Retained
2	Root Node	901	7.88	92.12	
	University GPA >= 2.49	846	5.79	94.21	Retained
3	Root Node	901	7.88	92.12	
	University GPA < 2.49	55	40.00	60.00	
	University GPA < 1.40	12	83.33	16.67	Not Retained
4	Root Node	901	7.88	92.12	
	University GPA < 2.49	55	40.00	60.00	
	University GPA >= 1.40	43	27.91	72.09	Retained

TABLE 5
VARIABLE IMPORTANCE FOR DE RETENTION LEARNER

Variable	Frequency of Splits
University GPA	2

TABLE 6
NODE INFORMATION FOR NON-DE RETENTION LEARNER

Node	Path	Observations	Not Retained (%)	Retained (%)	Classification
0	Root Node	6492	12.4	97.6	Retained
1	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	Not Retained
2	Root Node	6492	12.4	97.6	
	University GPA >= 2.12	6026	7.98	92.02	Retained
3	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA < 1.44	246	85.37	14.63	Not Retained
4	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	Not Retained
5	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA < 1.72	60	66.67	33.33	Not Retained
6	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	Retained
7	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA < 1.72	60	66.67	33.33	
	High School Core GPA < 2.74	12	33.33	66.67	Retained
8	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA < 1.72	60	66.67	33.33	
	High School Core GPA >= 2.74	40	75.00	25.00	Not Retained
9	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Asian, Non-resident Alien, White	113	53.98	46.02	Not Retained
A	Root Node	6492	12.40	97.6	

	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Two or More, Underrepresented Minority, Unknown	47	27.66	72.34	Retained
B	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Asian, Non-resident Alien, White	113	53.98	46.02	
	First Generation Student = Non-First Generation	78	44.87	55.13	Retained
C	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Asian, Non-resident Alien, White	113	53.98	46.02	
	First Generation Student = First Generation	78	44.87	55.13	Retained

TABLE 7
VARIABLE IMPORTANCE FOR NON-DE RETENTION LEARNER

Variable	**Frequency of Splits**
University GPA	3
Race Group	1
First Generation Student	1
High School Core GPA	1

TABLE 8
NODE INFORMATION FOR DE GRADUATION LEARNER

Node	**Path**	**Observations**	**Not Graduated (%)**	**Graduated (%)**	**Classification**
0	Root Node	901	18.31	81.69	Graduated
1	Root Node	901	18.31	81.69	
	University GPA < 3.18	217	45.16	54.84	Graduated
2	Root Node	901	18.31	81.69	
	University GPA >= 3.18	684	9.80	90.20	Graduated
3	Root Node	901	18.31	81.69	
	University GPA < 3.18	217	45.16	54.84	
	University GPA < 1.64	14	92.86	7.14	Not Graduated

4	Root Node	901	7.88	92.12	
	University GPA < 3.18	217	45.16	54.84	
	University GPA >= 1.64	203	41.87	58.13	Graduated

TABLE 9
VARIABLE IMPORTANCE FOR DE GRADUATION LEARNER

Variable	**Frequency of Splits**
University GPA	2

TABLE 10
NODE INFORMATION FOR NON-DE GRADUATION LEARNER

Node	**Path**	**Observations**	**Not Graduated (%)**	**Graduated (%)**	**Graduated**
0	Root Node	6492	24.89	75.11	Graduated
1	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	Not Graduated
2	Root Node	6492	24.89	75.11	
	University GPA >= 2.8	5368	16.86	83.14	Graduated
3	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA < 1.8	334	90.42	9.58	Not Graduated
4	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	Not Graduated
5	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA < 2.28	236	67.37	32.63	Not Graduated
6	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >=	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	Graduated
7	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	Graduated
8	Root Node	6492	24.89	75.11	

	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	Graduated
9	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	
	University GPA < 2.32	20	30.00	70.00	Graduated
A	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	
	University GPA >= 2.32	307	49.84	50.16	Graduated
B	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	Graduated
C	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = First Generation	100	48.00	52.00	Graduated
D	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	
	University GPA >= 2.32	307	49.84	50.16	
	High School Core GPA < 3.08	166	54.22	45.78	Not Graduated
E	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	

	Gender = M	327	48.62	51.38	
	University GPA >= 2.32	307	49.84	50.16	
	High School Core GPA >= 3.08	141	44.68	55.32	Graduated
F	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA < 3.20	71	23.94	76.06	Graduated
G	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA >= 3.20	56	46.43	53.57	Graduated
H	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA >= 3.20	56	46.43	53.57	
	High School Core GPA < 3.45	36	63.89	36.11	Not Graduated
I	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA >= 3.20	56	46.43	53.57	
	High School Core GPA >= 3.45	20	15.00	85.00	Graduated

TABLE 11
VARIABLE IMPORTANCE FOR NON-DE GRADUATION LEARNER

Variable	Frequency of Splits
University GPA	4
High School Core GPA	3
First Generation Student	1
Gender	1

TABLE 12
TERMINAL NODE INFORMATION FOR DE RETENTION LEARNER

Node	Path	Observations	Not Retained (%)	Retained (%)	Classification
2	Root Node	901	7.88	92.12	
	University GPA >= 2.49	846	5.79	94.21	Retained
3	Root Node	901	7.88	92.12	
	University GPA < 2.49	55	40.00	60.00	
	University GPA < 1.40	12	83.33	16.67	Not Retained
4	Root Node	901	7.88	92.12	
	University GPA < 2.49	55	40.00	60.00	
	University GPA >= 1.40	43	27.91	72.09	Retained

TABLE 13
TERMINAL NODE INFORMATION FOR NON-DE RETENTION LEARNER

Node	Path	Observations	Not Retained (%)	Retained (%)	Classification
2	Root Node	6492	12.4	97.6	
	University GPA >= 2.12	6026	7.98	92.02	Retained
3	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA < 1.44	246	85.37	14.63	Not Retained
7	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA < 1.72	60	66.67	33.33	
	High School Core GPA < 2.74	12	33.33	66.67	Retained
8	Root Node	6492	12.4	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA < 1.72	60	66.67	33.33	

	High School Core GPA >= 2.74	40	75.00	25.00	Not Retained
A	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Two or More, Underrepresented Minority, Unknown	47	27.66	72.34	Retained
B	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Asian, Non-resident Alien, White	113	53.98	46.02	
	First Generation Student = Non-First Generation	78	44.87	55.13	Retained
C	Root Node	6492	12.40	97.6	
	University GPA < 2.12	466	69.53	30.47	
	University GPA >= 1.44	220	51.82	48.18	
	University GPA >= 1.72	160	46.25	53.75	
	Race Group = Asian, Non-resident Alien, White	113	53.98	46.02	
	First Generation Student = First Generation	78	44.87	55.13	Not Retained

TABLE 14
TERMINAL NODE INFORMATION FOR DE GRADUATION LEARNER

Node	**Path**	**Observations**	**Not Graduated (%)**	**Graduated (%)**	**Classification**
2	Root Node	901	18.31	81.69	
	University GPA >= 3.18	684	9.80	90.20	Graduated
3	Root Node	901	18.31	81.69	
	University GPA < 3.18	217	45.16	54.84	
	University GPA < 1.64	14	92.86	7.14	Not Graduated
4	Root Node	901	7.88	92.12	
	University GPA < 3.18	217	45.16	54.84	
	University GPA >= 1.64	203	41.87	58.13	Graduated

TABLE 15
NODE INFORMATION FOR NON-DE GRADUATION LEARNER

Node	Path	Observations	Not Graduated (%)	Graduated (%)	Graduated
2	Root Node	6492	24.89	75.11	
	University GPA >= 2.8	5368	16.86	83.14	Graduated
3	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA < 1.8	334	90.42	9.58	Not Graduated
5	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA < 2.28	236	67.37	32.63	Not Graduated
9	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	
	University GPA < 2.32	20	30.00	70.00	Graduated
C	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = First Generation	100	48.00	52.00	Graduated
D	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Male	327	48.62	51.38	
	University GPA >= 2.32	307	49.84	50.16	
	High School Core GPA < 3.08	166	54.22	45.78	Not Graduated
E	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = M	327	48.62	51.38	
	University GPA >= 2.32	307	49.84	50.16	

	High School Core GPA >= 3.08	141	44.68	55.32	Graduated
F	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA < 3.20	71	23.94	76.06	Graduated
H	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA >= 3.20	56	46.43	53.57	
	High School Core GPA < 3.45	36	63.89	36.11	Not Graduated
I	Root Node	6492	24.89	75.11	
	University GPA < 2.8	1124	63.26	36.74	
	University GPA >= 1.8	790	51.77	48.23	
	University GPA >= 2.28	554	45.13	54.87	
	Gender = Female	227	40.09	59.91	
	First Generation Student = Non-First Generation	127	33.86	66.14	
	High School Core GPA >= 3.20	56	46.43	53.57	
	High School Core GPA >= 3.45	20	15.00	85.00	Graduated

FIGURE 1
VISUAL REPRESENTATION OF STUDENT SUCCESS RATES

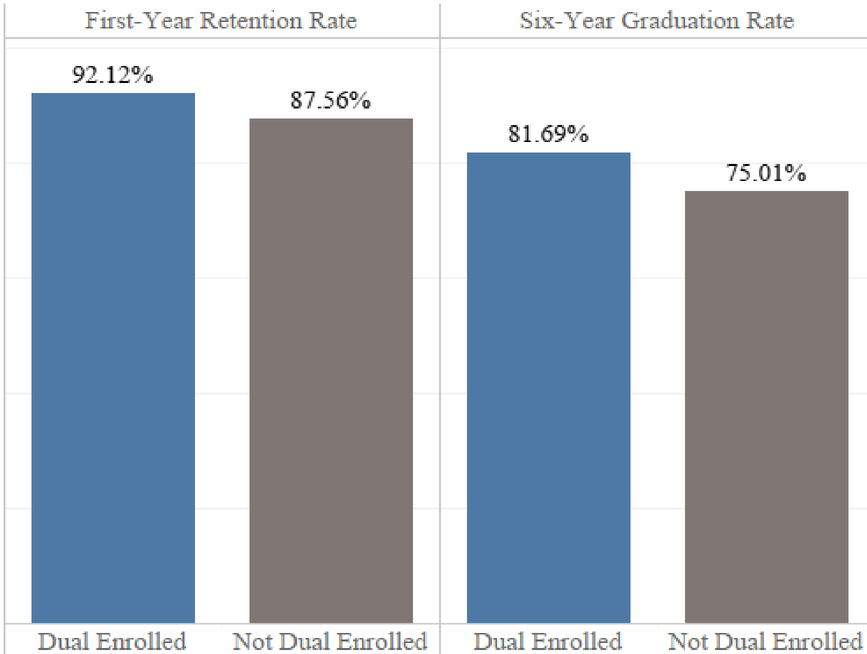

FIGURE 2
DE RETENTION LEARNER

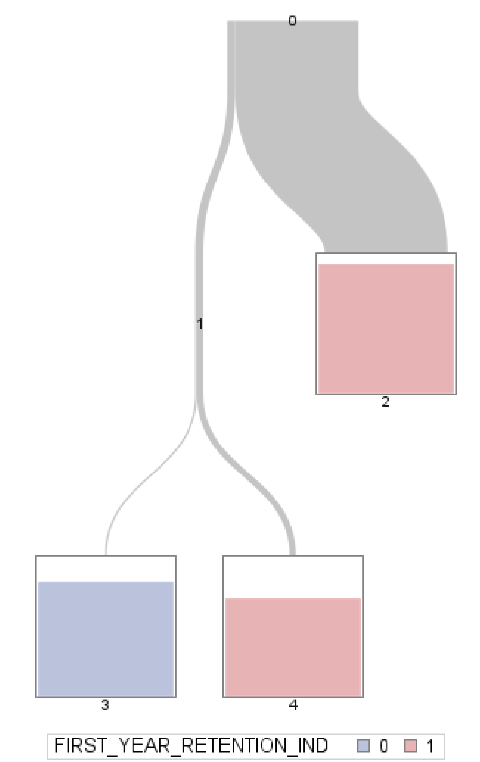

FIGURE 3
COST-COMPLEXITY ANALYSIS FOR DE RETENTION LEARNER

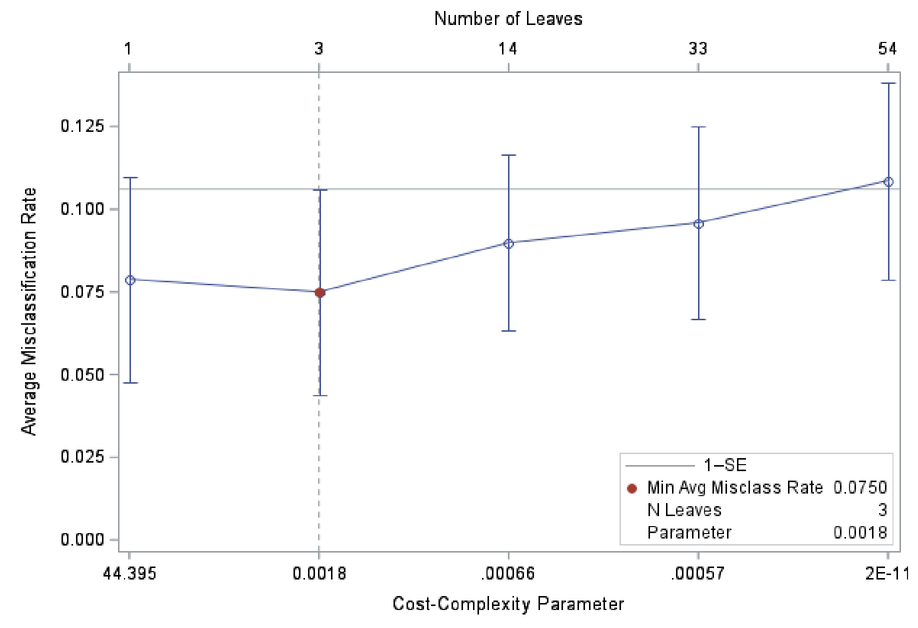

FIGURE 4
NON-DE RETENTION LEARNER

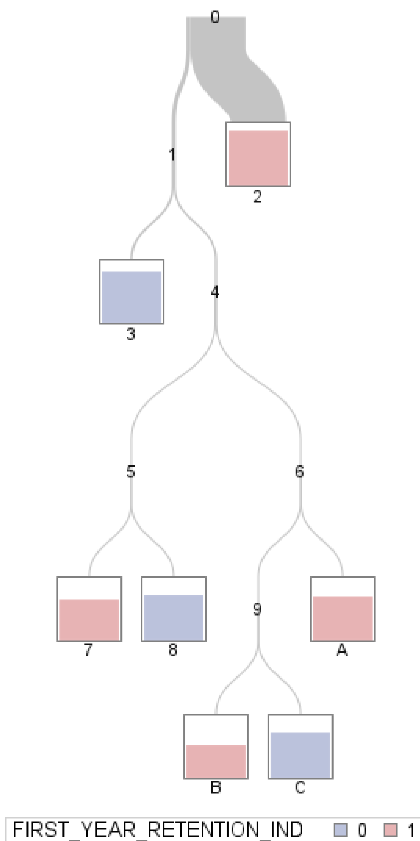

FIGURE 5
COST-COMPLEXITY ANALYSIS FOR NON-DE RETENTION LEARNER

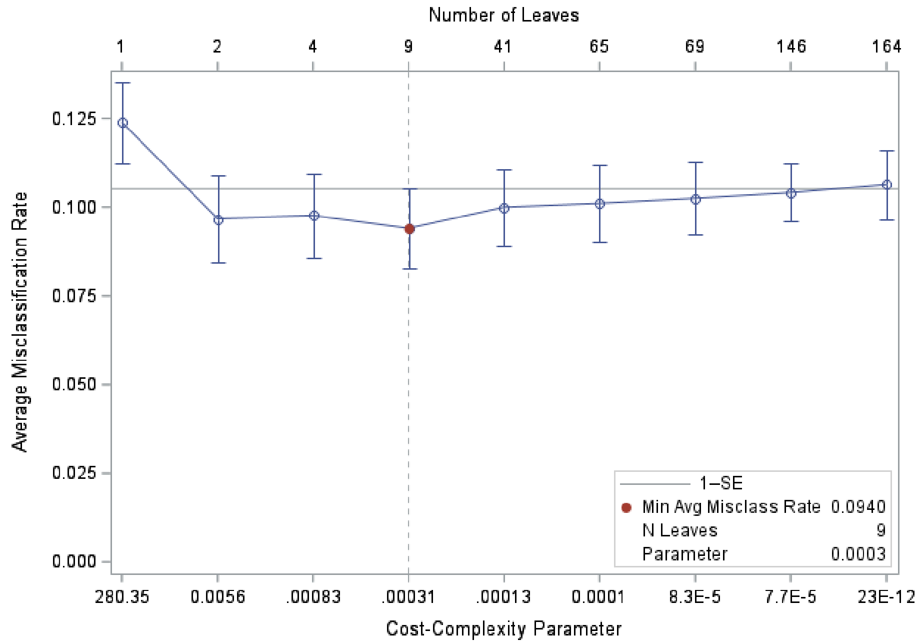

FIGURE 6
DE GRADUATION LEARNER

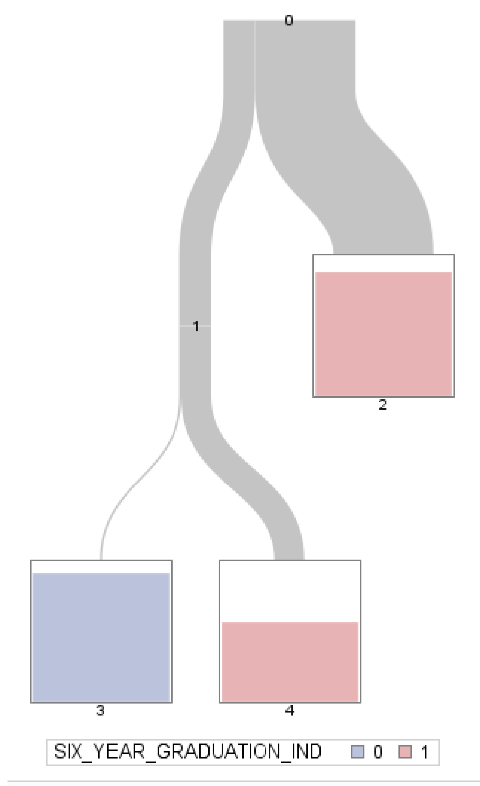

FIGURE 7
COST-COMPLEXITY ANALYSIS FOR DE GRADUATION LEARNER

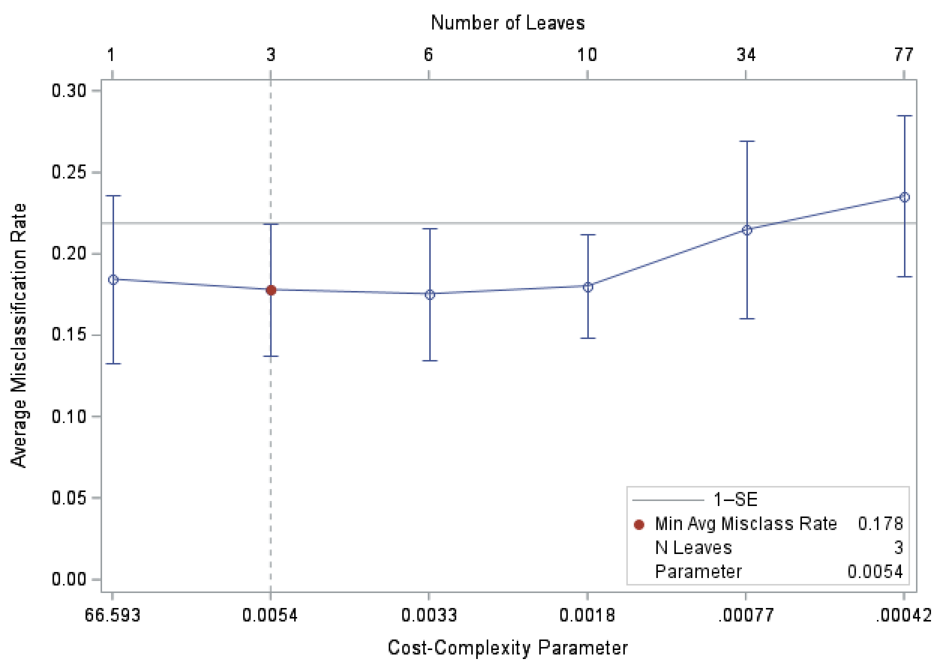

FIGURE 8
NON-DE GRADUATION LEARNER

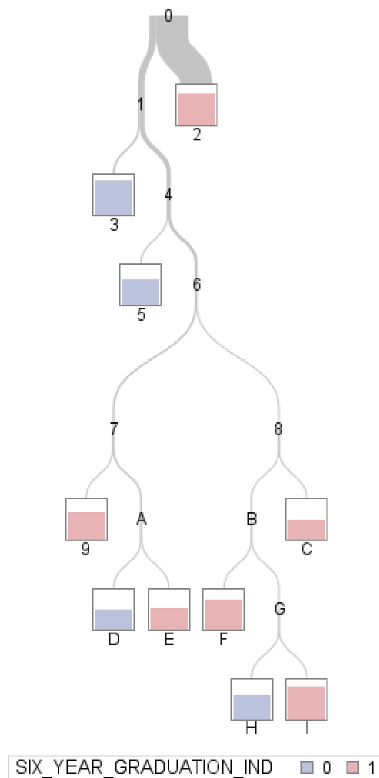

FIGURE 9
COST-COMPLEXITY ANALYSIS FOR NON-DE GRADUATION LEARNER