Exploring the Impact of Attendance on the Continuous Internal Evaluation Among the Students of Business School

Debraj Bhattacharjee
Assistant Professor, FLAME University, Pune Email:
debrajbhattacharjee26@gmail.com

https://doi.org/10.5281/zenodo.10598555

ABSTRACT

In the modern education system, a minimum class attendance becomes a necessary criterion for eligibility for the final semester examination in most educational institutes. This criterion of minimum attendance threshold value has no direct influence on the final grade of the students. Still, many institutes made it mandatory for students to keep a minimum attendance percentage. However, the introduction of a continuous internal evaluation process in the education system has increased the importance of class attendance as it directly influences class participation, which is one of the components of continuous internal evaluation. In this article, the impact of attendance has been assessed by applying a set of different techniques, namely, correlation analysis, regression analysis, clustering analysis, and decision tree for business school students on their continuous internal evaluation marks. Here three attendance-related variables, namely, the attendance percentage for the first half of the semester (FHA), attendance percentage for the second half of the semester (SHA), and total percentage attendance (TA) have been considered as independent variables. Tools like correlation analysis and linear regression analysis have been deployed to find the statistical association and type of statistical association between attendance-related variables and continuous evaluation marks. Further, k-means clustering has been applied and the two distinct groups are identified. These two groups or clusters are further analyzed using a decision tree to provide useful insight into the relation between marks and class attendance. The result indicates that the students who maintain good attendance at the initial half of the semester are more prone to get good continuous internal marks or grades and the students who have higher overall attendance are more consistent in academic performance.

Keywords: Business school student, class attendance, continuous internal evaluation, regression analysis, correlation analysis, clustering analysis

A. Introduction

Attending a physical class in different educational institutions has been considered as important as it provides an opportunity for the teacher or faculty members to directknowledge transfer to the students (Elmore &Gieskes, 2013). Keeping this point in mind, a major number of different academic institutions around the world make it compulsory to attend classes with a minimum attendance threshold (Cohn & Johnson, 2006). However, the research, conducted till now, provides evidence that there is no direct relationship between success in life and class attendance (Devadoss & Foltz, 1996; Berić-Stojšić et al., 2020). However the impact of class attendance on academic success is observed in different literature (Aldosary, 1995; Al Hazaa et al.,2021). However, in the case of school students, the impact of attendance the academic performance is very less (Büchele, 2021). These studies mostly dealt with the traditional evaluation process of the written exam process. In modern times, the evaluation methods have been changed drastically and components like multiple class tests, assignments, and class participation are used as measures of academic performance. These components are often called the continuous internal evaluation components, to evaluate the students by the day-to-day activities inside the class classroom. The main advantage of such a valuation process is that academic performance is not dependent on the examination day's performance. In such circumstances, attendance becomes more important as the absent students may miss some of the evaluation components. But, while talking about the high school study the findings show that class participation has a positive impact on academic performance, which contradicts the previous research as here more continuous evaluation components are used for academic performance evaluation (Durden, & Ellis,1995; Cohn, & Johnson, 2006; Silva et al., 2010). The research by Durden & Ellis, (1995) found that in the case of courses like economics, where the students need to understand and critically think about the core concepts of the subjects, higher attendance plays a crucial role in improving the average score in the exam as the students with higher attendance provide continuous exposer to the interlinked economic concepts. This type of similar impact of attendance could be found in the studies on college students, where a lot more understanding of core knowledge of the subjects and application of bookish knowledge is required. The study by Khan et al., (2003) reveals students with better lecture attendance have better final grades than the students, who have lower attendance. A quantitative measure of this effect of attendance on final exam scores can be found in the research by Dobkin et al., (2010). The authors estimate the 10% increment in the overall attendance increases by 0.17 standard deviation of the final exam score. A positive correlation between attendance and grade performance was observed in the research by Haque, (2012). The research by Bamuhair et al., (2016) indicates that the students who attend the lectures have higher grades in the college education system. Similar results are found in the studies by other researchers like Kwenda, (2011), Kwak et al., (2019), Kim et al., (2020), and Khan, (2022).

In literature, the different type of factors related to attendance is considered and these different factors have different types of importance on the student learning and development. These different types of factors are studied by the researchers to find their exact impact on the final grade or academic performance of the students. The positive impact of the lecture classes was described in the research by Paisey&Paisey,(2004) and the results of the study indicate that the students, who are regular in the lecture classes are more successful in academic activities and the indication of this academic excellence can be found in the literature (Marburger, 2001), which shows that the students with higher attendance are more responsive inside the classroom. Similar kind of findings are also found for blended learning-based pedagogy in the article by López-Pérez, et al. (2011). Another interesting finding can be observed in the research article by Lowder, et al., (2015), which concludes that attendance matters in the case of the courses, where the basic knowledge or the fundamentals of a subject is taught. According to Lowder et al., (2015), the impact of attendance is very less on the grade points as the grade points are awarded for a particular range of marks. The research conducted by Kaushik, et al., (2023), on the Batchelor of Technology (BTech) students, indicates that there should not be a long gap between the classes, which impacts the attendance of the students and further deteriorates the students' academic performance. Mainly in literature the attendance of lecture classes, tutorial classes, attendance in advanced courses, attendance in introductory courses, and the difference between two consecutive classes are considered along with other demographic factors. As this research only focuses on class attendance, attendance-related literature is more explored for the literature review. This article focuses on attendance-related factors and uses new factors like First Half Attendance (FHA) for the semester, which is nothing but the attendance till the mid-semester examination and another isSecond Half Attendance (SHA) which is the post mid-semester examination. These variables are rarely used in the available past literature and this produces an opportunity for exploring their impact on the continuous internal evaluation marks. Also, the Total percentage Attendance (TA) of the semester has been used for the study along with the previously mentioned two new variables or factors. The response variable considered here is the Percentage of Total Internal Continuous (CT) evaluation marks for business school students. All the factors are considered in percentage value for uniformity and as per the convenience of the analysis. The article mainly focuses on the business school academic activity as the courses offered to the students are highly practical and required proper knowledge transfer through proper class activity.

The organization of the article is started with an introduction to the research topic then methodology describes the collected data used for conducting the study and the tools used for data analysis in brief. The next section discusses the results and the final section concludes the article.

B. Methodology

The article deployed correlation analysis, regression analysis, cluster analysis, and decision trees for analysing the data. The correlation analysis is used to understand the association between the three attendance-related factors FHA, SHA, and TA, which may influence the response variable CT. This association information is then used to describe the insights obtained about the type of possible statistical relationship between the FHA, SHA, and TA and the response variable CT. Finally, a clustering analysis hasbeen performed to find multiple groups or clusters to know the hidden insights, which is useful to concretize the specific relationship between them. Here the decision tree method is used to find out the insight with the clusters. A schematic of the methodology of the article has been provided in Figure 1. The details of the used techniques in this methodology have been described in the consecutive subsections for a better understanding of the proposed methodology.

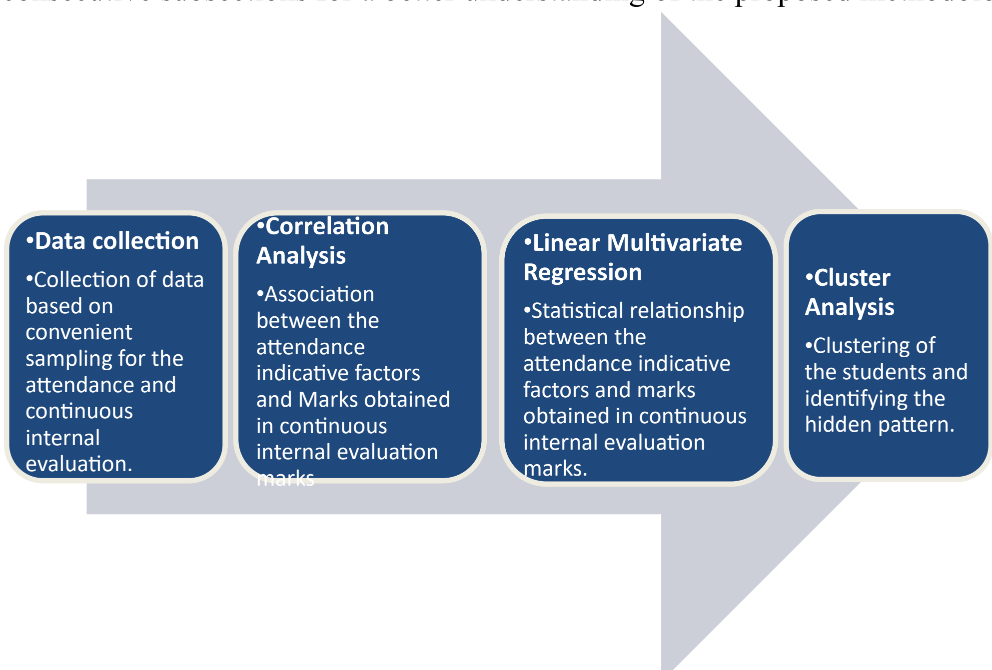

Figure 1. Methodological flow

Data Description

This research uses 448 data points from a business school for the study. The data contains information about attendance percentage for the first half (FHA), attendance percentage for the second half of the semester (SHA), Total attendance percentage (TA), and percentage of marks in continuous internal evaluation (CT). CT has several components like one class test, two quizzes, Assignments, and Class participation. The weightage of the class participation was 14.29% which is quite significant and this component is indirectly influenced by the attendance of the student. Convenient sampling has been deployed to collect the data and the data has been collected from a business school for both undergraduate and postgraduate students. Here the demographic, gender, and age information are not collected due to privacy concerns and the author tries to avoid the comparative study based on gender and age. The descriptive statistics for the data have been provided in Table 1. The marks concerning total attendance have been plotted in Figure 2. Similarly, the CT values for FHA and SHA are plotted in Figure 3 and Figure 4 respectively.

Statistic	FHA (%)	SHA (%)	TA (%)	CT (%)
No. of observations	448	448	448	448
Minimum	0.000	0.000	0.000	0.000
Maximum	100.000	100.000	97.730	99.290
Median	71.595	73.680	72.900	79.290
Mean	66.998	63.297	68.195	74.099
Standard deviation (n-1)	21.927	29.172	19.073	15.381

Table 1. Descriptive statistics for the collected data

Table 1 shows the mean of all three types of attendance is more than 60% and the median is more than 70%. Another fact regarding the continuous internal evaluation is the higher mean and median value, which is more than 70%. The standard deviation indicates 19.073% total attendance change, and the CT change is 15.38%. The three scatter plots show that in Figure 2 and Figure 3, the students with attendance of around 80% have very good scores. But in the case of Figure 4 the students with good scores are distributed over the different range of attendance percentage values.

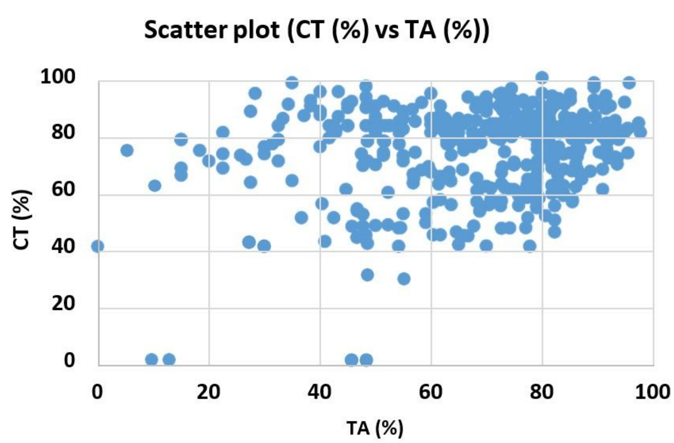

Figure 2. Scatterplot for TA (%) and CT (%)

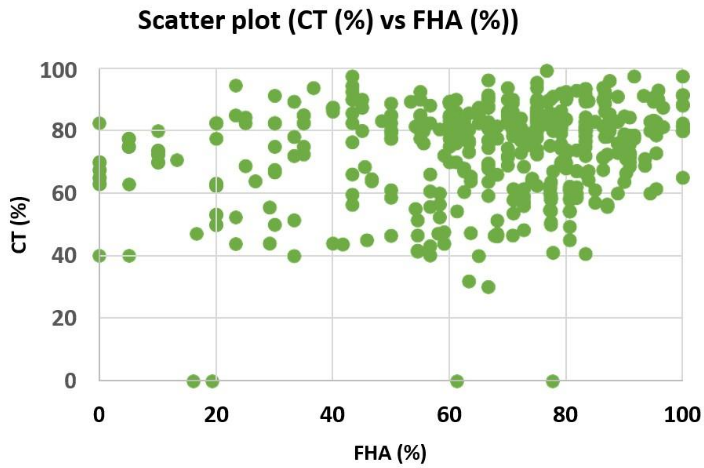

Figure 3. Scatterplot FHA (%) and CT (%)

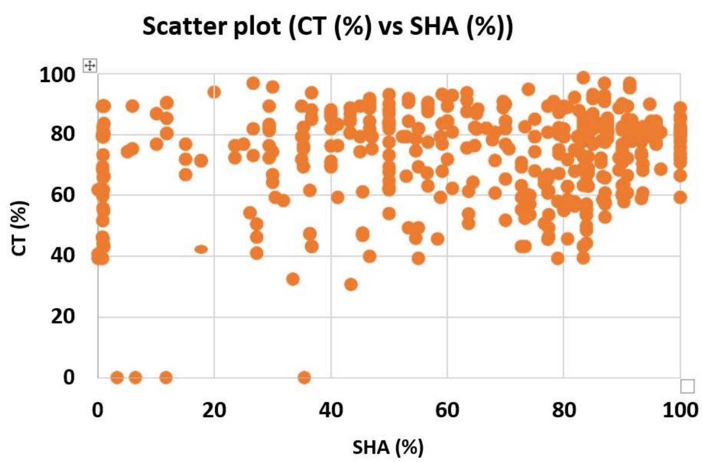

Figure 4. Scatterplot SHA (%) and CT (%)

The regression analysis is a method where a relationship between the independent variables and dependent variables are established as per the collected data to describe a phenomenon, which not necessary a causality. In this regression analysis, the data has been randomly divided into two parts one is training and another is testing part. This approach for using two different datasets for building the model and testing the model is required for testing the effectivity of the model in case of unknown set of information. The Training data statistical description has been provided in Table 2. The sample size for the training data is 326, where no missing data is observed. Here all the factors and response variable are summarized with the information like minimum value, maximum value, mean and standard deviation value. On the other side, similar information about the test data set has been provided in Table 3, which have the sample size of 122.

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
CT (%)	326	0.000	99.290	73.589	15.390
FHA (%)	326	0.000	100.000	66.609	22.193
SHA (%)	326	0.000	100.000	63.633	28.835
TA (%)	326	0.000	97.730	68.148	18.964

Table 2. The summary of training d ıta

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
CT (%)	122	0.000	97.500	75.461	15.338
FHA (%)	122	0.000	100.000	68.037	21.255
SHA (%)	122	0.000	100.000	62.398	30.155
TA (%)	122	9.680	95.700	68.320	19.439

Table 3. The summary of testing data

Correlation Analysis

Correlation analysis is a method to know the type of association between two variables. Here in this article, Pearson correlation analysis has been deployed to find the correlation coefficient value between the factors and response variable (Pearson, 1896). The Pearson correlation coefficient (r_p) calculation expression has been provided in equation 1, where x_i, y_i are two variables. The insight about the data needs to be known to use for better understanding of association between the independent variable and dependent variable. This has been provided in Table 4.

$$r_p = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2\sum_{i=1}^{N}(y_i-\bar{y})^2}} \quad (1)$$

	FHA (%)	SHA (%)	TA (%)
CT (%)	0.218*	0.138	0.194

**Indicates significance in 95% confidence level*

Table 4. Correlation analysis summary

Regression Analysis

Regression analysis is a process for establishing the statistical relationship between dependent variable and independent variables **(**Pearson, 1896)**.** Here the most common type of regression analysis, which is linear multivariate type, has been utilised for the understanding how the continuous internal evaluation mark is influenced by the attendance. The use of regression analysis for estimating or predicting the grade can be found in the research by Khan, & Al Zubaidy, (2017), Ayyappan, (2019), Alsariera et al., (2022) and Virtanen, &Tynjälä, (2022). The linear multivariate regression model is expressed in a form of mathematical expression provided in equation 1, where y ^ is estimated value of dependent variable, $c_0$is intersection and $\beta_i$indicates regression coefficient and $x_i$indicates the independent factors.

$$\hat{y} = c_0 + \sum_{i=1}^{N} \beta_i x_i \qquad (2)$$

The main assumption is that the independent and dependent variables are following normal distribution. The summary of the regression model has been provided in Table 5, where the information about the regression coefficients, intercept and p-values are tabulated. The regression model has been first derived based on the training data, which is summarized in Table 2. Further the described regression model in Table 5 has been tested in using the testing data mentioned in Table 3. The performance measurement of the derived regression model has been provided in the form of summary for goodness of fit in Table 6. The detailed insights form the Tables are described in the discussion section.

Source	Value	Standard error	t	Pr> |t|
Intercept	62.348	3.119	19.989	**<0.0001**
FHA(%)	0.157	0.068	2.315	**0.021***
SHA(%)	0.055	0.040	1.355	0.176
TA(%)	-0.040	0.097	-0.412	0.681

**Indicates significance in 95% confidence level*

Table 5. Regression model summary

Statistic	Training set	Validation set
Observations	326	122
DF	322	118
R^2	0.054	0.71

Table 6. Goodness of fit statistics (*CT (%))

Cluster Analysis

Cluster analysis is a process of grouping the data based on the similarity of each sample. The cluster analysis is an unsupervised type of learning and it helps the authors to find the similarities among the data, which reveals the hidden pattern inside the data (Sinaga & Yang, 2020). These hidden patterns are further analyzed and used to describe the possible conditional relationship among the response variable and the factors (Aggarwal, & Sharma, 2019). The K- means algorithm has been utilized for conducting the clustering analysis. K means algorithm is a simple clustering method based on the Euclidian distance, which helps to identify the nearest data points of the centroid values (Chang et al., 2020). The Euclidian distance measurement expression has been provided in Equation 3. The clustering of K means the algorithm is simple but for that the K or number of clusters should be predefined. Here two methods are used for identifying the optimal number of clusters one is within cluster inertia and another is the Silhouette scores calculation. The within-cluster inertia values are plotted with respect to the number of clusters in Figure 1. The Silhouette scores for a specific number of clusters are provided in Table 7. It can be observed that for two cluster values the within-cluster inertia and Silhouette scores are highest. As a result of this, the optimal cluster number is selected as two. The initial centroids and final centroids are provided in Table 8. The summary of the two clusters is provided in Table 9.

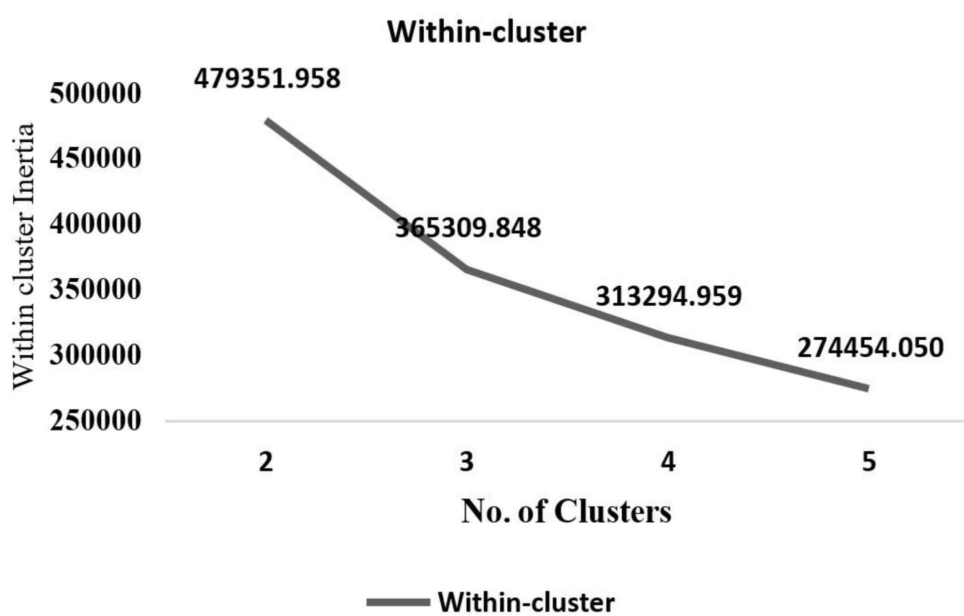

Figure 5. Within cluster inertia

Cluster	2	3	4	5
Silhouette scores	0.428	0.421	0.344	0.29 3

Table 7. Silhouette scores with respect to the number of clusters

In Table 8, The initial and final centroid information are provided for the understanding of the centroid change during the process of K means clustering. After clustering the standard deviation and coefficient of variance are calculated for both clusters for understanding the uncertainty and consistency within the clusters. The calculation formula for the standard deviation (σ) has been provided in Equation 3 and the coefficient of variance (P) has been calculated based on Equation 4. Here higher standard deviation indicates higher uncertainty and a lower coefficient of variance indicates better consistency within the data.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}} \quad (3)$$

Where x is the data point and N is number of observations.

$$\mathsf{P} = \frac{\sigma}{\mu} \quad (4)$$

Where μindicates the mean and $\sigma$is standard deviation.

Cluster	Centroid	FHA%	SHA (%)	TA (%)	CT (%)	Sum of Weights	Within-clustervariance
1	Initial	66.507	64.618	68.320	73.720		
	Final	55.845	32.841	51.722	72.030	180.000	1758.394
	σ	25.302	21.785	17.836	18.827		
	P	0.458	0.644	0.345	0.263		
2	Initial	67.327	61.937	68.060	74.505		
	Final	74.329	83.800	79.258	75.488	268.000	616.477
	σ	15.114	10.943	8.689	12.191		
	P	0.201	0.130	0.109	0.161		

σ: Standard deviation; P*: Coefficient of variance*

Table 8. Initial and final centroid of the clusters

Table 9 indicates the number of members within cluster 1 is 180 and in cluster 2, it is268. The within cluster variance indicates the span of the cluster and calculated average distance to centroid indicated how far the members are situated inside the cluster. Less value for both clusters is desirable for good cluster performance. Hence, cluster two performs better in comparison to the cluster 1.

Cluster	1	2
Number of objects by cluster	180	268
Sum of weights	180	268
Within-cluster variance	1758.394	616.477
Minimum distance to centroid	6.914	4.055
Average distance to centroid	38.553	22.176
Maximum distance to centroid	97.008	82.198

Table 9. Cluster details

The two clusters are further analysed using the decision tree. The grades for the continuous evaluation are considered as the target variable and the levels are defined based on the range of percentage and are provided in Table 10. According to these levels the decision trees are created and presented in Figure 6 and Figure 7. Here the focus is on the grades O, A+ and A. So, for these grades the rules are analysed separately and the attendance values are observed for useful insights.

Sl.No.	Grade	CT % range
1	O	$90\% \leq CT$
2	A+	$85\% \leq CT < 90\%$
3	A	$80\% \leq CT < 85\%$
4	B	$70\% \leq CT < 80\%$
5	C	$60\% \leq CT < 70\%$
6	D	$50\% \leq CT < 60\%$
7	P	$40\% \leq CT < 50\%$
8	F	$CT < 40\%$

Table 10. Grade definition with percentage marks

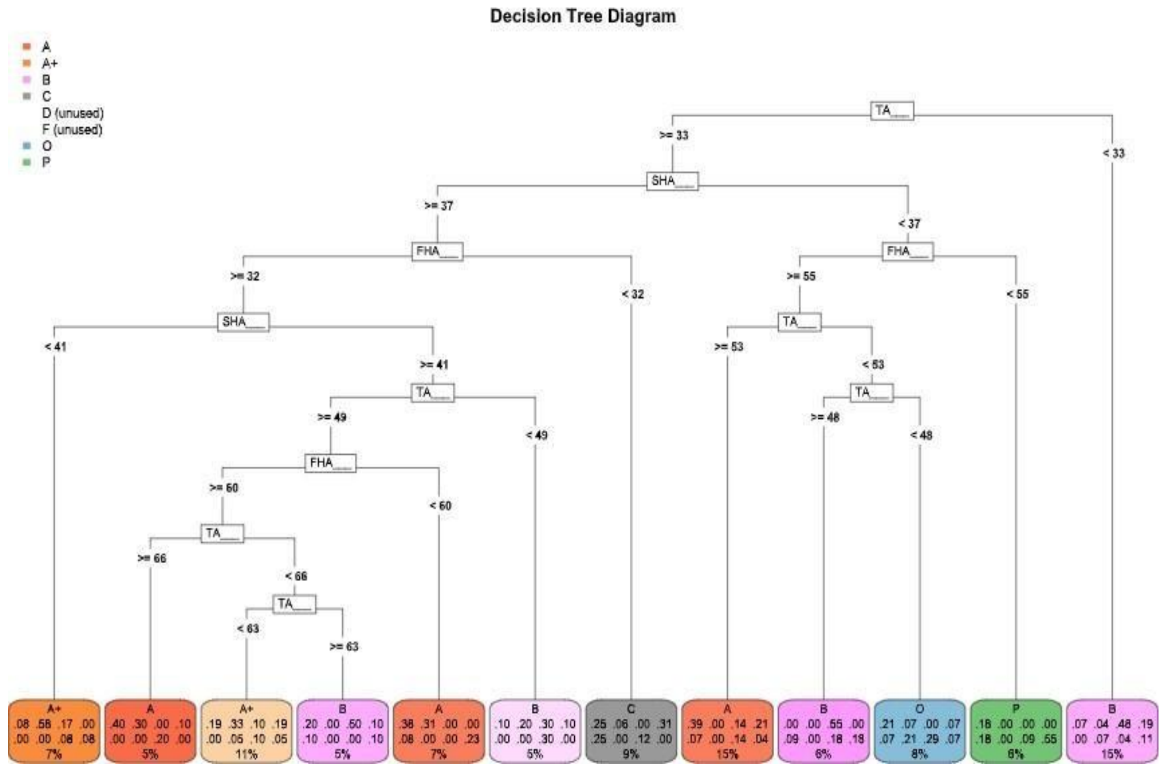

Figure 6. Rule visualisation for Cluster 1

Cluster 1 indicates that the student with total attendance from 60% to less than 63% achieved the score A+. The added condition is that they must maintain 49% attendance before mid-semester examination (refer Figure 6). On the other hand, if the total attendance is more than 53% and the attendance before mid-semester examination is more than equal to 55% then the chance of getting grade score A is higher. Total attendance of less than 48%, 55% attendance before midsemester examination and after mid-exam, the attendance of less than 37%, produce a higher chance of getting O.

In the Case of cluster 2, if the attendance percentage before mid-semester examination is less than 77% and after midsemester exam, the attendance is less than 71%, produce higher chance of getting a grade score of A+ (refer Figure 7). Similarly, if the attendance of the student before mid-semester exam is less than 89% and the attendance post mid semester exam is greater or equal to 87% then there is a high chance of achieving a grade score of A. In this cluster for getting O grade the total attendance should be less than 76% but the attendance before midsemester exam should be greater than or equal to 78% and the attendance in the post mid semester exam should be less than 87%.

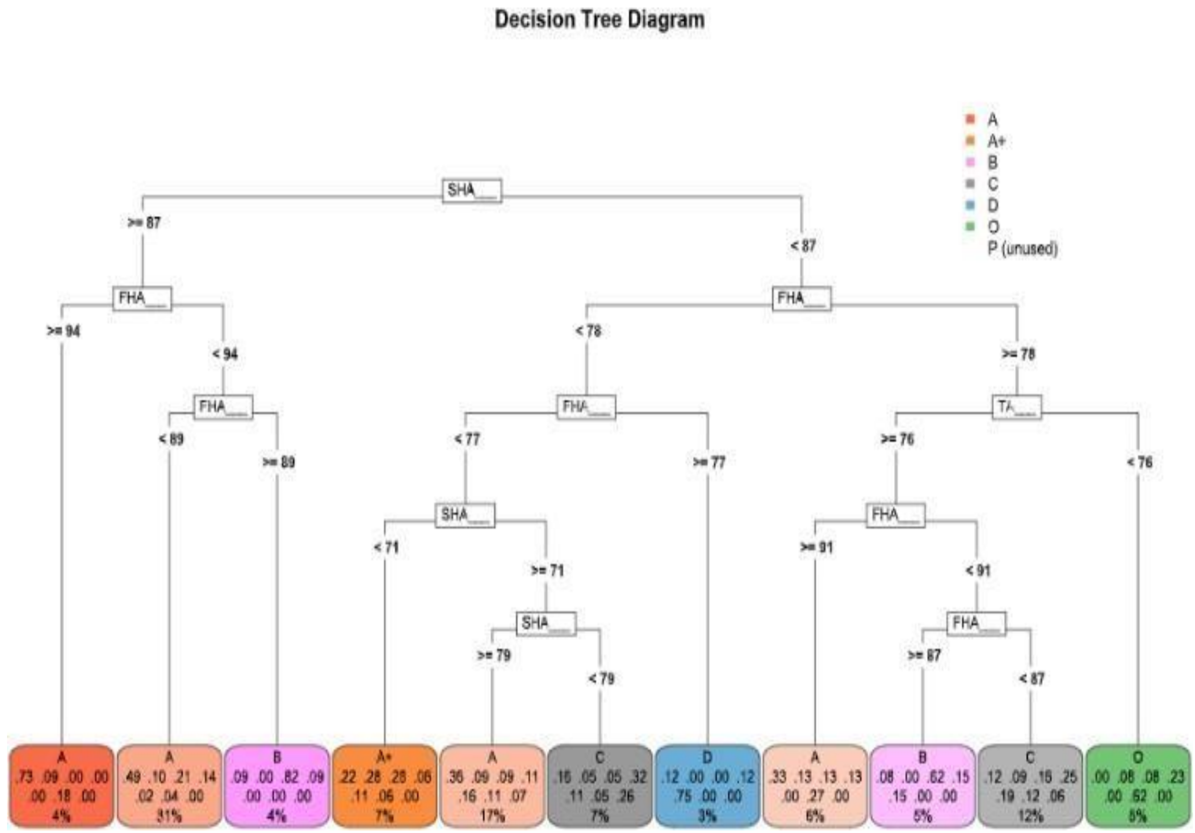

Figure 7. Rule visualization for cluster 2

In this article for the two clusters very few rules are generated for grades C and D. For example, only cluster 2 is the rule, which indicates the attendance in the first half of the semester should be more or equal to 77% and less than 78% can produce a score of D grade with additional condition is that the attendance in the second half of the semester should be less than 87%. For, C grade cluster 1 indicates the total attendance should be greater or equal to 33% and the attendance in the first half of the semester should be less than 32%. In cluster 2, similar grade two rules are there, here only the rule with a higher percentage of probability has been provided. The rule indicates that a student has a higher chance of getting a grade score of C if the total attendance is greater or equal to 76% and attendance in the first half of the semester is between 78% and less than 87%. The rule also indicates that attendance in the second half of the semester should be less than 78%. In the case of grade B also the rules having higher chances have been mentioned here. In cluster 1, the total attendance should be less or equal to 33%. On the other side, cluster 2 indicates if the total attendance is greater or equal to 76% and the attendance in the first half of the semester is within 87% to 91% along with less than 87% attendance in the second half of the semester provide a higher chance of getting a grade score of B.

C. Discussion

In the data analysis, it can be found that the attendance components have very little correlation between the attendance indicative factors, namely, FHA, SHA, TA, and the response variable CT. Here one important finding is the correlation between FHA and CT is significant (refer to Table 4). This indicates that attending the classes before the mid-semester examination has some statistical association with the result but the association strength is very low. Further for more insight can be gained from the regression analysis. Table 5 clearly indicates that there is a positive impact of attendance on the marks obtained in the continuous evaluation process. This positive impact is significant. The main reason for such observation is the students, who are attending the classes are more familiar with the topics that are taught in the class and because of that, they have a higher chance of scoring good marks in the mid-semester exam, which have a significant proportion in the continuous internal evaluation. Also, most of the assignments and class activities tend to happen in the first half of the semester. Also, the regression analysis indicates that in the case of training data and testing data the three attendance-indicating factors able to explain the variance in the CT only of 5.4% and 7.1% respectively (Refer Table 6). This also indicates other factors may be responsible for the larger amount of variance which is not explained by the three factors, namely FHA, SHA, and TA. So, this analysis failed to give any generic model for estimating the grade through attendance based on the data. As a result of that the decision tree is deployed to extract the attendance-based rules for getting good grades during the examination.

Based on the correlation and regression analysis it is very hard to estimate the grade of the students using the attendance information as there may be some cases where the individual students' merit plays a major role in getting a good grade. So, in this sample may be different types of groups, of different natures, are there and identification of them is necessary to conclude. The K means algorithm has been deployed in this article and two distinct groups are identified in this process. These two groups are very different in nature and that can be observed in the group information provided in Table 8 and Table 9. The cluster 1 is have average attendance significantly less than the cluster 2 or group 2. But the marks-wise the cluster 1 have an average of 72.03% and cluster 2 has average of 75.488%. If the standard deviation and consistency are checked then cluster 2 or group 2 is better than cluster 1. This concludes that the students with good attendance are more consistent than the students, who have less class attendance. This indicates that the students who have good class attendance have a higher chance of getting good grades in the continuous evaluation process. Finally, the decision tree indicates some remarkable rules for obtaining good grades (O, A+, A). The first cluster indicates (refer Figure 6) though the students have less than 48% total attendance still it can give an O grade if the attendance before the mid- semester examination is more or equal to 55%. So, just coming to the class in the first half of the semester can give a student a good grade. Similar types of rules can be observed for the A+ grade. The students should maintain total attendance with 60% to 63% and must maintain an attendance of a minimum 49% in the first half of the semester. Achieving an A grade is easier if the total attendance is maintained at 53% and the attendance should be 55% till the mid-semester examination. The second cluster (refer Figure 7), however, indicates that good grade can be achieved through maintaining good attendance. O grade can be achieved with total attendance should less than 76% but attendance before the midsemester exam greater than or equal to 78% and attendance in the post mid semester exam should be less than 87%. If the attendance percentage before the mid-semester examination is less than 77% and after the midsemester exam the attendance is less than 71%, it gives a higher chance of getting an A+ grade. Similarly, if the attendance of the student before mid-semester exam is less than 89% and the attendance post mid-semester exam is greater or equal to 87% then there is a high chance of getting an A grade.

D. Conclusion

The study concludes that the attendance components are very less statistically associated with marks. The regression analysis indicates that the attendance before the mid-semester examination has significant impact on the internal continuous evaluation. The cluster analysis indicates that the students with less attendance are less consistent in case of obtained marks in case of an internal continuous evaluation process in comparison to the students, who have higher attendance in the class. This makes the students with higher attendance more predicTable about good academic performance. Finally, the cluster analysis indicates it is very hard to predict the grades of the students as the obtained rules show some remarkable pattern which indicates a student can achieve good grades if the student maintains attendance of around 50%. The two clusters indicate the students who participate more in the classes before the mid-semester exam have higher chances of getting a good grade.

References

Aggarwal, D., & Sharma, D. (2019). Application of clustering for student result analysis. *Int J Recent Technol Eng*, *7*(6), 50-53.

Al Hazaa, K., Abdel-Salam, A. S. G., Ismail, R., Johnson, C., Al-Tameemi, R. A. N., Romanowski, M. H., ... &Elatawneh, A. (2021). The effects of attendance and high school GPA on student performance in first-year undergraduate courses. *Cogent Education, 8*(1), 1956857.

Aldosary, A. S. (1995). The correlation between final grade score, attendance, and homework in the performance of CED students. *European Journal of Engineering Education, 20*(4), 481- 486.

Alsariera, Y. A., Baashar, Y., Alkawsi, G., Mustafa, A., Alkahtani, A. A., & Ali, N. A. (2022). Assessment and evaluation of different machine learning algorithms for predicting student performance. Computational Intelligence and Neuroscience, 2022.

Ayyappan, G. (2019). Ensemble classifications for student academic performance data set.*Indian J. Comput. Sci. Eng, 10*, 31-34.

Bamuhair, S. S., Al Farhan, A. I., Althubaiti, A., ur Rahman, S., & Al-Kadri, H. M. (2016). Class attendance and cardiology examination performance: a study in problembased medical curriculum. International journal of general medicine, 1-5.

Berić-Stojšić, B., Patel, N., Blake, J., & Johnson, D. (2020). Flipped classroom teaching and learning pedagogy in the program planning, implementation, and evaluation graduate course: Students' experiences. *Pedagogy in Health Promotion, 6*(3), 222-228.

Büchele, S. (2021). Evaluating the link between attendance and performance in higher education: the role of classroom engagement dimensions. *Assessment & Evaluation in Higher Education, 46*(1), 132-150.

Chang, W., Ji, X., Liu, Y., Xiao, Y., Chen, B., Liu, H., & Zhou, S. (2020). Analysis of university students' behavior based on a fusion K-means clustering algorithm. *Applied Sciences, 10*(18), 6566.

Cohn, E., & Johnson, E. (2006). Class attendance and performance in principles of economic. *Education economics, 14*(2), 211-233.

Devadoss, S., & Foltz, J. (1996). Evaluation of factors influencing student class attendance and performance. *American Journal of Agricultural Economics, 78*(3), 499507.

Dobkin, C., Gil, R., & Marion, J. (2010). Skipping class in college and exam performance: Evidence from a regression discontinuity classroom experiment. *Economics of Education Review, 29*(4), 566-575.

Durden, G. C., & Ellis, L. V. (1995). The effects of attendance on student learning in principles of economics. *The American Economic Review, 85*(2), 343-346.

Elmore, M., &Gieskes, K. E. (2013, June). Attendance in large engineering classes and its effect on student performance. In 2013 ASEE Annual Conference & Exposition (pp. 23-231).

Haque, M. E. (2012, June). Effect of Class Absenteeism on Grade Performance: A Probabilistic Neural Net (PNN)-Based GA-Trained Model. In 2012 ASEE Annual Conference & Exposition (pp. 25-498).

Kaushik, A., Kumar, N., & Kumar, P. (2023). Effect of partial compulsory attendance on academic performance of technical education students. *Innovations in Education and Teaching International, 60*(1), 124-133.

Khan, H. U., Khattak, A. M., Mahsud, I. U., Munir, A., Ali, S., Khan, M. H., ... & Shah, S. H. (2003). Impact of class attendance upon examination results of students in basic medical sciences. *Journal of Ayub Medical College Abbottabad, 15*(2).

Khan, R. N. (2022). Attendance Matters: Student Performance and Attitudes. *International Journal of Innovation in Science and Mathematics Education, 30*(4).

Khan, W. Z., & Al Zubaidy, S. (2017). Prediction of Student Performance in Academic and Military Learning Environment: Use of Multiple Linear Regression Predictive Model and Hypothesis Testing. *International Journal of Higher Education, 6*(4), 152160.

Kim, A. S., Shakory, S., Azad, A., Popovic, C., & Park, L. (2020). Understanding the impact of attendance and participation on academic achievement. *Scholarship of Teaching and Learning in Psychology, 6*(4), 272.

Kwak, D. W., Sherwood, C., & Tang, K. K. (2019). Class attendance and learning outcome.*Empirical Economics, 57*, 177-203.

Kwenda, M. (2011). Factors affecting performance in an introductory sociology course.*College Teaching, 59*(2), 60-65.

López-Pérez, M. V., Pérez-López, M. C., & Rodríguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers & Education, 56*(3), 818-826.

Lowder, L., Khalid, A., Ferreira, D. R., Bohannon, J. L., Stutzmann, B., Atiqullah, M. M., ... &Colebeck, D. (2015). Student and faculty perceptions of attendance policies at a polytechnic university. In ASEE Annual Conference & Exposition, Seattle, Washington. American Society for Engineering Education.

Marburger, D. R. (2001). Absenteeism and undergraduate exam performance. *The Journal of Economic Education, 32*(2), 99-109.

Paisey, C., &Paisey, N. J. (2004). Student attendance in an accounting module–reasons for non-attendance and the effect on academic performance at a Scottish University. Accounting education, 13(sup1), 39-53.

Pearson, K. (1896). VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character, (187), 253-318.

Silva, E. T. D., Nunes, M. D. F., Queiroz, M. G., &Leles, C. R. (2010). Factors influencing students' performance in a Brazilian dental school. *Brazilian Dental Journal, 21*, 80-86.

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. IEEE access, 8, 80716-80727.

Virtanen, A., &Tynjälä, P. (2022). Pedagogical practices predicting perceived learning of social skills among university students. *International Journal of Educational Research, 111*, 101895.