Scientific Analysis of Scholastic Performance of Students using Markov Chain

Rajesh Menon Kozhimattathil, Shijo Pushpan, Abraham Daivanathan Egan

***Abstract*— The scientific study of scholastic performance of students over a period of time at College/University level has always helped academicians to enhance and refine the pedagogical practices in place and to guide as well as motivate the students for better performance. This paper aims to analyze the status of the student's scholastic performance in Middle East at college/University level using Markov chain. The findings of this analysis can be used in the implementation of innovative and effective techniques at the beginning stages of college/University level study to improve or bring up the scholastic performance of students as well as to optimize the number of graduates and as a result to reduce the dropouts in all over the Middle East.**

***Index Terms*— Homogenous Markov chain,Scholastic performance, stochastic process.**

I. Introduction

Introduction

EDUCATION is an essential element of the progress of any society. Therefore, changes in the field of education are inevitable. As in other developing countries, a number of fundamental changes have taken place at the core of the education system in the Middle East, especially in the field of College/University level education, to meet the need of the time. Educationists around the world are devising new techniques, methods and approaches to make changes in the classical approach to teaching methodology. An automated system called Student Management System (SMS) runs for the management of Registration, Mark entry, Attendance, Reports (all academic and social related activities of the students). It has an Information System Module that generates the statistical reports regarding the academics. Statisticalanalysis of the data generated by SMS will lead to strong conclusions that may help to make changes in the pedagogy itself which in turn can improve the academic performance of the students. The extracted data from SMS is subject to analysis that leads us to significant conclusions. Higher dropout rates in College/ Universitylevel students in the Middle East lead to reduced graduate throughput.

A sample of 950 College/ University level students in the Middle East who joined foundation programme in the semester 1 of Academic Year 2015-2016 is taken into consideration and their academic performance is analyzed through 14 consecutive semesters.

Rajesh Menon Kozhimattathil, UTAS-Salalah
Shijo Pushpan, UTAS-Salalah
Abraham Daivanathan Egan, UTAS-Salalah

Conceptual Framework:Absorbing Markov Chain

Theories and concepts of stochastic processes and Markov Chain are considered here for further investigation.

P(Um+1=j |U0= u0,….Um-1=um-1, Um= i) = P(Um+1=j |Um= i) = Pij where U_m is the state of the process at time 'm' and p_{ij} is the probability of moving from state i to state j in one transition.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

If $p_{ii} = 1$, then state i is an absorbing state. The canonical form of P is:

$P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$, Q is the constraint of P to the transient states and R is the probability matrix which includes the probability from transient state to the absorbing state.

Absorption probability

The probability of absorption is 1 in an Absorbing Markov chain.

That is $Q^v \to 0$ as $v \to \infty$

The Fundamental Matrix

N = $(I-Q)^{-1}$ = $I + Q + Q^2$.....

The entity$n_{ij} \in$ N denotes the anticipatedfrequencyof a process being in the transient state jhaving started at state i.

Absorption probabilities.

Let $B = [b_{ij}]_{t \times r}$ represents the probability matrixin which each elementdenotes the absorbing probability at absorbing state jhaving startedat the transient state i.

$B = NR$is in the canonical form and N is the fundamental matrix.

Database

The data has been analyzed for 4 years starting from academic year 2015-16 to 2018-19 in the MiddleEast. Main streams of the College/University level education include IT, Business and Engineering. Our sub group sample includesstudents who joined for foundation in the academic year 2015-16.Their performance is analyzedthoroughlyfor a sequence of 10 semesters with each semester being a period of 15 weeks of

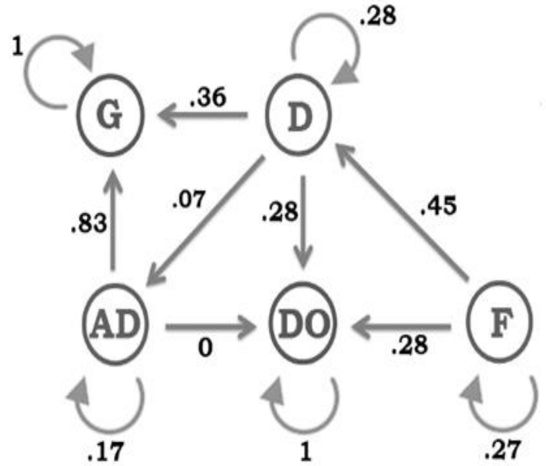
study. The system of College/ University level education in the Middle East is unique and hence this study leads to the evaluation of the percentage of students completed their graduation/ dropout in different levels.
For the student data analysis we use the following notations

- F = Foundation
- D = Diploma
- AD = Advanced Diploma
- G = Graduate
- DO = Dropout

Matrix 1: Transition Data Matrix:

$$\begin{array}{c} \\ G \\ DO \\ F \\ D \\ AD \end{array} \begin{array}{c} \begin{array}{ccccc} G & DO & F & D & AD \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 270 & 260 & 420 & 0 \\ 240 & 190 & 0 & 190 & 60 \\ 50 & 0 & 0 & 0 & 10 \end{bmatrix} \end{array}$$

From Foundation to Graduation state is impossible and hence the value is zero. We have 270 students who dropped out in foundation itself. There were 260 students absorbed in foundation. Number of students who moved from Foundation to Diploma is 420. Movement from Foundation to Advanced Diploma is not possible and hence the value is zero. From the frequency matrix, the probability value is obtained by dividing itself with the total of the corresponding row. Then, the Transition Probability Matrix is shown in Matrix 2 and Transition Probability diagram in Figure 1.
Matrix 2: Transition Probability MatrixP
Figure 1: Transition Probability diagram



Markov chain transition matrix rewritten in its canonical form is a large matrix that can be separated into quadrants: a zero matrix, an identity matrix, a transient to absorbing matrix, and a transient to transient matrix. Here the first quadrant I is an 2 × 2 Identity matrix, second quadrant 0 forms an 2 × 3 zero matrix, third is R a non-zero3 × 2 matrix and fourth is Q, a 3 × 3 matrix, shown in Matrix 3. Notethat in Matrix2, the probability that a student at foundation will remain in foundation stage itself is 0.27. A freshman student at foundation needsto complete four semesters with English language in all the levels, Basic Math and IT for foundation at different levels. The dropout probability at foundation levelis 0.28, which could be due to factors such as failure at any of the four levels of study, dismissal or student being suspended. The probability of a student getting progressed to the diploma level is 0.45. The probability that a student at diploma level to remain in the same level is 0.28. This may be attributed to the fact that some of the courses at this level are tougher than those at lower levels and also few coursesmay be highly technical for a student at that level.For diploma, the student needs to complete 25 courses for which a minimum of seven semesters are required, if the student passes every course he registers in each semester. A student can repeat the course in two cases: (1) If he fails the course (2) To improve his GPA. The probability of Diploma student to dropout is 0.28. The probability of a diploma student moving to Advanced diploma is only 0.16. It is being observed that, a student may go for On Job Training, once the diploma is completed and move to a permanent job further. Advanced diploma branches to specialized streams and student must satisfy the minimum GPA criteria for an entry and a minimum of four semesters are required for completion. But at higher levels since the courses are relativelytough, the studyduration may show significant variation. Here the dropout rate in Advanced Diploma is nil as the student becomes very serious in study at this stage. The probability for Advanced Diploma student to graduate is 0.83 and to remain in is 0.17

Matrices 3. Matrices I,O,R and Q.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0.28 \\ 0.36 & 0.28 \\ 0.83 & 0 \end{bmatrix} \qquad Q = \begin{bmatrix} 0.27 & 0.45 & 0 \\ 0 & 0.28 & 0.07 \\ 0 & 0 & 0.17 \end{bmatrix}$$

Until now wehave takenonly one step in a Markov chain.The probability for multiple step transitions can be obtained by multiplying the transition matrixto itself therequired number of times.

Here, P_{ij}^{n} represents the probability of passing from state *i* to state *j* in n-times step.The sum of probabilities in each row is unity and probability of staying in present state may be non-zero. $P^2 = P.P$ and $P^3 = P^2.P$ is shown in Matrix 4.

Matrix 4. Three step transitionprobability matrix

	G	DO	F	D	AD
G	1	0	0	0	0
DO	0	1	0	0	0
F	0.27725	0.57131	0.01968	0.10211	0.02268
D	0.57327	0.38035	0	0.02195	0.01084
AD	0.99509	0	0	0	0.00491

$P^3 =$

	G	DO
F	0.27725	0.57131
D	0.57327	0.38035
AD	0.99509	0

The R component of this matrix shows the probabilities of being in the absorbing states for the next batch of similar group after three years. See the Matrix5. The probability of a fresh man student to remain in foundation in the forthcoming batch will be 0.27 approximately and chance of dropout on the way is 0.57. Probability of diploma student to graduate is better 0.57 and dropout rate is 0.38. Here too a considerable surety is shown for the probability of an Advanced diploma student to be graduated 0.99 and no student to dropout at this level.

Matrix 5: The constraint matrix Q

$$Q = \begin{bmatrix} 0.01968 & 0.10211 & 0.02268 \\ 0 & 0.02195 & 0.01084 \\ 0 & 0 & 0.00491 \end{bmatrix}$$

$$I - Q = \begin{bmatrix} 0.98032 & -0.1021 & -0.0227 \\ 0 & 0.97805 & -0.0108 \\ 0 & 0 & 0.99509 \end{bmatrix}$$

$$N = (I - Q)^{-1} = \begin{bmatrix} 1.02008 & 0.10649 & 0.02441 \\ 0 & 1.02244 & 0.01114 \\ 0 & 0 & 1.00494 \end{bmatrix}$$

$$R = \begin{bmatrix} 0.27725 & 0.57131 \\ 0.57327 & 0.38035 \\ 0.99509 & 0 \end{bmatrix}$$

$$B = NR = \begin{bmatrix} 0.36815 & 0.62329 \\ 0.59727 & 0.38889 \\ 1 & 0 \end{bmatrix}$$

$$\begin{matrix} & G & DO \\ F & 0.36816 & 0.62329 \\ D & 0.59727 & 0.38889 \\ AD & 1 & 0 \end{matrix}$$

As per the above data, the probability that a student at foundation will eventually be graduated is 0.37 and the probability of a freshman student to dropout is 0.62329. Also there is a chance that a Dropout student may rejoin as a fresh student which we have not taken into consideration in this analysis. Lesser probability, 0.38889 for a Diploma student to dropout. A very promising value 1 is shown for an Advanced Diploma student to graduate and 0 for them to dropout.

Conclusion

Analysis has shown that, the dropout rate is much higher at the lower stages of study than at the higher stages.Though an exhaustive study was not done taking into consideration all the parameters that affect the study behavior of student, it was able to give a precautionary message to give ample boost in the beginning stages of study at College/University. Sudden change in the compulsory language medium from school to college and lesser transition period makes it difficult for the student to understand the courses. Also, the minimum marks to qualify into the level 1 at foundation can be fixed. It should be made sure that the students are given required hours of study for a course and tuned well to match with the requirements at Post Foundation level. In future, an organized and effective implementation of course help centers within the campus can help greatly in escalating the student academic performance.

References

[1] Aulck, L., Velagapudi, N., Blumenstock, J. & West, J. (2016). Predicting Student Dropout in Higher Education. Machine Learning; Computers and Society. Retrieved from http://arxiv.org/abs/1606.06364

[2] Brezavšček, A. & Baggia, A. (2015). Analysis of student's flow in higher education study programmes using discrete homogeneous Markov chains. In L. Zadnik Stirn (Ed.), SOR '15 proceedings, 13th International Symposium on Operational Research in Slovenia (pp. 473–478). Ljubljana: Slovenian Society Informatika, Section for Operational Research.

[3] Adam, R. Y. (2015). An Application of Markov Modeling to the Student Flow in Higher Education in Sudan. International Journal of Science and Research, 4(2), 49–54.

[4] Adeleke, R. A., Oguntuase, K. A. & Ogunsakin, R. E. (2014). Application Of Markov Chain To The Assessment Of Students ' Admi ssion And Academic Performance In Ekiti State University. International Journal of Scientific & Technology Research, 3(7), 349–357.

[5] Moody, V. R. & DuCloux, K. K. (2014). Application of Markov Chains to Analyze and Predict the Mathematical Achievement Gap between African American and White American Students. Journal of Applied & Computational Mathematics, 3(161). http://doi. org/10.4172/2168-9679.1000161

[6] Auwalu, A., Mohammed, L. B. & Saliu, A. (2013). Application of Finite Markov Chain to a Model of Schooling. Journal of Education and Practice, 4(17), 1–10.

[7] Victor Mlambo(Nov 2011)An analysis of some factors affecting student academic performance in an introductory biochemistry course at the University of the West Indies. Caribbean Teaching ScholarVol. 1, No.2, 79–92

[8] ShafiqahAlawadhi and MokhtarKonsowa (2010) Markov Chain Analysis and Student Academic Progress: An Empirical Comparative Study. Journal of Modern Applied Statistical Methods. Vol. 9: Iss. 2, Article26Availableat: http://digitalcommons.wayne.edu/jmasm/vol9/iss2/26

[9] Nicholls, M. (2008). Short term prediction of student numbers in the Victorian secondary education system. Australian and New Zealand Journal of Statistics, 24(2), 179- 190.

[10] Al-Awadhi, S. & Konsowa, M. (2007). An Application of Absorbing Markov Analysis to the Student Flow in Kuwait University. Kuwait Journal of Science and Engineering, 34(2A), 77-89.

[11] Wainwright, P. (2007). An enrollment retention study using a Markov model for a regional state university campus in transition. Master thesis, Department of Mathematical sciences, Indiana University, USA, 2007.