Isabela State University Linker:
Journal of Education, Social Sciences, and Allied Health
Volume 1, Issue 1
ISSN (Print): 3082-3692 ISSN (Online): 3082-3706
DOI: https://doi.org/10.65141/jessah.v1i1.n9

Predicting Critical Courses in Retention of Bachelor of Science in Architecture

Ladi Kyla N. Cole[1], Edwin R. Arboleda[2], Ma. Angelica S. Madriaga[3], Lyka Rossele R. Nuestro[4]
Department of Computer, Electronics and Electrical Engineering, Cavite State Univerisity[1,2,3,4]
edwin.r.arboleda@cvsu.edu.ph

RESEARCH ARTICLE INFORMATION	ABSTRACT
Received: May 26, 2023 **Reviewed:** June 25, 2023 **Accepted:** June 05, 2024 **Published:** June 28, 2024 	The study of architecture can be said to be challenging, interesting, and not an easy profession because having artistic skills is one of the most important parts of this course as it requires a lot of drawing skills and creativity. To predict and identify courses that are necessary for BS Architecture retention, this study proposed a unique architecture that makes use of educational data mining techniques. This paper presented an experimental investigation based on actual data from the BS Architecture course at Cavite State University. In this study, machine learning models were used to estimate the ranking of subjects that are important for students to pass based on the information on previous academic performance. Furthermore, mathematics performance had a significant impact on the academic progress of architecture students. By analyzing this proposition, this paper contributes to the current topic of the relationship between important subjects in the retention of BS Architecture and undergraduate academic success. The possibility of using grades in mathematics and major disciplines as a predictor of academic achievement in undergraduate architecture programs is one of the problems identified by these findings. In addition, this work uses WEKA as data mining software, and the modeling technique used is Random Forest. Overall, the model can be a very useful tool for completing the program.

Keywords: *academic success, BS Architecture, critical courses, retention, and WEKA*

Introduction

Every architecture school has a curriculum that must be followed, and after passing the tests, the student is qualified to practice architecture and receive a Bachelor of Architecture degree. Over the period of ten semesters and five years, multiple subjects are taught and studied, but there are a few areas that must be thoroughly studied, and these subjects serve as the foundation for one's architectural career. Some topics are extremely important and must be extensively studied by all architecture students. Neither a masterpiece in architecture nor an excellent architect ever existed without any knowledge of the subjects that needed to be implemented in a design. Architecture is not just about designing a building, it is about bringing a structure to life with all the necessary stuff just done right (Khodbole, 2020). Architectural education develops students' ability to conceptualize, design, understand, and realize the act of building within the context of architectural practice, which balances the tensions between emotion, reason, and intuition, and gives physical form to the needs of society and the individual (UIA, 2017). The profession of architecture faces challenges that threaten its traditional role. Academicians agree that some of the methods used in architectural education do not reflect real professional practice. Architectural education is considered to be a complex process and its creative demands must be supported by an understanding of art, science, psychology, mathematics, and engineering (Kurt, 2009). Furthermore, architects are facing a challenge that is characterized by creating environments that support, enhance, and celebrate human activities (Antoniades, 1992).

Many public colleges in emerging countries struggle to predict undergraduate students' academic success. Despite the fact that these institutions store enormous amounts of educational data, they are unable to identify students who are at risk of dropping out (Pradeep & Thomas, 2015). This paper discusses subjects that are critical for predicting and identifying students who are most likely to fail. The issue statement was to identify the subjects that have a significant impact on students enrolled in BS Architecture using the academic data that were obtained.

In the literature, there are different meanings of the term dropout and ways to calculate its rates. In Brazil, such a problem is not new (Lobo, 2012). The phenomenon of retaining students in university education is extensive and connected to policies concerning access and the selection of higher education, which is indicative of the reality that there are high school graduates who lack the necessary skills, conditions, capacities, aptitudes, or competencies to pursue their university studies (Donoso & Schiefelbein, 2007). On the contrary, public universities have explored dropout and its measurement with a unique approach in recent years (Catarina et al., 2009; Barroso & Falcão, 2004; Soares, 2006; Silva Filho et al., 2007; Andriola, 2009; Dias et al., 2006). As a result, those studies looked at dropouts from three perspectives: (i) socioeconomic reasons (keeping a job to live or support a family; the influence of family in career decisions); (ii) vocational reasons (disappointment with erroneous course selection); and (iii) academic reasons (failure in initial courses, poor academic background, difficulties in professor/student relationships or with colleagues).

WEKA is a collection of machine learning algorithms and data processing tools for data mining tasks. It includes data preparation, categorization, regression, clustering, mining of association rules, and visualization tools (Frank et al., 2017). Data

mining is the process of extracting significant (non-trivial, implicit, previously unknown, and possibly valuable) information from large amounts of data (Aldowah et al., 2019). Data mining is the process of extracting a huge quantity of information from a set of data. Educational systems are evolving, and various technologies are being used to keep up with how individuals learn. Several solutions are now utilized to handle massive amounts of data scattered across multiple databases. EDM is designed to help educators, students, academic administrators, governments, and society gain access to such information (Baker et al., 2011; Baker & Yacef, 2009; Mahajan & Saini, 2020; Romero & Ventura, 2013). Many studies that used EDM approaches to predict student performance have aimed to improve course quality (Alyahyan & Düştegör, 2020; Huang, 2011). Academic achievement, achievement of learning objectives, development of desired skills and competencies, contentment, tenacity, and post-college performance are all examples of academic success (York et al., 2015). The Bachelor of Science in Architecture program prepares students to continue their education in a professional graduate program and earn a Master of Architecture. These degrees enable graduates to engage in limited architectural practice but do not fully prepare them for architectural licensure. Students in this program are informed that in order to be certified for architectural registration and professional practice, the National Council of Architectural Registration Boards (NCARB) requires a recognized professional degree and a broad architectural education. (Tika, 2021).

In an educational setting, grading is a primary function, effect, and outcome of assessing and evaluating students' learning. This study covered three key aspects of student grading: definition, purpose, and rationale. The concept of acceptable grading in literature, as well as its practice, were described in the definition of grades. Grading was defined in terms of providing feedback, identifying exceptionalities, and motivating students (Domfeh, 2015). The recommended approach was to assess whether it is necessary to retake the subjects that are critical for passing the course and to provide further knowledge on the major impact of subjects taken for retaining a Bachelor of Architecture at Cavite State University. It is important to determine which classifier is the best to utilize in data mining.

Methods

Research Design

This studyn used a quantitative, quasi-experimental research approach to investigate the grades of students enrolled in the Bachelor of Science in Architecture program. The purpose of a quasi-experimental design is to identify the cause-and-effect relationship between an independent and dependent variable (Thomas, 2020). Quantitative research method is defined as the process of gathering numerical data and evaluating it using mathematical approaches, particularly statistics, in order to understand a problem or phenomenon (Apuke, 2017). The design was chosen as a means of identifying the significant association between passing their first year in their course and the most critical subjects that they are required to pass by studying the numerical grades of all participants enrolled in the years 2012–2019.

Research Participants

The study's participants were BS Architecture students from Cavite State University's College of Engineering and Information Technology in Cavite Province, Philippines. They were the students who took the subjects listed on the attributes from

the year 2012 to 2019. The study comprised 142 students, 72 of which were graduates and 70 were not graduates.

Data Description

The data from the Department of College of Engineering and Information Technology at Cavite State University in Cavite, Philippines, were used to train the created models. The data collected covered 142 individuals who finished their undergraduate studies between 2012 and 2019. This provided information about the previous passing of their first year in their chosen course and identified critical subjects for each student to succeed. Before entering the data, it was cleaned. Cleaning data prepares it for the data mining process, which extracts the most important information from the data collection (Mesevage, 2020). The information was tallied and entered into WEKA. WEKA has data preprocessing, classification, regression, clustering, association rules, and visualization capabilities (Mo, 2019). The researchers eliminated the data of those students who did not complete the subject from the attributes; thus, some information was missing. During the current study's model creation phase, only 142 students' data were employed.

Research Procedures

Permission was acquired from the respective academic department prior to data collection. In order to comply with ethical guidelines, the administrators in charge of the database anonymized each student's academic records. After that, the anonymized data was sent to the co-author. This made any kind of connection with anyone impossible. The concerns regarding admission criteria and access to student admittance information were handled by senior academics and top university authorities. There was also an internal method for cross-checking marks issued for each course taken by each student. This served as quality control that ensured objectivity in the measurements used to assess students' academic achievement. The information was tallied and entered into WEKA.

Data Analysis

The WEKA online tool was utilized to find a connection between successful first-year BS Architecture students and the most critical subjects they must pass. It consisted of 16 data records, 15 of which were selected subjects taken during their first year (characteristics), and the latter one was their graduation record (class).

Ethical Considerations

In conducting the study on predicting critical courses in the retention of BS architecture students, the authors applied ethical considerations regarding the involvement of the participants in this study. This includes taking steps to ensure that any personal information collected from participants was kept confidential and not disclosed to anyone who was not directly involved in the study. Participants were informed of the measures taken to protect their privacy and reassured that their personal information would not be shared with others. Additionally, prior to the study, full consent was obtained from the participants, who were students pursuing Bachelor of Science in Architecture at Cavite State University, Cavite, Philippines. In handling the data, all data collected were stored securely, and a high level of confidentiality was ensured. Any forms of misleading information, as well as biased representations of the main data findings, were avoided.

On the other hand, maintaining the highest level of objectivity in discussions and analyses throughout the study was considered. Any forms of affiliation and possible conflicts of interest were acknowledged. Furthermore, the study was done with honesty and transparency.

Results and Discussion

From the data shown in Table 1, there were 142 students involved in the study, 72 of which were graduates and 70 are not graduates.

Table 1. Class Discretization

Class	No. of students
Graduate	72
Not Graduate	70

Critical Subjects in Retention of BS Architecture

The data included 15 input attribute (independent) variables that were important in BS Architecture retention, as well as one output as a class (dependent) variable. The input variables were the grades they received in those subjects when they were still in school.

The record of whether or not each student is a graduate after completing the architecture program was utilized as a measure of academic success for the output variable. Each student was assigned to one of two categories: 'Graduate' or 'Not Graduate'. Table 1 shows the specifics of each class. It is important to remember that one of the criteria used to select students for the Architecture program included earning a minimal grade. As a result, this criterion was utilized to assess critical academic subjects. Table 2 lists the input and output variables that were utilized to generate the models.

Table 2. The Pre-Processed Input and Output

Attributes	Description (Domain)
ENGL 1	Student's grade (numeric from 5 to 1)
MATH2A	Student's grade (numeric from 5 to 1)
ARCH 21	Student's grade (numeric from 5 to 1)
ARCH 22	Student's grade (numeric from 5 to 1)
ARCH 50	Student's grade (numeric from 5 to 1)
ARCH55	Student's grade (numeric from 5 to 1)
PHED 1	Student's grade (numeric from 5 to 1)
ENGL 2	Student's grade (numeric from 5 to 1)

MATH9	Student's grade from 5 to 1)	(numeric
PHYS 1	Student's grade from 5 to 1)	(numeric
ARCH 23	Student's grade from 5 to 1)	(numeric
ARCH 24	Student's grade from 5 to 1)	(numeric
ARCH60	Student's grade from 5 to 1)	(numeric
ARCH 65	Student's grade from 5 to 1)	(numeric
PHED 2	Student's grade from 5 to 1)	(numeric

Processed Dataset in WEKA

Random Forest is the most appropriate classifier, as seen in Figure 2. In machine learning, a classifier is an algorithm that automatically sorts or categorizes data into one or more "classes" (Dutta, 2022). The data was cross-validated using four (4) methods. Cross-validation is a technique in which the model is trained on a subset of the data set and then evaluated on the other subset (Kumar, 2021). The correctly classified instances are 79.5775% while the incorrectly classified instances are 20.422%.

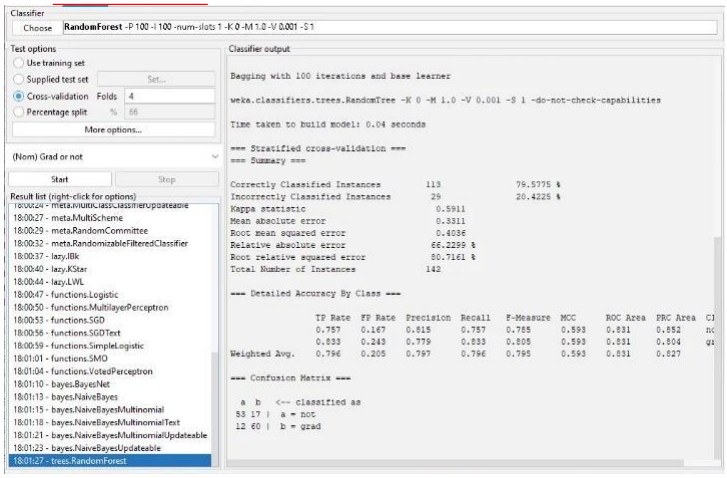

Figure 2. *Cross-Validation in WEKA*

Graphical Visualization of Processed Attributes

WEKA analyzed the processed data using various data mining techniques such as visualization. Data visualization is the visual presentation of data or information. It is typically visualized in the form of charts, photos, lists, infographics, and other visual elements (Nediger, 2020). The processed properties are visualized in a two-dimensional graphical representation in Figures 3 and 4.

Using two or more associative relations, information was extracted from the data set. During this process, the researchers attempted to visualize the essential subjects

in the retention of BS Architecture. After processing and computing the grades and population data, the attributes were sorted to find the most influential subject.

Figure 5 shows that ARCH 23 ranked first with a score of 0.17465, followed by PHYS 1 with a score of 0.15493, both PHED 2 and ENGL 1 with a score of 0.14789, ARCH 65 with a score of 0.14648, MATH 9 with a score of 0.12535, ARCH 24 with a score of 0.11127, ARCH 60 with a score of 0.09859, MATH 2A with a score 0.08451, ENGL 2 with a score of 0.06901, ARCH 55 with a score of 0.05493, PHED 1 with a of 0.0507, ARCH 50 with a score of 0.04085, ARCH 22 with a score of -0.00141, and ARCH 21 with a score of -0.01127.

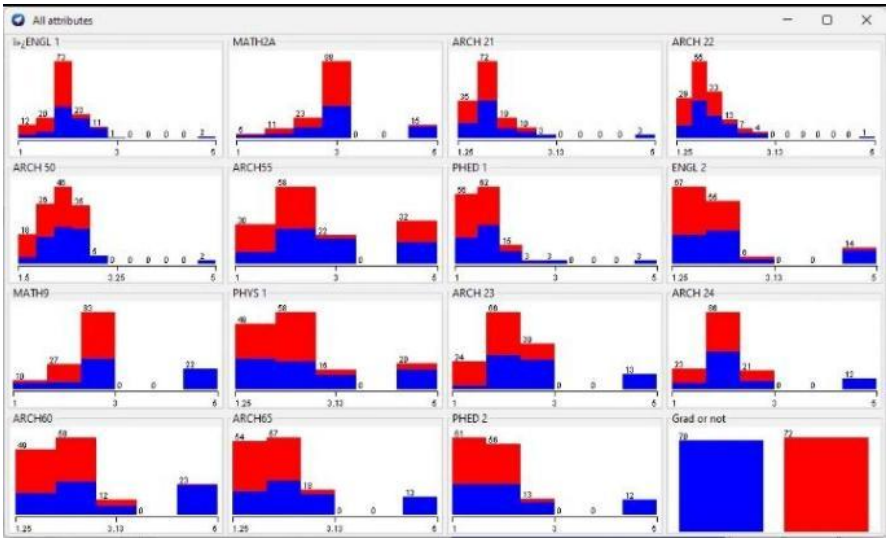

Figure 3. *Graphical Visualization*

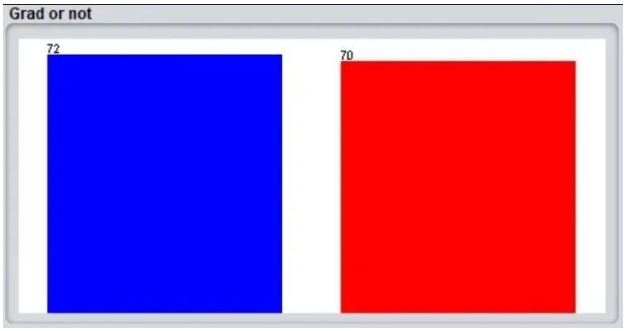

Figure 4. *Graphical Visualization of Graduate or Not Graduate*

```
Ranked attributes:
 0.17465  11 ARCH 23
 0.15493  10 PHYS 1
 0.14789  15 PHED 2
 0.14789   1 ï»¿ENGL 1
 0.14648  14 ARCH65
 0.12535   9 MATH9
 0.11127  12 ARCH 24
 0.09859  13 ARCH60
 0.08451   2 MATH2A
 0.06901   8 ENGL 2
 0.05493   6 ARCH55
 0.0507    7 PHED 1
 0.04085   5 ARCH 50
-0.00141   4 ARCH 22
-0.01127   3 ARCH 21

Selected attributes: 11,10,15,1,14,9,12,13,2,8,6,7,5,4,3 : 15
```

Figure 5. *Ranking of Attributes*

Conclusion and Future Works

Through analysis of data collected from previous academic years, the researchers had identified the essential subjects that contributed to student retention in the BS Architecture program. The result of the study showed that the courses related to mathematics had a substantial impact on undergraduate architecture students' academic growth, and the subject ARCH 23 had a stronger correlation with retention in BS Architecture. The results can be valuable for academic departments and educators in identifying the courses that are critical to students' success.

Future research in this area could expand the scope of the study to include factors that contribute to student retention, like non-academic factors such as socioeconomic status, students' characteristics, and other factors that may affect student retention. To gain deeper insights into the variables affecting students' academic performance and program success, qualitative research techniques such as interviews or surveys may also be used. Furthermore, future studies can explore different data mining techniques to analyze the data. Deep learning algorithms and other advanced techniques may improve the accuracy of predicting critical courses. In addition, it would be interesting to explore the potential applicability of the research findings to other courses in architecture or different academic disciplines.

References

[1] Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics, 37*, 37. https://doi.org/10.1016/j.tele.2019.01.007

[2] Alyahyan, E., & Düştegör, D. (2020). *Predicting academic success in higher education: literature review and best practices. International Journal of Educational Technology in Higher Education, 17*(1). https://doi.org/10.1186/s41239-020-0177-7

[3] Antoniades, A. C. (1992). *Architecture and allied design: An environmental design perspective* (3rd Ed.). Kendall/Hunt Pub Co.

[4] Apuke, O. D. (2017). *Quantitative research methods: A synopsis approach. Kuwait Chapter of Arabian Journal of Business and Management Review, 6*(11), 40–47. https://doi.org/10.12816/0040336

[5] Baker, R., Isotani, S., & Carvalho, A. (2011). *Mineração de dados educacionais: Oportunidades para o Brasil. Revista Brasileira de Informática Na Educação, 19*(02), 2–13. https://doi.org/10.5753/rbie.2011.19.02.03

[6] Catarina, S., Orientadora, T., & Prof, Q. (2009). *Florianópolis 2009.* 103. http://www.udesc.br/arquivos/id_submenu/102/diplomacao.pdf

[7] Domfeh, C. (2015). Reconciling mentor and residence lecturer's scores in grading students on teaching practice to be true reflection of students' performance. *Ghana Journal of Health Physical Education Recreation Sports and Dance (GJOHPERSD), 10*(1), 160–170.

[8] Donoso, S., & Schiefelbein, E. (2007). *Analisis de los modelos explicativos de retencion de estudiantes en la universidad: Una vision desde la desigualdad social. Estudios Pedagogicos, 33*(1), 7– 27. https://doi.org/10.4067/s0718-07052007000100001

[9] Dutta, B. (2022). *Six types of classifiers in machine learning*. https://www.analyticssteps.com/blogs/types-classifiers-machine-learning

[10] Frank, E., Hall, M. A., & Witten, I. H. (2017). The WEKA workbench. *Data Mining,* 553–571. https://doi.org/10.1016/b978-0- 12-804291-5.00024-6

[11] Huang, S. (2011). Predictive modeling and analysis of student academic performance in an Engineering Dynamics Course. *Thesis USA,* 136.

[12] Khodbole, P. (2020). *Eight extremely important subjects in architecture college. Rethinking the Future.* Retrieved December 2, 2020, from https://www.re-thinkingthefuture.com/architectural-community/a2277-8-extremely-important-subjects-in-architecture-college/#google_vignette

[13] Kumar, D. (2021). *Role of cross-validation data in machine learning.* https://www.linkedin.com/pulse/role-cross-validation-data-machine-learning-deepak-kumar

[14] Kurt, S. (2009). An analytic study on the traditional studio environments and the use of the constructivist studio in architectural design education. *Procedia - Social and Behavioral Sciences, 1*(1), 401–408. https://doi.org/10.1016/j.sbspro.2009.01.072

[15] Lobo, M. B. D. C. M. (2012). Panorama da evasão no ensino superior brasileiro: Aspectos gerais das causas e soluções. *Associação Brasileira de Mantenedoras de Ensino Superior Cadernos, 25*(08780 220), 1–23.

[16] Mesevage, T. G. (2020). *Machine learning classifiers - The algorithms & how they work.* MonkeyLearn. https://monkeylearn.com/blog/what-is-a-classifier/

[17] Mo, D. (2019). *WEKA: Open-source machine learning tools for developers.* https://medevel.com/WEKA-machine-learning-tools-for-java-developers/

[18] Nediger, M. (2020). *What is data visualization?* https://venngage.com/blog/data-visualization/#1

[19] Pradeep, A., & Thomas, J. (2015). Predicting college students' dropout using EDM techniques. *International Journal of Computer Applications, 123*(5), 26–34. https://doi.org/10.5120/ijca2015905328

[20] Swaen, B., & George, T. (2022). *Developing a conceptual framework for research.* https://www.scribbr.com/methodology/conceptual-framework/

[21] Thomas, L. (2020). *An introduction to quasi-experimental designs.* https://www.scribbr.com/methodology/quasi-experimental-design/

[22] Tika, F. (2021). Learning style and achievement learner madrasah ibtidaiyah. *Journal of Basic Education Research, 2*(1), 10–14. https://doi.org/10.37251/jber.v2i1.113

[23] UIA. (2017). UNESCO-UIA charter for architectural education. *International Union of Architects,* 12. https://www.uia-architectes.org/webApi/uploads/ressourcefile/178/charter2017en.pdf

[24] York, T. T., Gibson, C., & Rankin, S. (2015). Defining and measuring academic success. *Practical Assessment, Research and Evaluation, 20*(5), 1–20.

Acknowledgment

Academic Community – The authors would like to thank them for their insightful comments and conversations. The interchange of ideas with other researchers and experts broadened their comprehension of the topic.

Cavite State University – The researchers would like to express their gratitude for allowing them to conduct this research. This study would not have been possible without their cooperation. Cavite State University's cooperation had been outstanding throughout their endeavor. They truly appreciated their ongoing encouragement and belief in their abilities. The university's resources and facilities considerably benefited their research efforts.

Cavite State University Registrars - Special thanks to the University Registrars. Their contribution of essential data to this study was critical for its completion. Their willingness to provide this information was important to their investigation.

Dr. Edwin R. Arboleda - Their research adviser deserves special mention for his unwavering passion, patience, and wide expertise. His guidance and assistance were extremely beneficial in the development of this research project. With Dr. Arboleda's help, they were able to successfully write this research paper. His valuable input and feedback greatly enhanced the quality of their work. They were deeply appreciative of his dedication and commitment to their project.

Participants – The researchers would also like to thank the participants who were willing to participate in their research and contributed to the reliability of their findings. Their participation and collaboration were vital to their project's success. Also, they would like to extend their heartfelt appreciation to everyone who assisted them in any way during this research. Their words of wisdom, encouragement, and assistance were vital to them in their academic journey.