Exploring Academic Dishonesty Among University Students: University of Jammu, India

Sunil Kumar, Apurba Vishal Dabgotra

Department of Statistics, University of Jammu, Jammu, J&K, India

Corresponding author: Sunil Kumar, Email: sunilbhougal06@gmail.com

Over the years, academic cheating continues to be an endemic issue that has always been a threat to academic honesty and social values. Academic dishonesty among the students is a perplexing phenomenon, that exists in every stage of our education system, especially in colleges and universities. College and university administrators admit that academic dishonesty is an issue on campus but they often lack in preparing effective policies and procedures to monitor it and to deal with it. In addition, indecisive perceptions regarding academic dishonesty has adverse effects on paradoxical situation of education. The current study provides the details about the causes that motivate the students to cheat and describes the different forms of cheating practices performed by the students. The purpose of this study is to define the students of University of Jammu what is meant by academic cheating, to determine the factors associated with cheating behaviour and to classify of the students according to their cheating behaviour. From the current study, the main reasons due to which students go for cheating are not knowing/understanding the study material, performance pressure, inadequate exam preparations, etc. Also, around 42%, 6% and 52% of the respondents are found to be occasional, persistent and instantaneous cheaters, respectively. We suggest that students must understand that cheating is wrong not only for the society but also for their own knowledge because by indulging in cheating students prevent them from learning what they are studying and hence, deteriorate the intellectual human resource of the country.

Keywords:Academic cheating, Latent Class Analysis (LCA), AIC, BIC.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 629–647. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-60

1 Introduction

Academic dishonesty is an escalating phenomenon that is plaguing educational institutions around the world. Infact, it has become an inescapable activity, especially, in colleges and universities, where grades earned, directly effect the academic careers of students for many years to come. The rising pressure to get the best grades in school, get into the best college, and land the best paying job is a cycle that has made academic dishonesty increase exponentially. The subject of academic cheating has attracted the attention of not only academics but also public communities. Recent studies have proven that the issues of academic cheating among undergraduates have increased along the years. For example, 76 percent of the students confessed to having involved in academic cheating [16].

Rajendran defined cheating as an activity which is performed to complete a work in an unethical way by a person when he does not know how to do that work in a legal way ([27]). However, this is not the only way of defining cheating; different people have their different perception about cheating like [11] defined cheating as changing one or more answers when correcting own examination and/or not marking two or more in correct responses, [23] defined cheating as an immoral activity in the academic environment, [28] defined cheating as an act of using any means of unfair and unjust privileges that include: lying, concealing the truth, deceive, deceit and violation of trust to achieve something. But academic cheating/academic dishonesty may be defined as the students' behaviour of contradicting the fundamental values of their academic lives [11]. Academic cheating can be of various types, suggested by [11] and [19], such as copying from another's person, stealing examination papers and lecture notes, using prohibited material like crib notes, deception, sabotaging, impersonation, forgery, plagiarism, fabrication, data manipulation, padding bibliographies, and many more.

Various theoretical perspectives have been applied for understanding cheating. [25] examined the applicability of several theories of divergence to cheating, including deterrence theory, rational choice theory, social bond theory, etc. Thereafter, [4] proposed the integrated model of cheating, which shows that students' intentions to engage in disfunctional behaviours may be influenced by attitudes, subjective norms, perceived behavioural control and moral obligations. Some studies focused on the relationship between motivation and cheating behavior, some focused more on the relationship between socio-demographic factors and cheating behavior (e.g., [15]). Others (e.g., [8]) have carried out studies to determine what causes students to cheat. The existence of academic cheating has always been a major concern for various researchers. Many studies have

been conducted in order to identify various types of academics cheating among college students and this can be achieved by using anonymous questionnaire distributed or mailed to the students. The estimates of students that are involved in academic cheating during their college lives, ranges from 49% for marketing students [32] to 88% for premedical students [30]. Dishonesty in an academic setting has been a consistent and paramount problem for many years at all educational levels [9], and it is a serious educational issue([26]; [18]). Considerable progress has thus, been made in identifying factors that influence cheating behaviour. [3] have found that age and self esteem are the two major factors that are positively correlated with the cheating behaviour of students. [17] found that higher socio-economic status along with stress, depression and family crisis are the major factors that promote students to cheat in their academic lives.

Various guidelines for controlling cheating have also been proposed like [12], [13] and [1] suggested that cheating can be controlled by making it as difficult as possible for students to cheat, [14] suggested that the cheating can be avoided by discussing the consequences of cheating with the students, etc. Despite these findings and recommendations, the prevalence of cheating is on the rise ([6]; [5]). [22] suggest that academic integrity should be strongly assumed as an institutional concern, instead of just students' responsibility. Moreover, through a collaboration approach and using workshops and open educational resources settled to address paraphrasing, summarizing and quotation, [22] concluded that "better collaboration and co-operation among faculty staff, learning advisors and librarians is therefore essential" [22]. Still there are many questions that remain to be answered concerning the nature, cause and type of academic cheating.

The purpose of current study is to define the students what is meant by academic cheating, to identify the factors/circumstances associated with cheating behaviour of students and to classify of the students according to their cheating behaviour. The paper is structured as follows: The next section (i.e. section 2) explains the sample design in detail followed by data description in section 3. Section 4 explains the framework of LCA methodology along with the statistical hypothesis and results of the analysis in the section 5. Finally, we discuss conclusions of the study in Section 6.

2 Sample Design

Our target population comprised of all the undergraduates students enrolled in 3rd semester and above, BEd students, masters students of University of Jammu and its affiliated Colleges. So, the total population size in our study was 66091. We

had followed the NSSO methodology to select the sample with 95% confidence level, 2% margin of error and 50% of population share, using formulae:

$$n = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \frac{z^2 \times p(1-p)}{e^2 N}} \tag{60.1}$$

where, n= sample size, z= z score, p= population proportion, e= margin of error, and N= size of population size.

Thus, a sample of 2317 had selected for the collection of data. And we had selected sample based on the multi-stage stratified random sampling design. The first stage units (FSU) comprises of the main campus, off campuses and different colleges affiliated to the Jammu University. The ultimate stage units (USU) comprises of different courses offered by these institutes. Then, we apply proportional allocation for selecting sufficient number of sample units from each of the USU. For the sampling frame, we had collected information from the department of Statistical Planning and Research Unit of University of Jammu.

Stratum had formed at district level. Within each district of Jammu division, nine basic strata were formed (one for each district). However, within the Jammu district, different government degree colleges, government engineering college, private engineering colleges, private BEd colleges and private law colleges formed a separate basic stratum and the remaining off campuses of University of Jammu i.e., Ramnagar campus and Bhaderwah campus and campus of University of Jammu, itself was considered as another basic stratum. Different sub-stratification were done for the courses offered by each of the selected campuses and colleges. Seven (7) different streams were considered as the courses offered by these institute. They were: Science, Technology, Commerce, Arts, Management, Law and BEd. Within each stream(course) offered by University of Jammu, its Affiliated colleges and its different off campuses, the respective sample size was allocated to the different strata in proportion to the number of students enrolled in that stream.

The above discussed technique has employed on each of the district except for Jammu District. For Jammu District, 75% of the remaining sample, after selected from the other districts and off campuses, had selected from the campus of the University of Jammu, using proportional allocation. And the rest of sample had selected from the remaining colleges of Jammu district using proportional allocation. Finally, For the different districts (from each sub stratum) required number of sample were selected by simple random sampling without replacement (SRSWOR) procedure.

3 Data Description

A descriptive design is used to study cheating behaviour of the students. So, a questionnaire, comprises of 46 closed ended questions related to factors leading to students' cheating, types of cheating and other related questions, was designed to analyze the cheating behaviour of the students. The survey period of this study was from March to September 2019. The required information was collected from a selected number of students by visiting their respective departments, colleges or institutes, personally and by asking them to fill in an anonymous questionnaire based on different cheating behaviour. But due the disturbances in the state during the survey period, it was not possible to collect the data from the highly disturbed areas such as Doda, Kishtwar and Rajouri. So, the data was collected from only 1906 students which was the relevant sample size for our study. The whole questionnaire is divided into following 5 major sections.

Respondent's Details : This section tracks the record of the demographic characteristics of the students which includes gender, age, religion, status and stream.

General Questions : This section consist of general questions about cheating in order to get the general perception of students on academics cheating and to get idea from the students about the main reason why students cheat during the exams?

Situational Aspects : This section consist of questions about the situational factors that can influence the students to cheat during the exams.

Personal Aspects : This section consist of the questions about the personal factors that can influence the students to cheat during the exams.

Types Of Cheating : This section comprises of the questions about the types of cheating students are involved in. This section was prepared with the objective of making students aware of the types of academics cheating and to know about their own cheating behaviour.

4 Methodology

4.1 Latent Class Analysis

We have certain phenomenon that often cannot be directly observed or to analyze certain phenomenon not all variables can be measured directly. So, latent variable modeling, in which the value of the latent variable (unknown variable) cannot be directly measured, rather its value is deduced from observed (manifest) variables, can be used in those cases. Latent variables may be defined as an unobserved random variables which are hidden from us [31] and are unknown to us in any particular study, whereas, manifest variables may defined as variables which are the observable and are designed specially to measure the unknown

latent variable. Indicator variables are the manifest variables which can measure the unknown latent variable. The value of unknown latent variable can be estimated on the basis of the responses made by the individuals to the different indicator variables. We also have a third variable, referred to as a grouping variable such as gender (G), which is used to identify an individual's membership in two or more population subgroups. The grouping variable is based largely on subjective criteria, by considering whether the assumptions are plausible for a it or not.

Latent Class Analysis (LCA) is appropriate when the latent variable and all the indicator variables are discrete in nature. It establishes a relationship between a set of observed discrete variables (manifest variables) and a set of unknown discrete variables (latent variables). LCA is a methodology that allows us to identify hidden population subgroups/classes. A class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values. In this approach we envisage a relationship between discrete indicator variables and discrete latent variables across different number of groups. We have used poLCA [21] package of R software for performing the Latent class analysis (LCA).

4.1.1 Latent Class Models

LCA models comprises of two types of probabilities which include the probability indicating the likelihood of a response by respondents in each of the classes and the probability representing the latent class size or the proportion of individuals who are members of a particular latent class.

Following the notation used by [21], suppose we have J polytomous categorical manifest variables (the observed variable) each of which contain K_j possible outcomes, for individuals $i = 1, 2, 3, ..., N$. Let Y_{ijk} be the observed values of the J manifest variables such that

$$\left\{ \begin{array}{ll} Y_{ijk} = 1: & \text{if } i^{th} \text{ respondent give the } k^{th} \text{ response to the } j^{th} \text{ variable} \\ Y_{ijk} = 0: & \text{otherwise} \end{array} \right\}$$

where j=1,2,...,J and k=1,2,...,K_j.

The LC models approximates the observed joint distribution of the manifest variables as the weighted sum of a finite number, R, of constituent cross-classification tables. Let π_{jrk} denote the cross-conditional probability that an observation in

class r=1,2,...,R produces the k^{th} outcome on the j^{th} variable with $\sum_{k=1}^{K_j} \pi_{jrk} = 1$.

Let p_r be the prior probabilities of latent class membership, as they represent the unconditional probability that an individual will belong to each class before taking into account the responses Y_{ijk} provided on the manifest variables. The probability that an individual i in class r produces a particular set of J outcomes on the manifest variables, assuming conditional independence of the outcomes Y given class membership, is the product

$$f(Y_i; \pi_r) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{60.2}$$

The probability density function across all classes is the weighted sum

$$f(Y_i|\pi, p) = \sum_{r=1}^{R} f(Y_i; \pi_r) = \sum_{r=1}^{R} P_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (pi_{jrk})^{Y_{ijk}}, \tag{60.3}$$

The parameters P_r and π_{jrk} are estimated by the latent class model.

Given estimates $\hat{P}_r$ and $\hat{\pi}_{jrk}$ of P_r and π_{jrk} respectively, the posterior probability that each individual belongs to each class, conditional on the observed values of the manifest variables, are calculated by

$$\hat{P}(r_i|Y_i) = \frac{\hat{p}_r f(Y_i; \hat{\pi}_r)}{\sum_{q=1}^{R} \hat{p}_q f(Y_i; \hat{\pi}_q)'} \tag{60.4}$$

where $r_i \in (1, 2, ..., R)$. It is important that the condition $R\sum_j (K_j - 1) + (R-1) \leq n$ on the number of parameters should hold. Also, $R\sum_j (K_j - 1) + (R-1) \leq (3^{10} - 1)$ i.e. one fewer than the total number of cells in the cross-classification table of the manifest variables, as then the latent class model will be unidentified. Under the assumptions of multinomial distribution, the log likelihood function can be given as:

$$lnL = \sum_{i=1}^{n} ln \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{60.5}$$

LCA not only builds a classification model but it also explain a relation of the class membership to explanatory variables by including covariates [33] in the model. Grouping variables can be used in LC models in order to model the unexplained heterogeneity in the data. In that case latent class membership probabilities are predicted by covariates through a logistic link.

4.1.2 Parameter Estimation and Model Selection

The unknown parameters of the LC models can be estimated by maximizing (60.5) with respect to p_r and π_{jrk},using the expectation-maximization (EM) algorithm ([7], [24] and [21]). The EM algorithm, begin with arbitrary initial values of $\hat{p}_r$ and $\hat{\pi}_{jrk}$, and denote them $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$. The expectation step, calculate the missing class membership probabilities using equation (60.4), substituting $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$ in place of $\hat{p}_r$ and $\hat{\pi}_{jrk}$. The maximization step, update the estimates of the parameters by maximizing the log likelihood function given these posterior $\hat{P}(r_i|Y_i)$, with

$$p_r^{new} = \frac{1}{N}\sum_{i=1}^{N} P(r_i|Y_i) \quad and \quad \hat{\pi}_{jr}^{new} = \frac{\sum_{i=1}^{N} Y_{ij} P(r_i|Y_i)}{\sum_{i=1}^{N} P(r_i|Y_i)} \tag{60.6}$$

as the new prior and class conditional outcome probabilities, respectively; $\hat{\pi}_{jr}^{new}$ is the vector of length K_j of class-r conditional outcome probabilities for the j^{th} manifest variable; and Y_{ij} is the $N \times K_j$ matrix of observed outcome Y_{ijk} on that variable. The algorithm repeats these steps several times until the overall log-likelihood reaches a local maximum and further increments are less than some arbitrarily small value.

Different LCA models have different number of latent classes. Usually, models with more parameters (i.e, more latent classes) provide a better fit, and more parsimonious models tend to have a somewhat poorer fit. So, there is always very close agreement between goodness of fit and parsimony of the latent class models. We can test the goodness of fit of an estimated LCA models by the Pearson Chi-square(χ^2) or the Likelihood Ration Chi-square(L^2) . However, the likelihood ratio Chi-square test, although extensively used in statistical literature, has a number of important limitations. These limitations can be controlled by making use of several information criteria, such as the Akaike information criterion (AIC) [2] and Bayesian information criterion (BIC) [29], each of which is designed to penalize models with larger numbers of parameters. LC models with different number of latent classes are compared and a model with lower AIC and BIC is selected.

4.2 Statistical Hypothesis

The following null hypothesis were considered and tested, at the 5% level of significance, in the current study :

$H_{\circ 1}$: Cheating is independent of gender.

$H_{\circ 2}$: Cheating is independent of current Status of the student.

$H_{\circ 3}$: Cheating is independent of stream of the student.

$H_{\circ 4}$: Cheating is independent of students' perception of academic cheating.

$H_{\circ 5}$: Cheating is independent of students' understanding of material.

$H_{\circ 6}$: Cheating is independent of type of exam given by the teacher to the students.

$H_{\circ 7}$: Cheating is independent of performance pressure on the students.

$H_{\circ 8}$: Cheating is independent of the students' perception about humiliation due to failure.

$H_{\circ 9}$: Cheating is independent of the students' confidence lacking.

$H_{\circ 10}$: Cheating is independent of the students' parental pressure.

$H_{\circ 11}$: Cheating is independent of the students' attitude towards grades.

$H_{\circ 12}$: Cheating is independent of sitting plan of the students in the exam.

$H_{\circ 13}$: Cheating is independent of punishment severity in the educational atmosphere.

$H_{\circ 14}$: Cheating is independent of students' inadequate exam preparations.

$H_{\circ 15}$: Cheating is independent of students' ineffective time management skills.

$H_{\circ 16}$: Cheating is independent of students' habit of laziness.

$H_{\circ 17}$: Cheating is independent of students' perception about instructor vigilance.

$H_{\circ 18}$: Cheating is negatively correlated to the students' subject liking.

$H_{\circ 19}$: Cheating is negatively correlated to the students' interest in the subject.

$H_{\circ 20}$: Cheating is negatively correlated to the students' instructor liking.

$H_{\circ 21}$: Cheating is independent of the student's perception that cheating inevitable.

$H_{\circ 22}$: Cheating is independent of the student's perception of peer influence.

5 Results

We had conducted a pilot survey on 9 departments of the University of Jammu, in order to test the reliability and validity of our questionnaire, before conducting the full survey. The Value of Cronbach Alpha coefficient in this case comes out to be 0.735, which indicates a high reliability according to [10]. Cohen's Kappa index, in this case was 0.61, which shows that face validity holds for our questionnaire. We had also performed factor analysis utilizing principle component analysis with varimax rotation method for testing construct validity and correlation analysis for testing the criterion validity. Both the validities (criterion and construct) holds in this study.

Our data set consists of 1906 observations, but out of that 7 were deleted due to non response and a total of 1899 responses were considered for the further analysis. These responses were, then tested for the identification of missing data values. We had used multiple imputation method to deal with missing data values. With singular imputation methods, mean, median, or some other statistics is used to impute the missing values. However, using single values carries with it a level of uncertainty about which values to impute. Multiple imputation narrows uncertainty about missing values by calculating several different options. We had used SPSS software for performing multiple imputations on over dataset. The imputed dataset is thus, used for the statistical analysis.

5.1 Testing of Hypothesis

Table 1 provides the summary of the hypothesis testing. From table 1 it is clear that the cheating behaviour of the students is independent of the their current-status, type of exam given by the teacher, perception about humiliation due to failure, confidence lacking, parental pressure, attitude towards grades, punishment severity in the educational atmosphere, ineffective time management skills, habit of laziness and perception about instructor vigilance.

Also, Cheating behaviour of students depends on their gender and stream. It also depends on their perception about academic cheating, understanding of material, performance pressure, sitting plan in the exams, inadequate exam preparations, belief that cheating is inevitable and peer influence. Cheating behaviour of the students is found to be positively correlated to the their subject liking or disliking, interest in the subject and instructor liking.

Table 1: Hypothesis test summary

S. No.	Hypothesis	χ^2	Significant value	Decision
1	$H_{\circ 1}$	17.574	0.000	Reject the null hypothesis
2	$H_{\circ 2}$	5.504	0.138	Accept the null hypothesis
3	$H_{\circ 3}$	17.795	0.013	Reject the null hypothesis
4	$H_{\circ 4}$	8.309	0.004	Reject the null hypothesis
5	$H_{\circ 5}$	20.702	0.000	Reject the null hypothesis
6	$H_{\circ 6}$	6.919	0.140	Accept the null hypothesis
7	$H_{\circ 7}$	38.973	0.000	Reject the null hypothesis
8	$H_{\circ 8}$	1.735	0.188	Accept the null hypothesis
9	$H_{\circ 9}$	0.208	0.648	Accept the null hypothesis
10	$H_{\circ 10}$	2.023	0.155	Accept the null hypothesis
11	$H_{\circ 11}$	2.763	0.096	Accept the null hypothesis
12	$H_{\circ 12}$	9.199	0.002	Reject the null hypothesis
13	$H_{\circ 13}$	1.098	0.295	Accept the null hypothesis
14	$H_{\circ 14}$	15.933	0.000	Reject the null hypothesis
15	$H_{\circ 15}$	2.709	0.100	Accept the null hypothesis
16	$H_{\circ 16}$	0.314	0.575	Accept the null hypothesis
17	$H_{\circ 17}$	0.119	0.731	Accept the null hypothesis
18	$H_{\circ 18}$	6.906	0.009	Reject the null hypothesis
19	$H_{\circ 19}$	23.675	0.000	Reject the null hypothesis
20	$H_{\circ 20}$	14.010	0.000	Reject the null hypothesis
21	$H_{\circ 21}$	6.298	0.012	Reject the null hypothesis
22	$H_{\circ 22}$	11.629	0.001	Reject the null hypothesis

5.2 Latent Class Analysis

For the current study, the proportion of female in the sample is 0.60 and male is 0.40; whereas in population this proportion was 0.66 and 0.34. Also, the proportion of the students belonging to science, Technology, Commerce, Arts, Management, Law and other streams is 0.24, 0.08, 0.03, 0.45, 0.03, 0.05 and 0.12, respectively.

5.2.1 Latent Class Models/ Path Models

The path model diagram is the graphical method of displaying the causal relationships among variables in a LCA. A detailed description of the variable used

in the current study for performing LCA is in Appendix A. Given the relatively large number of observed variables measuring the latent variable and the number of response categories per variable, the number of parameters is fairly high. For this reason larger models were not considered. But out of those models, only 4 models provide the efficient results, consequently we proposed following 4 models for estimating the cheating behaviour of the students using LCA.

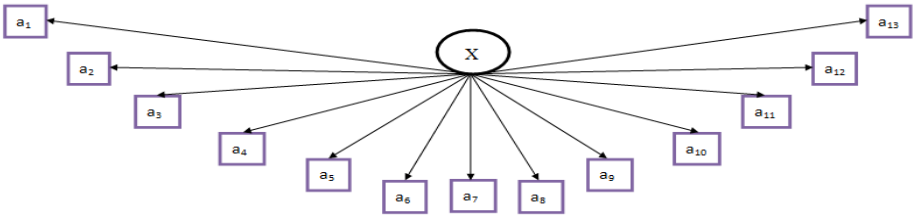

Figure 1: Model 0

Model 0 : Figure 1 is the path model for model 0. This model is the simple LC model without any grouping variable, preserving the assumption of local independence. It will estimates the Cheating behaviour (X) on the basis of the individuals' response pattern to the different indicator variables.

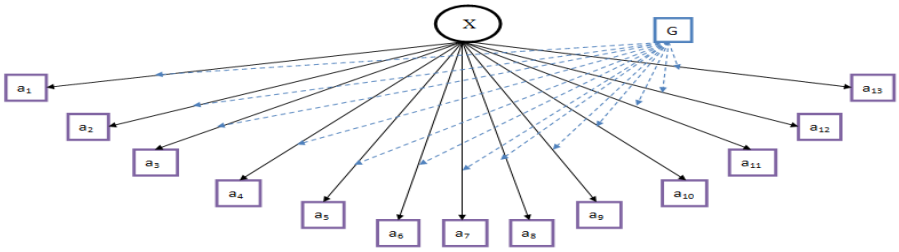

Figure 2: Model 1

Model 1 : Figure 2 is the path model for model 1. This model represents the variation in the indicator variables with the inclusion of grouping variable, gender (G) and estimates the Cheating behaviour on the basis of the individuals' response pattern to the different indicator variables through grouping variable.

Model 2 : Figure 3 is the path model for model 2. This model represents the variation in the indicator variables with the inclusion of grouping variable, stream (S) and will test the influence of stream on the response pattern of the individual to estimate the cheating behaviour.

Model 3 : Figure 4 is the path model for model 3. This model is a complex representation of the variation in the indicator variables on the inclusion of two grouping variables i.e., gender (G) and stream (S). It will estimate the cheating behaviour from the response patterns of the respondents to the different indicator variables under the influence of the gender and stream of the respondents.

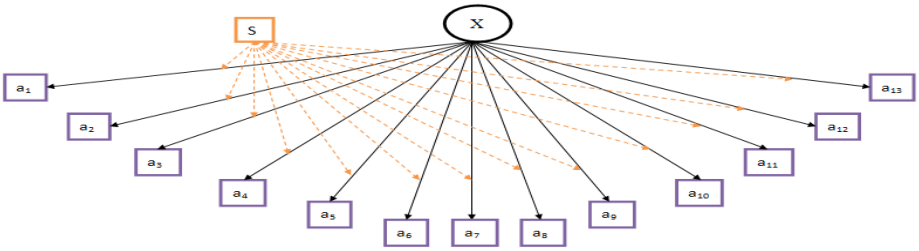

Figure 3: Model 2

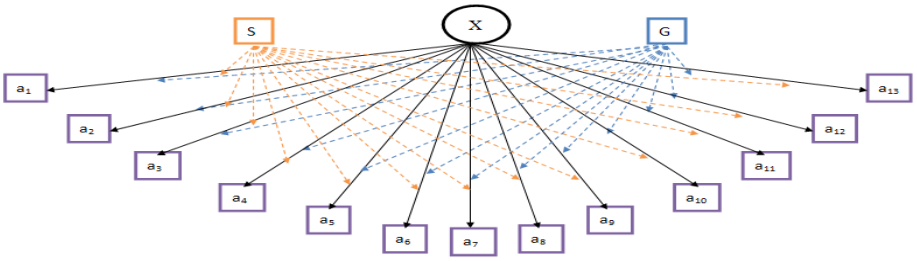

Figure 4: Model 3

5.2.2 Selection of best fitted model

Our data consist of 13 indicator variables which was used for predicting the cheating behaviour of the students. We had incorporated 2 grouping variables with these variables in order to estimate the LC models. As a result, we had to select one best model from the list of the proposed models, which provides the better fit and optimum number of latent classes. Table 2 provides necessary model statistics for different competing models, using poLCA package of R software. Best fitted model is selected on the basis of BIC value [20]. The model with lowest BIC value is always preferred since it provides the best balance between the two factors namely, model fit and model parsimony. From table 2, it can be seen that BIC value of model 3 is lowest i.e. 14954.7 . Also, it's AIC and LL value (14725.37 and -7315.686, respectively) is also satisfactory. Hence, model 3 is used for further analysis.

Table 2: Model diagnostics

Model	d.f.[1]	No. of parameters	LL[2] value	AIC[3]	BIC[4]
Model 0	931	41	-7344.114	14770.23	14970.28
Model 1	929	43	-7330.693	14747.38	14957.2
Model 2	929	43	-7330.405	14746.81	14956.62
Model 3	925	47	-7315.686	14725.37	14954.7

[1] $d.f.$: *Degrees of freedom of error term*
[2] LL : *Log likelihood value*
[3] AIC : *Akaike information criterion*
[4] BIC : *Bayesian information criterion*

5.2.3 Selection of optimum number of classes

A latent class or class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values depending on the responses of the respondents. It also forms the underlying subgroups of respondents based on the observed attributes. In the present study, classes specify the number of categories into which the responses about the personal cheating behaviour falls. In order to identify optimum number of latent classes, we had performed LCA on the selected model 3 to optimize the number of latent classes. This would help us to find a parsimonious model which provides better fit.

Table 3 provides the goodness of fit statistics of model 3 for different number of latent classes. From table 3 it is clear that the data-set was best fitted for Model 3 with 3 latent classes as the corresponding BIC as well as AIC values of that model were lowest.

Therefore, the underlying latent classes can be identified as "Occasional Cheaters" (latent class 1) which represents the group of students who are occasional cheaters and cheat rarely in their academic lives , "Persistent Cheater" (latent class 2) which represents the students who are frequent cheaters and cheat commonly in tests, exams or assignments and "Instantaneous Cheaters" (latent class 3) represents the group of students who are instant cheaters and cheats in their academic lives whenever got chance to do so.

Table 4 provides the Estimated conditional item response probabilities for each of the indicator variables. 1^{st} sub row of table 4 provides the results for the students who actually admitted that they are involved in the cheating activities and 2^{nd} sub row provides the results for students who have denied for being involved in the cheating.

Figure 5 provides the graphical representation of the class membership probabilities for estimation of the 3 class lc model. Each group of red bars represents the conditional probabilities of the indicator variables given the latent variable.

6 Conclusion and Discussion

Based on the objectives, defined hypothesis and the analysis of collected data, following conclusions are drawn:

- The current study defines students (by means of questionnaire) exactly what is meant by academics cheating and make them aware of different types of cheating in which they are involved knowingly or unknowingly.

- Also, it is observed that the main reasons because of which student opt for cheating are not knowing or understanding the study material as it is quite difficult and boring; performance pressure espe-

Table 3: Goodness of fit statistics of model.

Number of classes (n)	n=2	n=3	n=4
Estimated n-class population shares	0.6489 0.3511	0.4181 0.0659 0.516	0.3632 0.3087 0.2686 0.0595
Predicted n-class memberships	0.6605 0.3395	0.3981 0.0689 0.5329	0.356 0.356 0.2263 0.0617
No. of observations	972	972	972
No. of parameters	30	47	64
Residual degrees of freedom	942	925	908
Maximum log likelihood	-7417.268	-7315.686	-7371.238
AIC	14894.54	14725.37	14770.48
BIC	15040.92	14954.7	14982.76
χ^2	15781.95	11435.73	10963.64

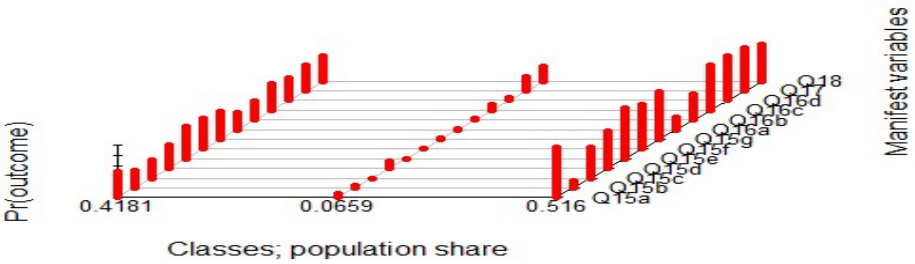

Figure 5: Graphical representation of class membership probabilities

cially because of fear of humiliation due to failure ; inadequate exam preparations; ineffective time management skills; Laziness; disliking of the subject/course and lack of interest in the topic.

- The latent classes, in present study, have been identified as "occasional cheaters", "persistent cheaters" and "instantaneous cheaters", consequently around 42%, 6% and 52% of our respondents are concluded to be occasional, persistent and instantaneous cheaters, respectively.

Academic cheating is a 'disorder' that should be taken seriously to restrain the behaviour of academic dishonesty. It can be challenging to overcome the behaviour of academic dishonesty but an ongoing effort must be taken to lessen its occurrence. The institutions of higher learning should organize programs to promote academic integrity and inculcating an ethical behaviour amongst tertiary students. Students should also be made aware on the negative implications they will receive if they are found to be involved in the academic cheating. Institutions of higher learning should also implement a clear and strict policy on the act of academic dishonesty.

Table 4: Estimated conditional item response probabilities

Indicator variables	Categories of indicators	Latent class 1	Latent class 2	Latent class 3
[t]2*a_1	1	0.5031	0.9208	0.0470
	2	0.4969	0.0792	0.9530
[t]2*a_2	1	0.6765	0.9530	0.8734
	2	0.3235	0.0470	0.1266
[t]2*a_3	1	0.6524	1.0000	0.4183
	2	0.3476	0.0000	0.5817
[t]2*a_4	1	0.5634	0.8724	0.2874
	2	0.4366	0.1276	0.7126
[t]2*a_5	1	0.3954	1.0000	0.0304
	2	0.6046	0.0000	0.9696
[t]2*a_6	1	0.4293	0.9870	0.1661
	2	0.5707	0.0130	0.8339
[t]2*a_7	1	0.4746	0.9751	0.1029
	2	0.5254	0.0249	0.8971
[t]2*a_8	1	0.6907	0.9871	0.7658
	2	0.3093	0.0129	0.2342
[t]2*a_9	1	0.6466	0.9797	0.5114
	2	0.3534	0.0203	0.4886
[t]2*a_{10}	1	0.5077	0.8925	0.1525
	2	0.4923	0.1075	0.8475
[t]2*a_{11}	1	0.5674	0.9468	0.1590
	2	0.4326	0.0532	0.8410
[t]2*a_{12}	1	0.5117	0.7281	0.1882
	2	0.4883	0.2719	0.8118
[t]2*a_{13}	1	0.5288	0.7153	0.3156
	2	0.4712	0.2847	0.6844

References

[1] Aiken, L. S. and West, S. G. (1991). Multiple regression: Testing and interpreting interactions. *Newbury Park*, CA: Sage Publications.

[2] Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In B Petrov, F Csake (eds), *Second International Symposium on Information Theory*, pp. 267-281 Akademiai Kiado, Budapest, Hungary.

[3] Alarape, A. I. and Onakoya, A. Y. (2003). Correlates of examination cheating behavior among university students. *IFE Psychologia: An International Journal*, 11: 71-79.

[4] Beck, L. and Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality*, 25: 285-301.

[5] Davis, S. F., Grover, C. A., Becker, A. H. and Mcgregor, L. N.(1992). Academic Dishonesty: Prevalence, Determinants, Techniques, and Punishments. *Teaching of Psychology* 19(1): 16-20.

[6] Daniel, L. G., Blount, K. D. and Ferrell, C. M. (1991), Academic misconduct among teacher education students: a descriptive–correlational study. Research in Higher Education, 32(6): 703–724.

[7] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B (Methodological)*, 39(1):1-38.

[8] Donse, L. and Van De. Goep, I. H. (2013). Academic dishonesty among college students. Predictors and interventions, Social Cosmos, 10(1): 40–50.

[9] Harding, T., Carpenter, D. , Finelli, C. and Passow, H. (2004). Does academic dishonesty relate to unethical behavior in professional practics? An exploratory study. Conference paper in *3rd conference on Ethics and Social Responsibility in Engineering*, 10: 311-324.

[10] Hinton, P. R., Brownlow, C., Mcmurray, I. and Cozens, B. (2004). SPSS explained. East Sussex, England, Routledge Inc.

[11] Holleque, K. L. (1982). Cheating behavior of college students. Doctoral dissertation, Montana State University. *Dissertation Abstracts International* 43: 88. University Microfilms No. 82-14,649.

[12] Houston, J. P. (1976). The assessment and Prevention of Answer Copying on Undergraduate Multiple-Choice Examinations. *Research in Higher Education*, 5: 301-311.

[13] Houston, J. P. (1983). Alternate test forms as a means of reducing multi-ple-choice answer copying in the classroom. *Journal of Educational Psy-chology*, 75: 572–575.

[14] Jendrek, M. P. (1989). Faculty reactions to academic dishonesty. *Journal of College Student Development*, 30: 401–406.

[15] Jung In, S. B., Jun, M. K. and Shane, C. B. (2015), Students' Unethical Academic Behaviors: A Self Determination Theory Approachm. *Journal of Tourism and Hospitality Management*, 3(2): 60–72.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 629–647. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-60

[16] Jeergal, P. A., Surekha, R., Sharma, P., Anila, K., Jeergal, V. A. and Rani, T. (2015). Prevalence, perception and attitude of dental students towards academic dishonesty and ways to overcome cheating behaviors. *Journal of Advanced Clinical & Research Insights*, 2: 2–6.

[17] Khodaie, E., Moghadamzadeh, A. and Salehi, K. (2011). Factors affecting the probability of academic cheating school students in Tehran. *Procedia-Social and Behavioral Sciences*, 29: 1587-1597.

[18] Koul, R., Clariana, R. B., Jitgarun, K. and Songsriwittaya, A. (2009). The influence of achievement goal orientation on plagiarism. *Learning and Individual Differences*, 19: 506-512.

[19] K12 Academics. Types of Cheating.
`https://www.k12academics.com/Academic%20Dishonesty/types-cheating`

[20] Lin , T. S. , and Dayton , C. M. (1997). Model - selection information criteria for non - nested latent class models. *Journal of Educational and Behavioral Statistics*, 22 , 249 – 264.

[21] Linzer , D. A. and Lewis J. (2011). poLCA: Polytomous Variable Latent Class Analysis. *Annals of Applied Statistics*, 11(4): 2085-2115.
`http://CRAN.R-project.org/package=poLCA`

[22] Marsh, J. D., and Campion, J. (2018). Academic integrity and referencing: whose responsibility is it?. *Journal of Academic Language and Learning*, 12(1), A213–A226.

[23] Mccabe, D. (1999). Toward a Culture of Academic Integrity. *Journal of higher education*, 46(8): B7.

[24] McLachlan, G. and Peel, D. (2000). Finite Mixture Models. *John Wiley & Sons*, New York.

[25] Michaels, J. W. and Miethe, T. D. (1989). Academic effort and college grades. *Social Forces*, 68(1): 309–319.

[26] Orosz, G., Dombi, E., Toth-kiraly, I., Bothe, B., Jagodics, B. and Zimbardo, P. G. (2016). Academic Cheating and time perspective: Cheaters live in the present instead of the future. *Learning and Individual Differences*, 52: 39-45.

[27] Rajendran, A. (unpublished), Research Paper on Cheating.

[28] Romney, M. and Steinbart, P.(2003). Accounting Information Systems. 9th Edition, Prentice Hall, Upper Saddle River, NJ.

[29] Schwartz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6: 461-464.

[30] Sierles, F., Hendrickx, I. and Circle, S. (1980). Cheating in medical school. *Journal of Medical Education*, 55: 124-125.

[31] Skrondal, A., Rabe-Hesketh, S. (2004). Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. *Interdisciplinary statistics* 1st ed., Chapman & Hall; Boca Raton, Fla.

[32] Tom, G. & Borin, N. (1988). Cheating in academic. *Journal of Education for Business*, 63: 153-157.

[33] Vermunt, K. and Jeroen k. (2010). Latent class modelling with covariates: Two improved three-step approaches. *Political Analysis*, 18: 450-469.

7 Detailed description of variables used in LCA

S.No.	Variables	Descriptions
S.No.	Variables	Descriptions
1	X	Cheating behaviour of the students
2	a_1	Used any sort of prohibited material in the exam
3	a_2	Deliberately looked at another student's test sheet or made someone else to look at your test sheets
4	a_3	Passed answers to another person during a test or take answers from them
5	a_4	Planned with another student how to cheat prior to exam
6	a_5	Obtain a copy of an exam paper or test paper before exam
7	a_6	Ever made attempt to obtain or accept assistance from any other person during exam
8	a_7	Lied to an instructor for conducting an exam or test again/ for not appearing in exams or tests
9	a_8	Copied another person's assignment/ research/ thoughts through online/ offline mode and passed it off as your own
10	a_9	Complete the work which is assigned to someone else or made any other person complete the work assigned to you
11	a_{10}	Illicitly obtain material or steal material needed to complete assignment
12	a_{11}	Misrepresenting a family or personal situation (made excuses) to get an extension in assignment
13	a_{12}	Ever prevented other students from from completing their work
14	a_{13}	Ever forged (copy) a faculty/ family/ friend's signature on permission form or add/ drop form
15	G	Gender of the student
16	S	Stream of the student