ABSTRACT

This study represents the portal for St. John Paul II College of Davao (SJP2CD). It is a responsive system to mobile and web platforms that enables the school community to operate efficiently throughout enrollment and enhances online access to academic information. The portal is built with Django and Python to create effective and timely management information and the database used is SQLite. Furthermore, the Classification and Regression Tree (CART) model is used for the automatic advising for the student enrollment process and the Linear Regression Model for the enrollment prediction. This portal is accessible via web browsers and mobile devices. During the deployment phase, when all required assessments are finished through extensive functionality testing and how the portal is helpful to all users, the development sprints’ goals were effectively achieved. In conclusion, the SJP2CD Portal has proved to meet all users' needs.

Key words: CART Model, Decision Tree, Linear Regression Model, Portal, Predictive Analytics

1. INTRODUCTION

Modern universities have transformed their websites by adopting advanced portal systems that extend beyond web-based interfaces and include mobile access [1]. These portals facilitate quick management of student records and data organization, improving service [2, 3, 4]. Universities make more informed strategic decisions by integrating data-driven decision-making into predicting and assessing critical performance measures [5]. The capacity of supervised learning algorithms, such as the Classification and Regression Tree (CART) model, to predict student outcomes and enhance decision-making makes them exceptional tools for interventions meant to improve school communities [6]. Aksenova (2006) emphasized enrollment prediction as a significant element in the planning process of educational institutions [7]. Viberg, Hatakka, Balter, and MAVOUNDI’S (2018) recent research highlights the transformative power of efficient data collection and analysis in guiding educational decisions, opening the door for predictive analytics to identify early risk behaviors among students and reveal complex data relationships [8]. This integrated approach emphasizes how crucial data analytics is to issue resolution and improving strategy for making decisions.

The preceding discussion emphasizes the crucial importance of data analytics in addressing difficulties within academic institutions, such as the case at SJP2CD, where the absence of enrollment forecasting creates substantial obstacles in resource allocation [9, 10]. Vanthienen and De Witte’s (2018) study reveals the significant influence of efficient data usage that improves the decision-making process in line with institutional needs [11]. In this case, predictive analytics is a substantial tool because it is used to forecast future occurrences, allowing for better decisions and interventions to enhance outcomes [9, 12].

As integrated platforms, portals reduce the costs and difficulties of creating standalone website development while simultaneously improving stakeholder collaboration, streamlining information retrieval, and diminishing overload among stakeholders [13]. Students greatly benefit from this system as it accelerates submission processes and expedites departmental evaluations [14]. Anent to this, laborious processes such as the lengthy and inefficient grade inquiry process, which, in older systems, require excessive labor, needs more data protection and lead to enormous paperwork [4, 15]. Thus, managerial strategy can be performed by effectively utilizing data analytics in these portals, which reveal patterns that simplify work, conserve resources, and reduce information overload in academic contexts [11, 16]. In addition to addressing current issues, this comprehensive approach shows how portals with data analytics support can change the frameworks used in educational settings for making decisions.

Portal technology has grown increasingly ubiquitous in the Philippine higher education environment, indicating a
considerable transition motivated by the digital revolution [15]. These portals serve as critical platforms for Higher Education Institutions (HEIs), delivering student services, including online course registration, curriculum access, admission monitoring, and secure communication channels [12]. Research shows that users are delighted with these technologies, demonstrating their effectiveness [17, 18]. In this case, the SJP2CD relies on traditional, manual management of student records and lacks a centralized portal system despite experiencing an increase in enrollment. While some departments retain their systems, a unified and easily accessible system leads to information overload and data access issues. This project introduces an integrated online solution to address these difficulties and boost efficiency by reducing the institution's reliance on traditional processes.

Organizations nowadays confront rising demands for information to support managerial decisions, which is vital to the productivity and development of educational institutions. Schools should use technology to improve decision-making processes in light of this. The portal technology is essential since it is a data repository for decision-making reports. The fundamental goal of this project is to develop an accessible online portal compatible with both mobile and web platforms, emphasizing providing students with convenient services and leveraging available data to make informed decisions. The Classification and Regression Tree (CART) method is employed for student advising, while Linear Regression forecasting is used to predict enrollment trends. The system also generates reports on student retention, which are used to inform future activities and academic preparation. Using these reports, the school administration can effectively intervene in teacher preparation, room assignments, section allocations, and other relevant decision-making procedures.

As outlined in the prior discussion, the development of the SJP2CD Portal is in keeping with the primary objective of improving academic operations and student convenience. By streamlining the remote enrollment and document submission procedures, eliminating the necessity for recurrent campus visits, and digitizing the document inventory, this portal aims to improve student engagement and satisfaction. These goals align with the primary objective of using available data for decision-making and student convenience. The portal's functionality, which allows system-generated reports and leverages data analytics, notably CART and Linear Regression predictive analytics, aligns with the emphasis on using data insights to make educated decisions. The description of user responsibilities and access fits well with the emphasis on system administration, guaranteeing that the portal serves multiple users and providing access to generated reports essential for making decisions. Furthermore, the focus on user login, report query filtering, and feature access is consistent with its primary purpose of expediting access to academic services and information for staff and students.

The project uses specific development tools to execute its implementation. Python is the programming language, Django is the web framework, SQLite is the database, and Matplotlib is used to generate graphical reports. The emphasis on security features such as Cross-Site Request Forgery (CSRF) prevention and reCAPTCHA version 3 ensures secure access and authentication for genuine users, which is essential for maintaining data integrity and user interactions within the portal. Furthermore, by requiring strong and unique passwords for access, password authentication is a vital security precaution for user accounts. Auto logout is also employed to improve security while the screen is not in use.

Online student record submission has several advantages for the institution, including increased efficiency, accessibility, security, and environmental benefits. It allows students to submit documents remotely and proceed with self-registration without visiting the premises during the enrollment period. It expedites administrative procedures by eliminating long queues of records submission on the day of enrollment, guarantees prompt access to records, upholds security guidelines, and supports environmental initiatives. However, regulatory compliance, backup and verification requirements, and the need for historical records may all contribute to the demand for tangible documentation at the beginning of classes. Furthermore, the document authentication mechanism, which allows for remote submission and subsequent verification of genuine hardcopy documents, aligns with the portal's purpose of permitting distant document submission while maintaining information correctness and authenticity. This strategy helps the admissions office manage the document inventory while assisting students in tracking their submissions.

**Objectives of the Study**

This project aims to create a predictive analytics-enabled mobile-web responsive portal for St. John Paul II College in Davao. To achieve the general objectives, the portal aims explicitly to:

1. Allow remote submission of admission requirements for students using Python programming language.
2. Automate student advising using CART Algorithm.
3. Access the student’s grade online using the SQLite database.
4. Send queries to the particular offices using application-to-application messaging.
5. Generate reports on student demographics, student retention, submitted requirements, all answered and unanswered queries filtered by date, frequently raised concerns, and enrollment trends using Matplotlib API.
2. METHODOLOGY

Figure 1: Conceptual Framework
As shown in Figure 1, when the user reaches the portal, he or she is directed to the login page, where a username and password are required. Users can log in using the given access credentials if they have a laptop that can access the web and iOS or Android devices that are connected to the internet. The password change occurs once users are able to reach the portal. Students may log in using their student ID number and password.

For the process of each user, the system admin creates, reads, updates, and deletes users and other data, which then be stored in the SQLite database. The enrolling personnel controls the courses which shows in the current semester once they have logged into the portal. This user finalizes the publishing of the class schedules, which are subsequently saved in the database. The IT personnel provides this user with temporary login information, which is changed once they have logged in. If a student is brand-new, they should be able to register remotely after they have logged into the portal. Suppose the student belongs to the old curriculum. In that case, the system offers the course evaluation that is analyzed using the CART algorithm, allowing the user to determine whether they should be promoted, placed on probation, given a final warning or recommended to shift—course prerequisites. In this sense, the school administration can offer appropriate action regardless of what the results reveal. The information necessary for report generation is the possible number of enrollees for the following semester, wherein the Classification and Regression Tree (CART) model is utilized for this functionality. This allows the school administration to attain an overview and prepare for the potential number of teachers, rooms, and sections needed.

2.1 Decision Tree Algorithm

Figure 2. Decision Tree Algorithm
The study employs the Decision Tree Algorithm in constructing a training model aimed at predicting the class of the target attribute through decision rules inferred from past data. This algorithm, a form of supervised learning and predictive analytics, offers an efficient framework to delineate choices and explore potential outcomes. Based on decision rules that are developed from the link between input and output properties, the dataset is divided into discrete branching segments [19].
Among various predictive analytics techniques, decision trees stand out as the most prevalent method in data prediction. They manifest as a structured flowchart tree wherein root node, splitting, decision nodes, leaf or terminal node, pruning, branch or sub-tree, and parent and child node denote internal nodes.

Figure 2 depicts the decision tree architecture for course evaluation. The term "course" here refers to the subject. The node is the offered course for the current semester. Each offered course should go through the entire process several times until all of them have been evaluated. Subsequently, the course's prerequisites, if any, are determined. If there are no prerequisites for the course, the student may enroll in it.

If the course has a prerequisite, it determines whether it has only one or more. If there is only one prerequisite, it is evaluated whether it is passed (grade greater than or equal to 75 or 3.5) or failed (grade is less than 74 or 5.0). If the student passes, he or she can enroll in the course; otherwise, the student cannot enroll. If it has more than one prerequisite, all grades must pass the prerequisite in order to enroll in the course; otherwise, the course cannot be enrolled.

CART, an adaptation of the decision tree algorithm, serves to address both regression and classification challenges. Decision trees, commonly employed in machine learning, are adept at tackling regression and classification issues. In classification, the aim is to forecast a label, whereas in regression, it forecasts a quantity [20]. The predictive model was implemented to anticipate the target variable—course evaluation. CART functions by progressively segmenting the sample space, aiming to augment the uniformity within the resulting subsets. The iterative process of division and assessment of each split's purity is pivotal in refining the predictive accuracy of the model [21]. This iterative refinement ensures that the model optimally captures patterns within the data, enhancing its predictive capabilities while minimizing potential overfitting, thereby promoting its adaptability to new datasets.

2.2 Linear Regression Model

Enrollment prediction accuracy is assessed by comparing the model and dataset's suitableness. One approach is to merge 44 data points from the first and second semesters into a single dataset for linear regression analysis. Alternatively, the data from each semester can be examined independently to create different datasets. This model assists in predicting enrollment changes within the institution, particularly by using data from the first and second semesters of each academic year to forecast the subsequent semesters.

For the unified dataset analysis for enrollment prediction, the relationship between time and the number of enrollees is described using the statistical measures Multiple R and R-squared (R^2) from a linear regression.

A perfect positive linear relationship, denoted by a correlation value of 1, means that the number of enrollees rises proportionately with the passage of time. On the other hand, a value of -1 denotes a perfect negative linear relationship, meaning that the two variables decline proportionately as one grows. There isn't any obvious linear relationship between the variables when the value is 0.

The multiple correlation coefficient (Multiple R) in this analysis is 0.93, suggesting that time and enrollment have a substantial positive linear relationship. The number of enrollees tends to expand positively with time. This analysis is consistent with the data shown in Figure 3, which shows a persistent upward trend in enrollment, indicating a continuous increase in the institution's student population. Apparently, the pandemic had a major impact on the enrollment reduction recorded between periods 31 and 38. However, since the outbreak, enrollment has progressively recovered. Furthermore, there is a regular trend in which enrollment falls every second semester, followed by an increase in the first semester. This pattern emphasizes the seasonality of enrollment data analysis.

The R-squared (R^2) value is a metric used to gauge the extent to which variations in the dependent variable (number of enrollees) can be accounted for by changes in the independent variable (time) which variations in the dependent variable (number of enrollees) can be accounted for by changes in the independent variable (time) within the linear regression model. With an R-squared value of 0.86 in this research, the linear regression model can account for almost 86% of the variations in enrollment over time. This demonstrates the linear regression model's good fit to the data and emphasizes the significant predictive value of time on enrollment.

In conclusion, the Multiple R-values of 0.93 suggest a strong positive linear relationship between time and the number of enrollees, indicating that as time increases, the number of enrollees tends to increase. The R-squared value of 0.86 indicates that a substantial portion (86%) of the variability in the number of enrollees can be explained by changes in time according to the linear regression model, reinforcing the strength of this relationship.
The F-statistic shows the degree to which the regression model well describes the data. It evaluates how the explained variation and the unexplained variance relate to one another. The value of 249.42 in this analysis signifies a strong fit of the regression model. This suggests that the independent variables collectively significantly impact the dependent variable, influencing its behavior.

The Significance F, often referred to as the p-value for the F-statistic, evaluates the likelihood of observing such an extreme F-statistic value if the null hypothesis were true. With a reported Significance F of 0.00, an exceedingly low value, it strongly supports rejecting the null hypothesis. This implies that the independent variables hold considerable and statistically significant influence over the dependent variable within the regression model.

In the final analysis, the F-statistic of 249.42, with a Significance F of 0.00, highlights two major findings: an excellent overall fit of the regression model and a substantial influence of independent factors on the dependent variable. These findings highlight the combined impact of the independent factors and show how well the regression model predicts the dependent variable.

As illustrated in Figure 4, for semester-wise enrollment prediction analysis, the multiple R coefficient for the first semester is 0.93, indicating a strong positive linear relationship between the number of enrollees; the R square value is 0.87, indicating that variations in time account for around 87% of the variability in the number of students. While these appear to be strong points, the F-statistic and p-value cast doubt on the interpretation. The regression model is not statistically significant at the traditional significance level of 0.05, as indicated by the F-statistic value of 134.10 and the accompanying p-value of 0.258. This suggests that there is not enough data to reject the null hypothesis, indicating that there is not enough variation in the dependent variable that the regression model can account for. Stated otherwise, there might not be a significant correlation between the predictors and the outcome variable.

In the second semester shown in Figure 5, the multiple R coefficient remained at 0.93, demonstrating a significant positive linear relationship between enrollment numbers and time. In a comparable manner, the R-squared value of 0.87 suggests that time accounts for about 87% of the variation in the number of participants. The F-statistic and p-value, on the other hand, change its significance. With an F-statistic of 131.62 and a corresponding p-value of 0.25, the regression model is statistically insignificant at the usual threshold of 0.05. This suggests that the predictors could not have a significant association with the outcome variable because there is not enough data to reject the null hypothesis.

Enrollment trends in the first and second semesters are similar, showing consistency in the observed patterns. However, while the first semester has much greater enrollment numbers than the second semester, both semesters show similar trends. The unified dataset analysis yields significant results, with a multiple R-value of 0.93 suggesting a strong positive linear association between time and enrollee numbers, as well as an 86% R-squared value indicating significant time-dependent variability. The regression model's excellent fit is indicated by the F-statistic of 249.42, which also suggests that the independent factors' combined influence on the dependent variables was significant. The null hypothesis was rejected with a p-value of 0.00, highlighting the substantial impact of independent factors on the dependent variable in the regression model. This analysis surpasses semester-wise analysis, demonstrating the model's reliability in forecasting enrolment trends and emphasizing the general impact of independent variables.
2.3 Web System Implementation

Figure 6. Dashboard page of the portal

Students, as users, can view personal information, view grades, view statements of accounts, enroll in courses, and send direct messages to specific offices as shown in Figure 6.

Figure 7. Grade page of the portal

Students can also see a more detailed breakdown of their overall grades as presented in Figure 7. Additionally, the student can print their grades. The tab on the left side shows the automated course availability for students, taking into account prerequisites and the retention policy. Moreover, the student is able to select which courses they want to enroll in. The system streamlines the enrollment process, enabling students to register for the available courses easily.

Figure 8. Enrollment page of the portal

Figure 9. Remote submission of documents page of the portal

Figure 9 shows the remote submission of requirements, allowing students to view the list of admission requirements, upload files, and await approval from admission personnel. On the other hand, admission personnel will review students' submitted requirements. The admission staff has the option to accept or reject documents. In case of rejection, students are prompted to submit a new file.

Figure 10. In-app message page of the system

Figure 10 demonstrates that the concern was successfully relayed to the appropriate office or department. This user can filter student concerns by date. The student shall check the in-app messages to view replies.

Figure 11. Report generation page of the system

Figure 11 illustrates the report on student demographics, where the number of female and male students is shown in a line graph; the pie chart shows the percentages for the current school year’s demographics for both male and female students.
for visual presentation. The admin can also generate statistics reports from this page and other decision-making-related reports.

### 2.4 Mobile System Implementation

Figure 12 displays the mobile responsiveness of the SJP2CD Portal particularly on a Realme 7 Android 10 device. The testing covered various portal features, confirming effective responsiveness and a positive user experience. The results highlight the portal's adaptability for smooth user interaction on diverse mobile platforms. The portal was also tested through various versions of mobile devices including Samsung Galaxy S22 Ultra, iPhone 14 Pro Max, and Huawei Nova 9, which all resulted to swift responsiveness.

### 3. RESULTS AND DISCUSSION

The White box testing results were positive, demonstrating the product's effectiveness and reliability. This testing helped to determine the project's internal functionality and response to inputs. Particularly, user functional tests conducted during the Module testing phase yielded positive results, showing the portal's usefulness. Moreover, the Black box testing evaluated the project's performance from a user's perspective, without inspecting its core code. It successfully emulated a wide range of user operations, such as downpayment processing, admission submission, student enrollment, communication submission, grade uploading, and course registration. Hence, testing confirmed the system's ability to handle user actions seamlessly, indicating that it is ready for real-world deployment.

The compatibility and functionality of the SJP2CD Portal were demonstrated by testing in many web browsers, including Google Chrome, Mozilla Firefox, Microsoft Edge, and Apple Safari. All functionality, including login processes, user interactions, and report generating, were successfully tested and worked as expected across all browsers. As a result of the findings, the portal performs well in a variety of web settings.

The mobile responsiveness test results for the SJP2CD Portal across multiple operating systems show strong performance. All important functions, including login and grade viewing, were thoroughly tested and exhibited effective responses. The portal demonstrated versatility across several mobile platforms and operating systems, as proven by successful results on devices including the Realme 7 with Android 10 OS, the Samsung Galaxy S22 Ultra with Android 12, the iPhone 14 Pro Max with iOS, and the Huawei Nova 9 with EMUI OS.

The SJP2CD Portal's internet speed test displays its performance on both web and mobile platforms under varied network conditions. This evaluation analyzes page loading times, upload and download speeds, and latency. The results indicate the portal can perform efficiently at various internet speeds while maintaining consistent responsiveness. The speed test on HighSpeedInternet confirms the system's capacity to execute all functions smoothly without experiencing delays or speed-related concerns.

The SJP2CD Portal's report creation capacity has been extensively assessed with a focus on both web and mobile platforms. The system consistently meets set standards across a variety of test situations, including demographics, student retention measures, and addressing concerns, indicating proficiency. Thus, the portal generates information on demographics, student retention, concerns, queries, submitted requirements, and enrollment forecasts, using historical data to make reliable predictions. However, it acknowledges instances of overfitting in prediction models, implying that adhering to the stated criteria is essential for obtaining accurate predictions.

Before moving on to beta testing, the consultant evaluated the alpha testing phase to make sure the program satiated quality standards. Following an in-depth assessment of its performance issues and reliability, the application was prepared for future improvements. After a successful demonstration and considerable testing, it was declared ready for widespread deployment. Beta testing was conducted in a real-world setting, with students, program leaders, IT personnel, registrars, admission staff, cashiers, enrollment personnel, and system administrators evaluating the portal's performance, dependability, quality, and efficiency. User training sessions improved familiarity with the application, with a focus on testing the web-mobile responsive portal, which met all beta testing requirements. Dr. April B. Evangelista, President of St. John Paul II College of Davao,
authorized the testing process, emphasizing the college’s dedication to quality assurance.

Moreover, the Beta testing resulted in better functionality and user satisfaction across multiple modules. User roles such as students, admissions personnel, IT personnel, system administrators, registrars, cashiers, and system administrators acknowledged sufficient functionality with opportunities for improvement. The range of input emphasizes the system’s overall performance, providing useful insights for optimization before widespread deployment.

4. CONCLUSION

From the testing and deployment that were done, it is evident that allowing students to submit requirements remotely through Python programming, reduces the need for physical presence and minimizes the time and effort spent on paperwork enabling students to submit requirements remotely. It also eliminates the inconvenience of visiting the registrar’s office when accessing grades online. Furthermore, the effectiveness of the CART algorithm in automating student advising starts from its systematic assessment of course prerequisites and other relevant factors. Using this algorithm, it guides informed enrollment decisions by evaluating parameters such as previous grades and available courses. Moreover, accessing grades online enhances student satisfaction by eliminating the process of waiting in long lines on the school premises. The immediate response to student queries through in-app messaging provides timely assistance and support, contributing to a smoother academic experience. Lastly, the report generation using the Matplotlib API establishes a strong basis for future planning of both academic and non-academic activities.

5. RECOMMENDATIONS

The recommendations below are based on the study’s findings and should be used by upcoming studies to enhance the SJP2CD Portal.

1. The SJP2CD Portal should be accessed using the Google Chrome web browser. For maximum performance, a minimum internet bandwidth of 5 Mbps is necessary. Furthermore, a download speed of 40 Mbps or greater is suggested for smooth downloads and optimal system performance.

2. The system’s improvements could include the implementation of a school feeder monitoring report to identify students’ former schools, thereby improving the admission process. In addition, to improve registrar and cashier tasks, the portal should make it simpler for teachers to submit grades and incorporate payment monitoring to track payments and outstanding balances. Customized reports for each program and department are recommended to improve enrollment efficiency. Furthermore, implementing a two-way authentication factor would improve system security.

REFERENCES


