International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-3, February 2020

Identifying the Dependency Outline of Students' Attendance in Tertiary Level using Markov Chain and Logistic Regression Model

Sayma Suraiya, Gouranga Chandra Debnath, Sayedul Anam

Abstract: ***The schooling of tertiary level is a source of culture for any nation and this issue has always been a major public concern. The success of this level depends on various issues that can influence the students to gather knowledge. Among them, class participation is vital for understanding the course contents. The education system of Bangladesh is almost sound and well-organized in all levels including tertiary stage. Even so sometimes, students are unwilling to attend in the classroom for a number of reasons. In this study, we measured how previous absence of a student impact on their further absence. Markov Chain is a mathematical tool that identified the chance of previous absence effect on the recent absent. In another methods, logistic regression showed the dependency of today's absence on previous absence. In this work, we got, the previous day absence influenced the student to continue his/her absent on the current day. Therefore, students' continuous class participation is important and any sorts of discontinuity makes a barrier to participate in the next class.***

Keywords: Tertiary level of education, Students' attendance, Markov chain model, Logistic regression.

I. INTRODUCTION

Bangladesh's education sector is alienated into four divisions, namely primary (years 1 to 5), secondary (years 6 to 10), higher secondary (11 and 12 years) and tertiary or university level, [3]. Higher education is playing a vital role in preparing professional, competent and far-sighted people to take on various higher responsibilities, to make potential for supporting prosperity in the developing nations [5]. The world absenteeism means a student's absence from school when he is expected to attend that academy. University students' absenteeism may the cause of negative consequences of academic ending of students as well as many societal matters, [12]. It is assumed that, the proportion of higher education attainment in Bangladesh is comparatively low but thanks to the phenomenal growth of private universities. In recent years, the achievement in the education of tertiary level in Bangladesh has increased rapidly due to the growth of several universities. For meeting the growing demands of higher studies, we have 27 public universities (National University and Open University are the exceptional) and 54 private universities in Bangladesh according to the Bangladesh Bureau of Educational Information and Statistics (BANBEIS, 2006) and [9] stated these universities have continued to grow over the past 10 years and growing on. As the growth of a modern civilization enormously depends on the standard of higher education, our students should complete it truly and sincerely. They should be present in the classroom with the course instructor to understand to the contents, to remain up-to-date all the time, to relate the contents with real-life context as so many examples are also analyzed in the classroom. But in our society students are less motivated about joining their classroom rather than using social media, sleeping, idleness, shyness, insincerity, lack of seriousness. Some other issues of absenteeism are media issue, family issue, class atmosphere issue, student's attitude issue, peer issue, student's activity issue, illness, demanding social life, and stress due to load in learning task and assignments and so on,[12]**.** Sometimes this turns to practice. They become habituated not to attending the class/es. For one day, two days three days and going on. As a result some social crime such as teasing, drug addiction, cyber-crime, sex crime, stealing, robbing even jongibad etc. take place in a regular manner.

In our study we are trying to show, is there any impact of one day's absence on further absence?

And it is shown that if any is absent for one day or more than one days then there grow a tendency of remaining absent and which should not be. Because we know participating in the classroom in a regular manner is very helpful to build up their friendship bonds, communal skills, team values, life aids, social awareness, and career pathways and so on.

II. LITERATURE REVIEW

Ahmad [1] shows throughout his research, absenteeism had a significant impact on academic performance. Their results showed that there was a modest negative association between the absenteeism along with the final achievement of assessments. Their research can also create awareness on the academic success for learners regarding the drawbacks of not being attended in their course/es. Monem [11] considers improved understanding between instructors and learners, the adoption of new teaching methods and the commitment of teachers and learners can strengthen the higher education community, and a appropriate educational calendar can make it disciplined. Qutub [13] suggested there should have a system for supporting students towards successful school-to-personal management.

Revised Manuscript Received on February 05, 2020.

* Correspondence Author

Sayma Suraiya, Department of General Educational Development, Daffodil International University, Dhaka, Bangladesh.

Gouranga Chandra Debnath, PhD*, Department of Business Administration, Daffodil International University, Dhaka, Bangladesh. Email:

Sayedul Anam, PhD, Department of General Educational Development, Daffodil International University, Dhaka, Bangladesh .

Their academic performances will also be optimized by this. Suraiya. S [14] shows that socialization

affects the quality and efficiency of workers, motivation, negative turnover ratio and organizational sustainability, including in the case of students. Hossain [10] is using Markov Chain to seek the preceding two-day state transition effect on today's country. The flood event was defined by Anam [2] using the Markov chain. Researchers Al-Zoubi [4] recommended that parents be aware of the problems of their kids and support their academic attainment at every step, improve their concern in the student's physical, emotional and wellbeing, provide a calming school environment and use enticing strategies and educational activities to inspire students to learn.

Since all the above mentioned papers have said regarding the negative correlation between absenteeism and academic performances we then therefore tried to identify the influence of previous absence on today's absence and their dependency pattern to make aware for not being absent and obviously by making the environment friendly.

III. METHODOLOGY

A. Discrete markov chain model

According to P. Billingsley [7], this is a probabilistic model which is on the basis of the supposition that the status of any day depends on the status of previous day. The relationship of dependency is denoted as the dependenc e of first order where of one experiment depends on the consequence of the previous exp eriment and these switching chances are constant. To study the outcome based on the dependence on this conventional process is a two-state di screte Markov Chain Time model. There are some attempts based on m order Markov chains to model such dependency.

Am order's Markov chain is a sequence of trials and all that trials depend on the outcomes of each trial. According to the random variables sequence, if a fixed m is given, $\{X_n\}$forms a Markov chain of order m, for all possible values of the variables X_n (n=0,1,2…..) it is true that

$$P[X_n = j / X_0 = i_0, X_1 = i_1, ..., X_{n-m}] = P[X_n = j / X_{n-m} = i_{n-m}]$$

multiple logistic regression model

$$P(X) = \frac{e^{\beta_0 + \sum_{i=1}^{p} \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^{p} \beta_i x_i}}; i = 1,2,....,p$$

Then the multiple logistic regression model is:

$$f(X) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...+ \beta_p x_p$$

$$in\ which\ case\ P(X) = \frac{e^{f(x)_i}}{1 + e^{f(x)}}$$

B. Data collection

Secondary data has been used for the research which has been collected from the daily class attendance of 50 students from a specific course of Daffodil International University. We have collected the records from a time period of one semester (from January 2019 to April 2019). About 27/28 classes are under consideration.

IV. ANALYSIS AND DISCUSSION

A. Markov chain

Basu [4] states that the explanatory variables can be calculated in various types of scales and are categorized in a dichotomous manner, according to the diversity of subject to be studied.

Now we are defining the variables;

The dependent variable index is (Y) which is dichotomous, assuming with wert '1' and probability P (say) when it is in 'present' status along with the value '0' by the probability (1-P) when it's status is 'absent'.

The independent variables are;

X_1 = Yesterday's status
= 0 if the status is absent
= 1 if the status is present

X_2 = Day before yesterday's status
= 0 if the status is absent
= 1 if the status is present

The transfer counts for Markov's first order model are obt ained by considering the state of student's engagement in the classroom for today and yesterday, where status is perceived to be presence and absence status.

The next table-1 displays the frequency of the changes of the first order taking into account the current status and abse nce of today, followed by the present status and absence of y esterday.

Table -1: Occurrences for the first order transition counts.

Students' Yesterday's Status	Students' attendance status Today's		Total
	Absent (0)	Present (1)	
Absent (0)	281	216	497
Present (1)	205	247	452
Total	457	463	949

In the above table, it is seen that the maximum quantity (0.57) goes to conversion from absent status to absent and minimum proportion (0.43) is in conversion from absent to present status.

Table-2 gives the greatest likelihood estimation of conversion probabilities of a first order Markov Chain attained by the formula;

$$P_{ij} = \frac{n_{ij}}{n_i},\ n_i = \sum_j n_{ij} \quad and \quad P_{ij} = P[X_t = j / X_{t-1} = i]$$

Table-2: The maximum likelihood estimates for the first order model

Transition Probabilities	P_{00}	P_{01}	P_{10}	P_{11}
Greatest Likelihood estimates	0.5184	0.4816	0.504	0.496

From table stated above, we see that being in absent status it is stated that the student was absent also at the preceding class is the highest (0.5184) and moving from absent to present status is the lowest (0.4816).

It is essential to consider the condition of attendance for the three consecutive classes in order to calculate the amount of conversions for the second order chain. That means it is observed whether current day is present or absent given the situation of index at the immediate precedent two classes. The next chart-3 is showing the frequencies of second order Markov chain.

Table-3: Frequencies for second order transition counts.

Students' attendance Status Immediate past two classes	Students' Today's Attendance		Total
	Absent (0)	Present (1)	
0-0	155	106	261
0-1	73	115	188
1-0	95	91	186
1-1	109	120	229
Total	432	432	864

From the above table, it is to be noted that among 864 classes, 58.9% stay in the absent status for three successive classes whereas 52.1% of the classes stay behind present index class state. The remaining of the situations have transformed their status at least once in the three consecutive classes. The maximum part (0.612) is in alteration mode from absent status to present for the next two successive classes whereas the smallest part (0.39) belongs to conversion type of absent status for the day before yesterday to present for yesterday to today's absent status class state. The greatest likelihood estimation of conversion probabilities of a second order Markov Chain attained by the formula;

$$P_{ijk} = \frac{n_{ijk}}{n_{ij}}, n_{ij} = \sum_{k} n_{ijk} \quad and \quad P_{ijk} = P[X_t = k / X_{t-1} = j, X_{t-2} = i]$$

The estimates are shown in Table-4.

Table-4: The maximum probability estimates for the second order model

Transition Probabilities	P_{000}	P_{001}	P_{010}	P_{011}	P_{100}	P_{101}	P_{110}	P_{111}
Maximum Likelihood estimates	0.5572	0.4043	0.4449	0.5551	0.5508	0.4492	0.4368	0.5632

In this table it is seen that the likelihood (0.563) is in present status which says that the status was also in present index for the previous two time points as well as it is the high proportion among all. And the lowest is (0.4368)of being in the absent state where the past two classes were present which is affirmed that if one is used to remain present in the class/es, it is quite unusual for his/her of being absent.

Significance test of the Markov Chain:

Significance test of hypothesis for first order, i.e.

H_0: Zero order Markov chain ($P_{ij} = P_j$)

H_1: First order Markov chain ($P_{ij} \neq P_j$)

The chi-squared test statistics is

$$\chi^2 = 2\sum_{i=0}^{1}\sum_{j=0}^{1} n_{ij}[\log_e \frac{n_{ij}}{n_i} - \log_e \frac{n_j}{n..}]$$

The calculated result of the above test is 35.838821, which is larger than $\chi^2_{0.05,1}(3.84)$. Therefore we can conclude that alternative hypothesis is accepted. That means first order Markov chain is very significant for this case.

Again significance test of hypothesis for second order, i.e.

H_0: First order Markov Chain ($P_{ijk} = P_{jk}$)

H_1: Second order Markov Chain ($P_{ijk} \neq P_{jk}$)

The chi-squared test statistics is

$$\chi^2 = 2\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{1} n_{ijk}[\log_e \frac{n_{ijk}}{n_{ij}} - \log_e \frac{n_{jk}}{n_{.j}}]$$

Where $S^{2-1}(S-1)^2=2$ degrees of freedom. The calculated number of the above test is -1.3837, which is not more than $\chi^2_{0.05,2}(5.99)$. Hence we can conclude that alternative hypothesis is not accepted that means first order Markov chain is significant for that case, [8].

B. Significance test for parameters of logistic regression and status dependence

I. Identification change

Table-5 shows the logistic regression study of the frequent status changes. The total number of observation is 1019. Observation 0 means that the status is 'absent' and the observation 1 means that the status is 'present'. Out of total 949 observations, 452 observations are in the present status and 497 are in absent status. The proportion of students' presence is 47.629%. Conventionally log likelihood is used to determine how the model is appropriate for the data. Here log likelihood is 701.12827 that show the model suits the data perfectly.

Table-5: Logistic Regression output

```
Logistic regression                     Number of obs   =       1019
                                        LR chi2(1)      =       8.56
                                        Prob > chi2     =     0.0034
Log likelihood = -701.12827             Pseudo R2       =     0.0061

------------------------------------------------------------------------------
  Attendance |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Lastday |   .3680231   .1260254     2.92   0.003     .1210179    .6150283
       _cons |  -.2613647   .0875348    -2.99   0.003    -.4329297   -.0897997
------------------------------------------------------------------------------
```

The above table is showing that the value of z-statistics is 0.0034 (<=0.05) which indicates that the model is significant. Here the independent variable last day state is very much significant with today state. From table we also can see if s/he was absent yesterday then the chance of being absent today is 36.8%.

V. CONCLUSION AND RECOMMENDATIONS

This study is based on two key statistical procedures; one is Markov Chain model that determined both the occurrence of daily absence and the present status where another is the logistic regression which explains one binary dependent variable's dependency on the further categorical independent variables. In this research the Markov chain has shown that if one student is absent for one day then s/he would have a tendency for being absented for following day(s). And here the dependency proportion is high. As we know students' presence in the classroom is vital for their academic performance as well as building friendship, team value, social awareness, communication skill development etc. Therefore our target is to make awareness among students regarding their further advancement. Therefore, by means of this study we are recommending to get students in the classroom in a regular basis for the ultimate betterment of the country.

REFERENCES

1. Ahmad, N., Ul-Saufie, A. Z., Mohamed, S. A., Ahmat, H., &Zahari, M. F. (2018, June). The impact of class absenteeism on student's academic performance using regression models. In AIP Conference Proceedings (Vol. 1974, No. 1, p. 050012). AIP Publishing.
2. Anam, S., Parvez, M., Khan, A. R., & Uddin, M. S. (2018). Developing Stochastic Linear Programming Model to Optimize Agricultural Production under Uncertain Flood Influence. RESEARCH JOURNAL OF PHARMACEUTICAL BIOLOGICAL AND CHEMICAL SCIENCES, 9(5), 1358-1369.
3. Al Noor, S., Mustafa, G., Chowdhury, S. A., Hossain, M. Z., &Jaigirdar, F. T. (2010). A proposed architecture of cloud computing for education system in Bangladesh and the impact on current education system. IJCSNS International Journal of Computer Science and Network Security, 10(10), 7-13.
4. Al-Zoubi, S. M., &Younes, M. A. B. (2015). Low academic achievement: causes and results. Theory and Practice in Language Studies, 5(11), 2262-2268.
5. Bangladesh University Grants Commission (2006). Strategic Plan for Higher Education in Bangladesh: 2006-2026, Dhaka.
6. Basu, A. N. (1971). Fitting of a Markov chain model for daily rainfall data at Calcutta. Indian Journal of Meteorology and Geophysics, 22, 67-74.
7. Billingsley, P. (1961). Statistical methods in Markov chains. The Annals of Mathematical Statistics, 12-40.
8. Good, I.J. (1955). The likelihood ratio test for Markov chain. Biometrika, 42, 531-533.
9. Haque, J. H. M., Das, D., &Farzana, R. (2011). Satisfaction of student services in tertiary level: perspective Bangladesh. European Journal of Social Sciences, 19(2), 286-296.
10. Hossain, M. M., &Anam, S. (2012). Identifying the dependency pattern of daily rainfall of Dhaka station in Bangladesh using Markov chain and logistic regression model. Agricultural Sciences, 3(03), 385.
11. Monem, M., & Muhammad, H. (2010). Higher Education in Bangladesh: Status, Issues and Prospects. Pakistan Journal of Social Sciences (PJSS), 30(2).
12. Mohamed, S. A., Ul-Saufie, A. Z., Ahmad, N., Ahmat, H., &Zahari, M. F. (2018, June). Factors influencing student absenteeism in the university. In AIP Conference Proceedings (Vol. 1974, No. 1, p. 050013). AIP Publishing.
13. Qutub, M. F., Bafail, M. A., Alomari, A. S., Azahrani, A. A., Abuznadah, W. T., Munshi, F. M., &Alsaywid, B. S. (2018). Absenteeism among Saudi Medical Students. Egyptian Journal of Hospital Medicine, 70(8).
14. Rahman, S., &Suraiya, S. (2008). An investigation into the Practice of socialization at workplaces in Bangladesh: Human Resource Departments and Employees point of view. Daffodil International University Journal of Business and Economics. Vol. 3, No.2, 131-143.

AUTHORS PROFILE

Sayma Suraiya is an Assistant Professor of the dept. of General Educational Development. She has completed her B.Sc. (Mathematics) and M.S. (Applied Mathematics) from University of Dhaka. Currently she is a PhD researcher at the department of Mathematics, Faculty of Sciences, University of Dhaka. Her current research interest is Operations Research, Management and Mathematical Model.

Gouranga Chandra Debnath, is an Associate Professor and Head of the Department of Business Administration. He has completed his PhD in 2019. He also completed his MBA in International Business from United Kingdom and another in Human Resource Management from University of Dhaka. His current research interest is management, human resource management, supply chain management, leadership and more.

Sayedul Anam, PhD is an Assistant Professor of General Educational Development department. He has completed his PhD in Mathematics in the faculty of Mathematical and Physical Sciences from Jahangirnagar University. MBA- Major in Quantitative Business Analysis, Institute of Business Administration, Jahangirnagar University. B.Sc (Hons.) & M.Sc in Mathematics, Jahangirnagar University His current research interest is development network model.