Deep learning on curriculum study pattern by selective cross join in advising students' study path

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

Muhammad Resha
Electrical Engineering Department
Universitas Hasanuddin
Makassar, Indonesia
resha@unhas.ac.id

***Abstract*— Advising engineering students in their study path need to understand the curriculum structure, student capabilities and challenge that commonly appear in courses. This paper offered the simple method to help student advisor in analyzing student performance in their study path based on academic progress record of the student it-self and pattern that been built from other students that have taken the courses. Using selective cross join for each possible permutation of pair courses with respect to courses' grade to create knowledge base. This knowledge base will be used to construct complex tree of any possible study path that might be taken by student to reach the end of study including course that must be retaken. Finding the best suggestion for study path using Monte Carlo tree search style**

Keywords*— *Educational data mining; Student learning path; Monte Carlo algorithm; Deep learning; cross join association

I. INTRODUCTION

Advising undergraduate students in their study path need to understand the curriculum structure, student capabilities and challenge that common appear in courses. It is very usual in engineering courses that some student surpasses greatly in some courses but not in others. Some student has best competency in one skill but lack in other skills. The lecturer or instructor grading process also could affect the outcome of the students' grade result. This research took place in Indonesia and before we are discussing further on the method, let's have a brief look in Indonesia's Higher Education System. It is common that one study program redesigns the curriculum after 4 or 5 years, and the new curriculum could have additional new courses or replacing old courses with new one whether the content has great similarity or totally different. In some universities, new curriculum means old curriculum become obsolete and no longer been applied and for the senior students all old courses that been passed should be transform into new curriculum using the equivalence table. Other universities don't have single curriculum policy causing new curriculum only applied to new students. The number of credits that one student could apply in next semester will depend on GPA of the recent semester. Some courses only offered in odd or even semester. To simplify the complex situation the assumption been made and discussed in later section.

There are number of research on predicting study behavior like in [6] using extended Bayesian Knowledge Tracing to estimate the student knowledge that claim to be significantly better prediction, [5] also using Bayesian Knowledge Tracing to detect student skill in knowledge component (KC). Study in tutoring Algebra [4] using causal model discovery on behavior in learning where all possible behavior is drawn to estimate linear structure with the weight value. Other explore learning exhibit behavior in problem solving [7] while in [8] using activity sequence clustering to model student behavior, [2] using Markov chain to describe the path, [9] developed a Hidden Markov Model to discover student behavior trends with different learning processes in problem solving.

Beside the study on learning behavior, this paper also looks on UCB that discuss in [1][3] which implemented to choose the best possible path

II. THE CRITICS

This paper improved the previous study that using cross-join to create any possible study path [2]. Although the method implements cross join that will cause exponential time consuming as the data increase, it needed to draw any possibilities course path that had been taken by student with the course grade result. The cross-join process doesn't need to be recalculate each time new data been added to avoid draining CPU resources. All processed information could be stored in database and new information could be added to continue from the stored information.

Even though storing the processed information in persistence media, the problems of exponential time $O(n^2)$ still exist especially when a student took a lot of courses in the last semester. The number of useless processing data from cross-join also increase as the pair has repeated elsewhere with exactly same result or only happened once or don't make sense and for that has wasting resource and time.

978-1-5386-0549-3/17/$31.00 ©2017 IEEE

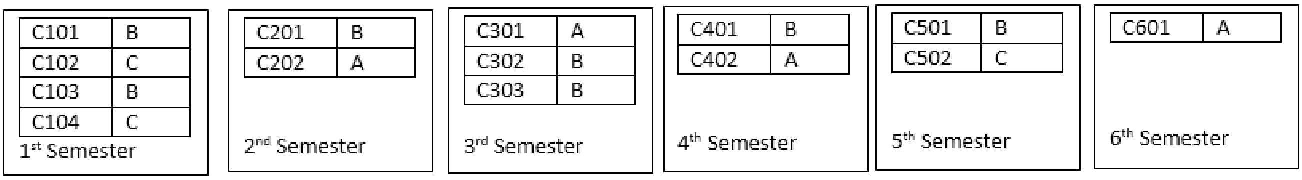

Fig. 1. Example of student's records

Realized the drawback on this problem, while needed to map all possible study path with the outcome grade, it should be avoided by using selective cross-join which repeated pair with exact result does not need to be processed. Using the selective cross-join to create base knowledge of study pattern that will be used later in finding best possible study path. More on this will be discussed later.

III. The assumption and limitation

This research has been conducted in Indonesia Higher Education system and the assumption had to be made to acknowledge that there are differences in other countries Higher Education system. The assumption also should be aware on the simplified complex problems that could distort the result. The assumptions that been implied in this research are:

- There is correlation among courses whether it is weak or strong correlation. [2] This will create study pattern.
- The grade assessments are assumed to be consistent across the span of time of the analyzed data.
- Although lecturer personal interest or personal judgment could bias the result of one student but for overall result might giving fair model in course study patterns. In some extreme cases that could happened where the lecturer giving the prejudice grade for the whole class, for cases like this we might ignoring the data.
- The curriculum might change significantly from previous version to respond with current demand but if it change drastically (which is almost the whole content are new courses) then it cannot be continued process from previous data and should be stop and recreate the new analysis tree.
- Creating new analysis tree would be ineffective if the curriculum courses is the first time executed with no previous history. New courses that replace old courses but with major similarity content with the old courses could be assumed same and need not to create new analysis tree.
- Courses contents are assumed to be less modified. Changes in the courses contents to be tougher in response to recent situation may create hardship for student to pass the course. This situation also should be treated as new courses rather than current courses because content and assessment has changed and if not been treated as new courses will make confusion in the historic pattern of the courses.
- All courses should be treated first in equivalence table so the courses from different curriculum with similar content or objective could be assume as the same.

IV. The process

Similar with the previous study, the step in the process start from data preparation which select and process the required attributes of the record and put it on one table, then do the selective cross join to pair the course from semester n with the next semester n+1 for each individual in each semester. The result will be stored as base knowledge that later will be used in Monte Carlo tree search. Monte Carlo tree search has been implemented many on computer games with repeated trial and error. In this case, the repeated trial and error (win lose) been replaced by knowledge base. More details of each phase will be discussed next.

A. Data preparation

In data preparation, we going to make preparation table based on academic record of all eligible students in one study program within certain time. Eligible students are students that going thru or finish with target curriculum and its equivalence (if any). The target curriculum is the current implementation but could also using past curriculum in case want to measure the error with cross validation.

For easiness, the example of curriculum is assumed has the same amount of credits for all courses and each semester offer 4 courses with first character represent the semester and the number, e.g. C401

In Fig.1 represent example of a record from a student "St01" that has completed in 3 years from 2012 to 2015. The curriculum that the student undergo has been changed once and the Course ID has changed also. The figure showing the result of transforming the old courses' ID to new ID based on equivalence table from previous curriculum. If some old courses that do not appear in new curriculum and have no equivalence with new courses in the new curriculum then these courses should not be included in the preparation table. The preparation table containing the information on time of course taken by student, student's ID, course's ID, student's course result (best in numeric value for the sake of speed on process). This whole information should be placed in one table according to target curriculum of target study program with no null grade (withdraw, courses in progress or incomplete).

Some study program permits the students to take course in other study program. If this the case, then before listing in the preparation table, should check the number of students that had taken the course from outside study program. If the number is considered too little then it should not be included in the

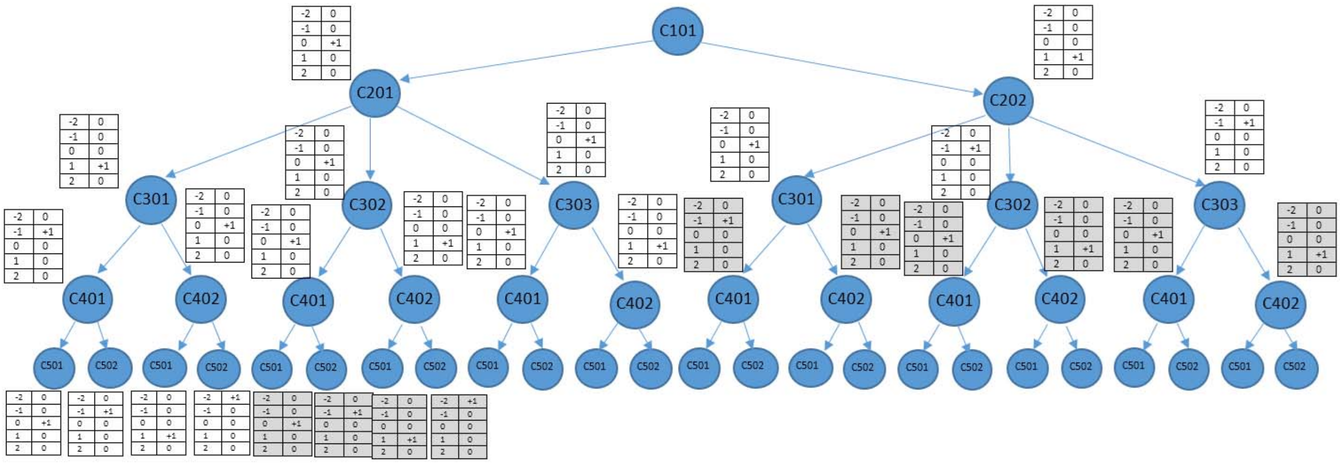

Fig. 2. Creating knowledge base from selective cross-join operation using example records of one student

preparation table. Also, if one student has failed one course more than 2 times then the only the last 2 could be include in preparation table since this might the problem on student and not represent the course study pattern.

TABLE I. PREPARATION TABLE FROM YEAR 1 TO YEAR N WITH EQUIVALENCY

Semester	Student ID	Course ID	Student Result
20121	St01	C101	B
20121	St01	C102	C
20121	St01	C103	B
20121	St01	C104	C
20122	St01	C201	B
20122	St01	C202	A
…	…	…	
Semester S_n	Student St_m	Course C_i	Grade $S_nSt_mC_i$

B. Pairing the courses and building knowledge base

This phase would create pair courses that is course taken in semester n with all possible courses taken in semester n+1. Using selective cross join to pair the courses from semester n with the next semester n + 1. The selective mean not all possible courses should be pair with all other courses. This could be done by filtering out courses that is not need in process. The processing starts on each individual student record. Example in fig. 1 show a record of students who have graduated in 3 years. Pairing the first and second semester will give 8 pair courses. Pairing the second and third semester will give 6 pair and so on. The pairing process also will record the grade from left side pair and grade from right side pair, differences of grade among two courses and number of this grade model appeared.

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 possibility, where Support is times of occurrences divided by total number of events.

TABLE II. COUNTING DIFFERENCES IN GRADE OF THE PAIR COURSES AND NUMBER OF OCCURRENCE OF THE SAME MODEL FROM WHOLE

Pair Courses	Grade	Differences	Occurrences
C101-C201	B→B	0	1
C101-C202	B→A	1	1
C102-C201	C→B	1	1
C102-C202	C→A	2	1
C103-C201	B→B	0	1
C103-C202	B→A	1	1
C104-C201	C→B	1	1
C104-C202	C→A	2	1
C201-C301	B→A	1	1
C201-C302	B→B	0	1
C201-C303	B→B	0	1
C202-C301	A→A	0	1
C202-C302	A→B	-1	1
C202-C303	A→B	-1	1
..	..	..	..
C502-C601	C→A	2	1
..	..	..	..
C_i-C_j	G_i→G_j	diff (G_j, G_i)	n

From the second table, after all data been processed, grouping the record based on pair courses and differences to become knowledge base. For instance, the pair course C_i-C_j with the differences (*diff* (G_j, G_i)) on grade C→B, B→A and D→C are just 1, so they sum up in the same group and number of occurrences are the total from those grades. The forth column is the Upper Confidence Bound as in (1) that will compute the chance of grading.

$$\bar{x}_i \pm \sqrt{\frac{2\ln n}{n_i}} \quad (1)$$

Where

$\bar{x}_i$ is the average of occurrences of differences in one pair

n is the number of occurrence

n_i is the sum of occurrence in one pair

This knowledge base would be used to create tree and should be store in persistence media. Any new information will be added to the correspondent data in knowledge base without having to recalculate everything from beginning.

TABLE III. GROUPING THE PAIR COURSES, DIFFERENCES, NUMBER OF OCCURRENCE AND UPPER CONFIDENCE BOUND

Pair Courses	Differences	∑Occurrences	UCB
C401-C501	-2	12	-1.24+0.49
C401-C501	-1	3	-1.24+0.32
C401-C501	0	5	-1.24+0.39
C401-C501	1	1	-1.24+0
C401-C501	2	0	-1.24+0
C401-C502	-2	1	0.79+0
C401-C502	-1	1	0.79+0
C401-C502	0	10	0.79+0.40
C401-C502	1	8	0.79+0.38
C401-C502	2	9	0.70+0.39
C401-C503	-2	1	1.17+0
C401-C503	-1	0	1.17+0
C401-C503	0	5	1.17+0.37
C401-C503	1	5	1.17+0.37
C401-C503	2	12	1.17+0.46
..	..	..	
Ci-Cj	diff (G_j,G_i)	n	

C. Building the tree

Using the knowledge base to create tree could produce the duplicate branch which could be copy from the first copy or put the pointer to the first copy (in programming). In Fig. 2 display the example tree created from one student that we used as sample here. The gray boxes show the duplicate parts and also give the idea that these should not be processed repeatedly in cross-join. This would save a lot of resource power. The table box in fig. 2 show the information on differences and number of events at first and second column respectively.

When creating the tree, it could expand more complex since some failed courses could be retaken in next year and could lead to branch looping.

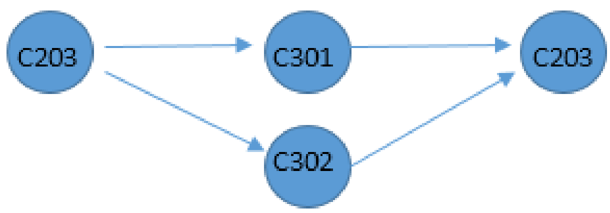

Fig. 3. Path with failed course

If this the case then it should be pruned if the number of occurrences showing insignificant or too little. This also could be divided into 2 separate branches represent 2 separate case because it might be some case C203 been took the first time after C301 without necessary means the student failed previously. The complete knowledge base on study pattern has finished. Using this tree is a little modified from Monte Carlo tree search by not doing the trial and error to learn but recalculate the weight of occurrences with highest sum of differences number from the end or goal.

D. Selecting the potential path

When deciding the best path for a student e.g. student "Stx" that going to 5th semester with result on 4th semester as following C401 with B, C402 with C and C403 with Fail. With current result, this student could only take 2 courses on 5th semester. This student still need 3 courses in 5th semester and 1 final project course in 6th semester. All mandatory courses have been passed except C301 that failed in third semester and C403 that failed in past semester. The goal has to reach C601 that is final project for the end of study. From knowledge base tree, it would select the all branch from passed courses in 4th semester as starting point and ignoring the duplicate branch. The selected tree branch might look like fig. 4 for C401 to C601.

Using the knowledge base, and the rule that this student could only take 2 courses then the option would be rank from the best choice to second best: C503 and C502. Let's assume that for C402 the options are C504 and C502. The course C301 should be retaken, and that's why this tree is not give the final solution or reach the C601.

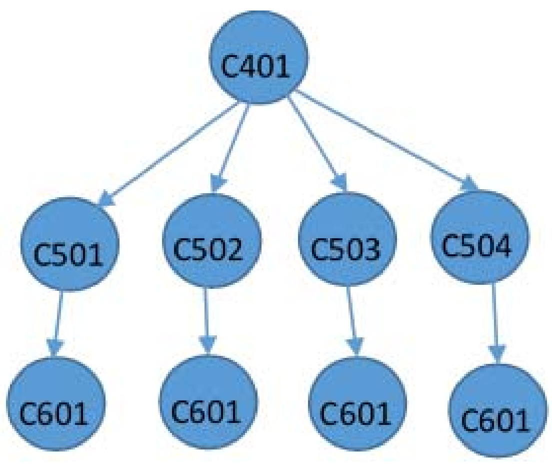

Fig. 4. Potential path from start point to finish the study

E. Expanding the path and choosing the best alternative

Because C301 only offer in odd semester, that's mean it should appear in odd semester. Expanding the tree by copy from C301 branch and pruning all sub-branch that the course already been taken and passed. Also in even semester, there is C403 that should be retaken. That is mean on the 6th semester, C601 cannot be taken and cut from this part. Because there is no other course in 4th semester except C403 then it is the only course chosen in 6th semester and the branch copied from C403 with purging all sub branch that already been taken and passed.

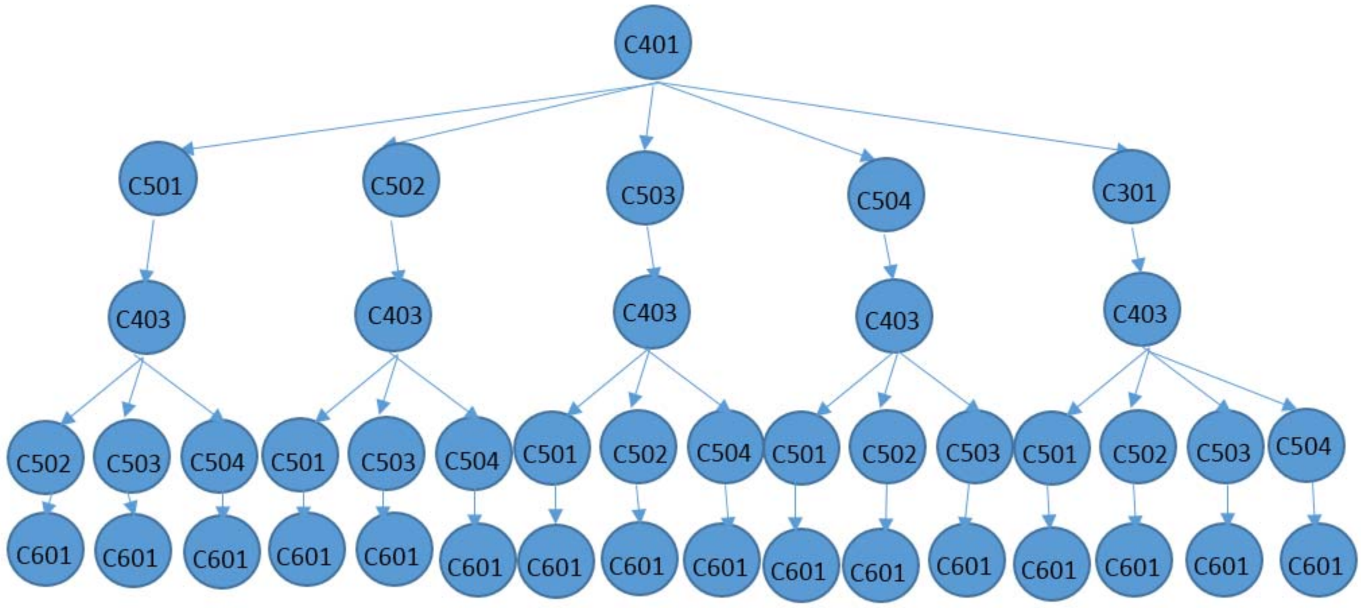

Fig. 5. Path expansion

The tree eventually will end up more like fig. 5. It is also showing the time needed to graduate more than 3 years as expected. The candidate rank is path with the highest sum of UCB to the lowest from each course in path.

F. BackPropagation

After completing ranking all path from the each starting course then we need to sum the highest UCB from bottom (the goal) up to starting point and then store in value in the bottom (goal) so it will show which path is the highest UCB. Student advisor could use this information to compare the highest value from all the possible path that means it is a big chance for the student to gain better mark.

V. Conclusion

This method will select the best path for student to gain higher GPA while the consequences is the easiest course is tend to preferred first then the hardest one. This might be violated the curriculum tree that has prerequisite system. Even though the algorithm has not been measured for the error, the algorithm only suggests the best scenario for student to choose courses in new semester based on the common trend in study pattern in particular study program with target curriculum. The suggestion might be wrong if the student study behavior also change.

References

[1] A. Peter, N. Cesa-Bianchi and P. Fischer, "Finite-time Analysis of the Multiarmed Bandit Problem," *Machine Learning,* vol. 47, pp. 235-236, 2002.

[2] T. Matulatan and M. Bettiza, "Building Student's Study Path using Markov Chain Process with Apriori Cross Join Pearson Correlation," in *Proceeding of the Electrical Engineering Computer Science and Informatics Vol 2*, Palembang, 2015.

[3] P. Auer, "Using Confidence Bounds for Exploitation-Exploration Trade-offs," *Journal of Machine Learning Research,* vol. 3, pp. 397-422, 2002.

[4] S. E. Fancsali, "Causal Discovery with Models: Behavior, Affect, and Learning in Cognitive Tutor Algebra," in *Proceedings of the 7th International Conference on Educational Data Mining*, London, 2014.

[5] R. S. Baker, A. B. Goldstein and N. T. Heffernan, "Detecting Learning Moment-by-Moment," *International Journal of Artificial Intelligence in Education ,* vol. 21, no. IOS, pp. 5-25, 2011.

[6] Z. A. P. D. D. T. Heffernan, "Learning What Works in ITS from Non-traditional Randomized Controlled Trial Data.," in *The 10th International Conference on Intelligent Tutoring Systems,*, Pittsburgh, 2010.

[7] D. Davis, G. Chen, C. Hauff and G.-J. Houben, "Gauging MOOC Learners' Adherence to the Designed Learning Path," in *Proceedings of the 9th International Conference on Educational Data Mining*, North Carolina, 2016.

[8] M. Kock and A. Paramythis, " Activity sequence modelling and dynamic clustering for personalized e-learning," *User Modeling and User-Adapted Interaction,* vol. 21, pp. 51-97, 2011.

[9] M. Kock and A. Paramythis, "Towards Adaptive Learning Support on the Basis of Behavioural Patterns in Learning Activity Sequences," in *The 2nd International Conference on Intelligent Networking and Collaborative Systems*, Greece, 2010.

[10] H. Jeong and G. Biswas, "Mining student behavior models in learning-by-teaching environments," in *First International Conference on Educational Data Mining.*, Montreal, 2008.