Building Student's Study Path using Markov Chain Process with Apriori Cross Join Pearson Correlation

Tekad Matulatan, Martaleli Bettiza
Computer Science Department, Engineering Faculty
Universitas Maritim Raja Ali Haji
Tanjung Pinang, Indonesia
tekad.matulatan@umrah.ac.id, mbettiza@umrah.ac.id

***Abstract*—Student's study path could be advised by using best possible path from Markov Chain rule based on student's academic performance records with several assumption on the current curriculum. Finding the Markov's rule is crucial process because it will determine study path's scenarios which rely on student current performance to choose the next best possible path. The rule would be built using the whole student's academic performance on the same curriculum by implementing Apriori Cross Join Pearson Correlation Test on two consecutive semesters. It will then create path consist of paired courses A→B with Pearson value that would be implemented as rule in Markov Process**

Keywords*— *Educational Data Mining, Markov Process, Apriori Cross Join Pearson Test, Student Learning Path

I. Introduction

Much in educational data mining researches have focus in finding pattern of learning behavior in accord to predict the student's academic result that could be used by academic advisor in suggesting courses should be took.

Previously in [1], we develop a simple method to find behavior pattern among courses of the same curriculum using cross join of any combination between semesters and calculate the correlation value of the differences grade value of each combination. In [2] and [3] suggest that the prediction of future learning is based on method called Moment by Moment Learning Graph (MBMLG) while in [4] claimed that R-FPA more advance in predicting the study result. The problem on student with no learning or shallow learning that would failed the PFL is handle by detector using K* machine learning as in [5]. Future learning could be accelerated by feature recognition using Probabilistic Context Free Grammar Induction as suggest in [6]. The prediction of student's future performance using an automated detector LOOGCV that claim to be better than Bayesian Knowledge Transfer as in [7].

While the efficiency is main problem in data mining which [8] claims done the information mining efficiently in the data stream using THUI (Temporal High Utility). Some studies were also conducted to improve the performance in an association rules, as in [9] [10] claimed finding a technique that is more efficient to extract the information with a high degree of confidence by association rules, which apply the model of upper bounds and lower bounds in determining the sub-rule apply in data mining. A survey has been taken around this problem and their development in [11]

In the field of educational data mining, where data mining is used as tools in curriculum analysis as applied to college in [12] where each student in each semester that has been in undertaken, using the following data: courses, credits, grades, student id and found the results of the adaptive data mining due to the historical results of the student. Other studies in education data mining: in [13] using Decision Tree and more to the application system, [14] using multiple selection, [15] uses clustering K -Mean, while in [16] using data mining to predict a person's GPA student and the student fails possibilities through the application of regression analysis and C5.0. Reference [17] using a non-linear correlation techniques in analyzing the course management system to find the necessary information from a given dataset where student activity becomes input for EDM to design items that match the student's ability. In [18] proposed matrix factorization method for predicting student performance.

II. The assumption

This assumption of learning path on current curriculum is made to acknowledge there are problems that we should be aware on several things. These problems will distort the result.

- The whole courses grade are given in fair consistent system or assuming no human bias judgment involved (e.g. the whole process of grading come from computer aided assessment system).
- Student personal affairs that could also interfere with the result. The algorithm does not take into the account of this problem.
- There is no significant changes in course's material that could impose student capability to pass.
- There is should no changes in course ID (this could be overcome with equivalence process in table preparation phase)
- The range of time provided in data and number of student involved in one combination of courses must be sufficient enough to produce more valid information.

- Learning path would be effective to be advice for student who has complete minimal the first 2 semesters.

The process contain several steps starting with preparation of the data, then pairing the item set with cross product of one student's transactional record in semester, testing the combination in Pearson's Correlation Test, and last the value putting in to the rule of Markov Chain process.

III. The Preparation

Some academic records system has multiple disperse tables containing the information on student's ID, course's ID, student's study result (Grade's numeric value) and the academic's period of the courses taken. These data should be already prepared in one table based on current curriculum of target study program with no null grade (cancel courses, courses in progress or incomplete) of the same curriculum of study program. If the course is offered to other study program, then the grade result of the student from other study program also be processed. If the course is offered by other study program but is not listed in current curriculum should not be considered, otherwise should. Also cleaning the repeating rows that will distort the result.

TABLE I. PREPARED TABLE OF ACADEMIC RESULT RECORD OF CURRENT CURRICULUM

Semester	Student ID	Course ID	Student Result
Semester S_0	Student St_1	Course C_1	Grade $S_0St_1C_1$
Semester S_2	Student St_1	Course C_2	Grade $S_2St_1C_2$
…	…	…	
Semester S_n	Student St_m	Course C_i	Grade $S_nSt_mC_i$

IV. Pairing and counting the differences

The next process would be creating apriori association by making cross join pair courses from semester n with the next semester n + 1. This could be simplified by stating the specific range of time that would be used in process e.g. the last 4 years records (for current 2015, the last 4 years would be 2010). Filtering the table based on time range could speed up the process.

The process start from the beginning of defining range of time e.g. from Odd Semester 2010 (in odd even semester system) as S_0, Even Semester 2010 as S_1 and soon.

The whole courses in S_0 would be cross-product paired with the whole courses in S_1 (the nest semester) of the same student ID with differences of both grade value $S_1St_xC_j - S_0St_xC_i$ (the next semester value subtract with previous semester grade value of pair courses). The differences then would be count as group result of pair courses.

$$Ci \rightarrow Cj = (\forall Semester\ n :\ \forall Student\ m\ (C_i\ x\ C_j)) \quad (1)$$

$$\forall\ (Ci \rightarrow Cj) = Group\ Count(\ (S_{n+1}St_mC_j -\ S_nSt_mC_i)) \quad (2)$$

For example, in odd semester 2010 (S_0), student St_0 took subject A, B, C with the grade value result consecutively 4, 3, 4. Student St_1 also took subject A, B, C with result 3, 3, 4.

In the next semester, even semester 2010 (S_1) Student St_0 took subject D, E with grade 3, 4 and student St_1 took subject D, F with grade 3, 4. The cross join pair would be A→ D, B→ D, C→D (comes from St_0 and St_1), A→E, B→E, C→E (from St_0), A→F, B→F,C→F (from St_1).

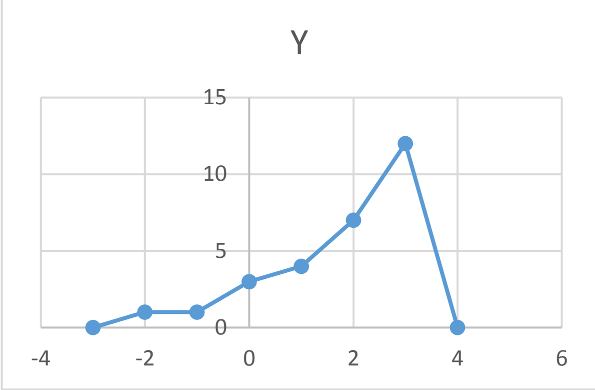
Table II shows the group counting from student St_0 and St_1. The differences grade in A→D for student St_0, would be -1 (grade value 3 – 4) that appears 1 times and the differences 0 is from student St_1 (grade value 3 – 3) that occurs 1 times. Differences 0 mean there is no different in grade result in both pair courses. We can now calculate the Support for each possibilities, where Support is times of occurrences divided by total number of events.

TABLE II. GROUPING THE DIFFERENCES AND COUNTING THE OCCURRENCES, WITH SUPPORT

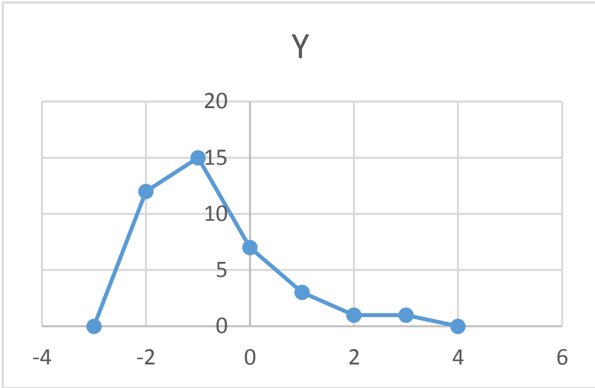
Pair Courses	Differences	Occurrences	Support
A→D	-1	1	½
	0	1	½
B→D	0	2	2/2
C→D	-1	2	2/2
A→E	0	1	1/1
B→E	1	1	1/1
C→E	0	1	1/1
A→F	1	1	1/1
B→F	1	1	1/1
C→F	0	1	1/1

V. Testing the correlation in Pearson-R

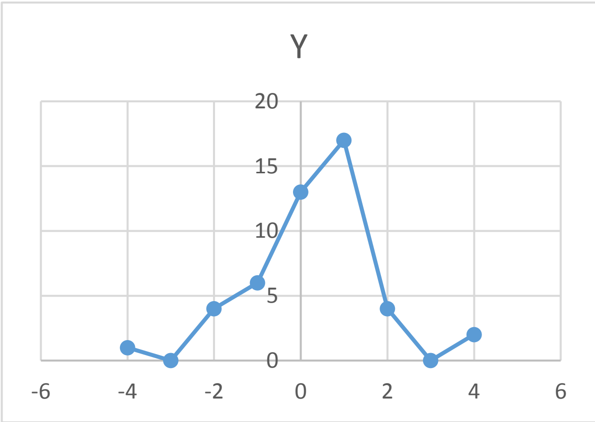
For each pair courses, the differences group count would now be calculated in Pearson-R correlation test to see if the pattern would be in strong or weak, positive or negative or no relation at all. With the value range from 0 to 1 where 0 mean no correlation and more than half to 1 mean strong positive correlation or -1 (strong negative correlation) or any number closer to 0 as week correlation.



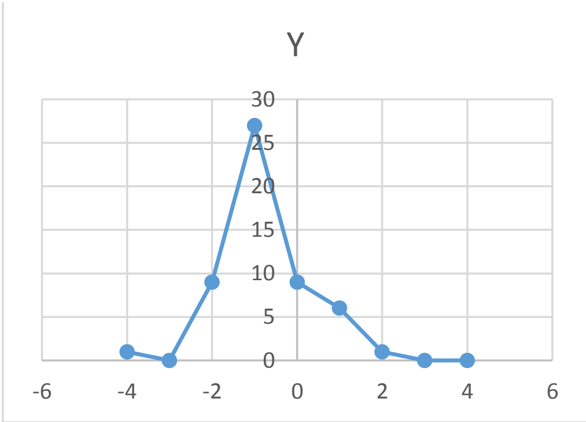
(a)



(b)



(c)



(d)

Fig. 1. The graph of the correlation result; (a) strong positive with Pearson value = 0.52, (b) strong negative with Pearson value = -0.5, (c) weak positive with Pearson value = 0.11 and (d) weak negative with Pearson value = - 0.21

The correlation could be plotted in to Cartesian graph as show in fig. 1, but because the previous example has too little information, could not be used as example chart, instead we use illustrative data that could be seen on table III. The illustrative data in fig.1 show strong or weak of positive or negative correlation on Ci→Cj. The y axis is the number of events and the x-axis is the differences values.

TABLE III. ILLUSTRATIVE COURSE CORRELATION RULE

Ci	Cj	Pearson	Differences	Occurrences	Support
F	H	0.52	-2	1	1/28
			-1	1	1/28
			0	3	3/28
			1	4	4/28
			2	7	7/28
			3	12	12/28
F	I	-0.5	-2	12	12/39
			-1	15	15/39
			0	7	7/39
			1	3	3/39
			2	1	1/39
			3	1	1/39
G	H	0.11	-4	1	1/47
			-3	0	0/47
			-2	4	4/47
			-1	6	6/47
			0	13	13/47
			1	17	17/47
			2	4	4/47
			3	0	0/47
			4	2	2/47
G	I	-0.21	-4	1	1/53
			-3	0	0/53
			-2	9	9/53
			-1	27	27/53
			0	9	9/53
			1	6	6/53
			2	1	1/53
			3	0	0/53
			4	0	0/53

VI. Building Markov Chain Process

Based on the current condition, we could now create rule based on previous step of the Apriori rules and their Pearson's values. The rule itself is a bit different from original Markov Chain which the current state is the last semester with the result, and the future state is the current new semester where student wants to choose the new courses. So the current state would be the last semester's courses with grade, and the future state would be all courses offered in new semester with all already taken non-fail courses will be ignored. The table would be consist of Course Ci an Course Cj, with Pearson value all possible differences (with Support value). For example Table III showing illustrative course F, G offered in Odd Semester and course H, I offered in Even Semester

Using the Table III, we now be able to construct the Markov Chain as illustrated in fig. 2 for F→H and F→I. Same process also for the rest of courses. For each current state, all positive correlation would be put inside the list in rank order with the highest Pearson value. The iteration start from first highest value and repeated to the next high value until number of selected courses satisfy the academic regulation.

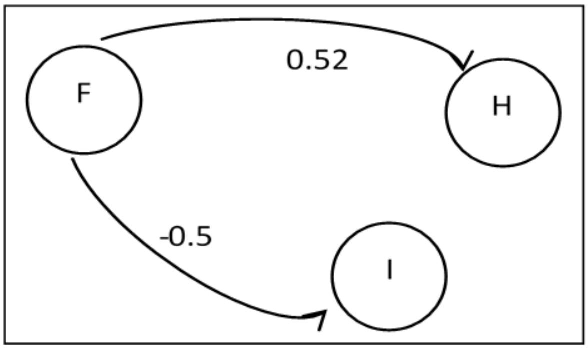


Fig. 2. Illustrative Constructing Markov Chain for Course Path

In figure 2, if the student have pass the F course with "C" mark or 2 in value , then it would be strongly suggest to choose H with possibility to achieve "A" mark (2 marks differences) in low support (7/28).

VII. Conclusion

Student Learning Path could be obtained by implementing Apriori Cross join Pearson Test to find the Curriculum learning pattern and the finding could be devised in other algorithm in this case is Markov Chain Process. The outcome will give the best scenario for student to choose courses in new semester while the actual result of student grade from following the suggested path is not the subject of this paper since the assumption that has been explained before.

References

[1] Matulatan, T., Bettiza, M., Nikentari, N., "Finding student learning pattern using cross join Pearson correlation test", Proceedings ICCS, pages 167-170, Indonesia, July , 2014

[2] A. Hershkovitz, R. S. Baker, S. M. Gowda, and A. T.Corbett, "Predicting future learning better using quantitative analysis of moment-by-moment learning",. In Proceedings 6th EDM Conference, pages 74–81, 2013.

[3] R.S.. Baker, A. Hershkovitz, L.M. Rossi, A.B. Goldstein, S.M. Gowda, "Predicting robust learning with the visual form of the momentby-moment learning curve", Journal of the Learning Sciences, Volume 22, Issue 4, pages 639-666, Taylor-Francis, October, 2013

[4] A. Galyardt, I..Goldin, ."Recent-performance factors analysis", In Proceedings of 7th International Conference on Educational Data Mining. London, UK, 2014

[5] S.M. Gowda, R.S. Baker, A.T. Corbett, L.M. Rossi "Towards Automatically Detecting Whether Student Learning is Shallow", International Journal of Artificial Intelligence in Education, November 2013, Volume 23, Issue 1-4, pp 50-70 , Springer

[6] N. Li, W.W. Cohen, K.R. Koedinger,"A Computational Model of Accelerated Future Learning through Feature Recognition", Lecture Notes in Computer Science, vol. 6095, pp 368-370, Springer, 2010 [10th International Conference, Intelligent Tutoring Systems, 2010, Pittsburgh, PA, USA, June 14-18, 2010, Proceedings, Part II]

[7] R.S. Baker, S.M. Gowda, A.T. Corbett,"Towards predicting future transfer of learning", Lecture Notes in Computer Science, vol. 6738, 2011, pp 23-30, [15th International Conference, Artificial Intelligence in Education 2011, Auckland, New Zealand, June 28 – July 2011]

[8] V. Tseng, C.J. Chu, T. Lyane, "Efficient Mining of Temporal High Utility Itemsets from Data streams", UBDM'06, August, USA, 2006

[9] B. Goethal, J. Muhonen, H. Toivonen, "Mining Non-Derivable Association Rules", Conference on Data Mining, USA, 2005

[10] W. Pedrycz, M.H. Smith, "Granular Correlation Analysis in Data Mining", IEEE, 1999

[11] A. Ceglar, J.F. Roddick, "Association Mining", ACM Computing Survey, 2006

[12] R. Knauf, Y. Sakurai, K. Takada, A. Tsuruta, "A Case Study on Using Personalized Data Mining for University Curricula", IEEE International Conference on Systems, Man, and Cybernetics, Seoul, 2012

[13] V.T. Chau, N.H. Phung, "A Knowledge-Driven Educational Decision Support System", IEEE, 2012

[14] N. Delavari, M.R. Beikzadeh, P. Somnuk, "Application of Enhanced Analysis Model for Data Mining Processes in Higher Educational System", ITHET 6th Annual International Conference, 2005

[15] S. Parack, Z. Zahid, F. Merchant, "Application of Data Mining in Educational Databases for Predicting Academic Trends and Patterns", IEEE, 2011

[16] M. Nasiri, F. Vafaei, B. Minaei, "Predicting GPA and Academic Dismissal in LMS Using Educational Data Mining: A Case Mining", 6th National and 3rd International conference of e-Learning and e-Teaching, 2012

[17] B. Liu, J.H. Zhao, "Non-linear correlation techniques in educational data mining", Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009

[18] N. Thai-Nghe, L. Drumond, A. Krohn-Grimberghe, L. Schmidt-Thieme, "Recommender System for Predicting Student Performance", Procedia Computer Science, vol 1 issue 2, p.2811-2819, Elsevier, 2010