A Model of Bayesian Network Analysis of The Factors Affecting Student's Higher Level Study Decision: The Private Institution Case

Poh Choo Song[1], Hui Yee Chong[1], Hong Choon Ong[2], Sing Yan Looi[1]
[1]Department of Physical and Mathematical Science, Faculty of Science, Universiti Tunku Abdul Rahman, 31900, Kampar, Perak, Malaysia.
[2]School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Gelugor, Penang, Malaysia.
songpc@utar.edu.my

Abstract— **A growing number of universities and colleges established in Malaysia have provided many options for students to further their tertiary education either in public university or private institution. Therefore, this study tends to identify the factors that affect the decision made by students to pursue their Degree in a private institution based on Bayesian networks (BNs). BN provides a unique graphical structure that indicates the dependency between variables. A total of 150 valid responses are collected from a private university by the means of questionnaire as the survey instrument in this study. The study found that reputation of the university is not the main concern of the students, meanwhile, undesired course offered from the public university and family advices are the significant factors influencing their decision to pursue their degree in a private university.**

Index Terms— **Bayesian network; Factor analysis; Higher level study; Study decision.**

I. INTRODUCTION

There are over hundred registered universities and colleges establishing and providing plenty choices for students to pursue their tertiary education. Selecting an appropriate university or institution becomes an important concern for those who want to continue their tertiary education. However, the competition among higher education industry is getting higher since there are various universities or institution with interesting courses available for students' choices of university selection. Note that the general use of the terms private institution and private university and college can be considered interchangeable in most contexts below.

Good reputation and high ranking university has the higher chances in influencing students' university selection. Some universities attracting students with their highly qualified lecturers and facilities that are especially suitable for students study environment. Besides, parents will always inspire their children in selecting the university. Parents will feel secure to let their children to study in this environment. Moreover, some of the students prefer to follow their siblings or friends so that they can study in the same university. However, the selection process from government for public university admission has stopped many students in enrolling in public university. This may due to students had been offered to the course that they do not desire or the public university is very far from their hometown. As a result, they look for other alternatives to pursue their tertiary education such as enrolling in private universities or colleges, or giving up on studies in the worst case.

The main focus of this study is to identify what factors influence the most for the students to pursue Degree in a private institution. Factors that influence students' selection of university can be course fees, location of university, expenses, institution's reputation, recommendation from family, quality of lecturers, the graduate employment rate, career opportunity, financial aids, facilities, siblings or friends are studying in certain institution, course offered by public university is unwanted and etc. Throughout this study, we would able to get a graphical view on identifying the main factors that affect the students' selection. It is also important in helping some private institutions to understand the student needs.

The rest of the paper is organized as follows. Section 2 focused on the literature review and some explanation or concepts of Bayesian network. In section 3, we introduce the method that is being carried out in this study. The description of the statistical data used in this study is presented in Section 3 as well. Results of findings and discussion are reported in Section 4. Finally, conclusion is discussed in Section 5.

II. LITERATURE REVIEW

A. Factors affect the students' university selection

First of all, reputation of an institution is one of the factors affecting students' decision on higher level study. According to [1], the reputation of a university can be passed on from students through conversation with other people, or through advertising activities in promoting the university. Study show that the increase of university's ranking will influence the student's enrolment [2]. Therefore, good reputation of a university with higher ranking will affect the decision of students' selection ([3], [4] and [5]). As a result, university needs to build a distinct image in order to have a higher competition with other universities [6].

We believe that the increase of course fees and registration of new intake students has a negative correlation. The increase of course fees will noticeably decrease the amount of student

enrolls in a university. However, students may overcome the financial problem with the financial aids such as scholarship or government loan [7]. Provision of financial aids can decrease the burden of students in term of cost of fees [8]. According to [9], the ways of payment, provision of financial aids and affordable accommodation fees will affect the decision of institution selection. Therefore, course fees and financial aids are probably the part of the consideration of students to choose a university to further studies. In most of the private institution they do provide scholarship or government loan for students to apply when they enroll in the particular institution as long as they fulfill certain criteria.

The high quality of facilities provided by university that are able to affect the decision of students' enrolment includes library, computer labs and self-learning areas [10]. In addition, students also look forward that university can provide unlimited internet access facilities, sufficient and adequate library books and computers for them to access. Provision of other facilities can attract students' attention such as recreational facilities and sport equipment [11]. Besides facilities concern, the distance between hometown and institution has the significant reason to affect the decision of students to pursue their higher education [12]. Studies showed that most of students prefer to study at the university that is nearer to their hometown. Therefore, expenses and living cost are in the consideration of students as well when choosing a university to further studies. Moreover, this can reduce the academic and the economic crisis among the students [13].

Recommendation from family has always influence students' selection because family members have closer relationship with them. Besides, siblings or friends who are studying in a certain university are another significant factor for students in choosing a university. Research showed that there is a very close relationship between the communications of secondary school students with undergraduate who are still pursing at university. If secondary school students interact frequently with undergraduate students, they will be influenced and choose to further their studies in the same university ([11] and [12]). As a result, there is a high tendency that students will follow to pursue their tertiary education with their friends who are studying in a university so that they can get guidance or accompany by siblings or friends who are studying in the same university. They might feel comfortable and safe when studying in the same university with them.

The great emphasis on teaching quality has important implication in lecturers' attitude. Often, students complain that some lecturers have the attitude of "it is just a job" in teaching such as reading from the slides or reference books and did not prepare well before entering the lecture class. Students, in fact, prefer qualified, patience, knowledgeable and motivated lecturers to enhance the learning environment. In addition, the lecturer's expertise in teaching specific courses and having adequate working experience in related course field are more favorable by students [14].

The graduate employment rate is another concern of students in choosing university. Students wish that university can provide advice and information to help them in finding future job according to their major course [14]. Therefore, students expect the course of their study field can provide better job opportunity in the future. As a result, some students

have the thought of "studying for better career" subsequently they are more concern on the course they are taking and whether they can engage with the future employment.

From the previous studies, most researchers have several methods such as economic model, sociological model, factor analysis Chi-square analysis and descriptive analysis in theirs papers to identify the factors meanwhile in our study we apply Bayesian network in order to identify the factors ([2]; [15], [16] and [17]).

B. Bayesian Network

Bayesian network (BN) is a directed acyclic graph (DAG) with joint probability distribution which consists of graph structure and computational architecture. BN can also be defined as the set of nodes and directed edges where it represents the variables and causal relationship respectively [18]. BN is unique in relation to other graphical and probabilistic analysis tool as the uncertainty is solved mathematically in an effective and simple way [19].

In 1985, Pearl officially characterized BN and designed an effective algorithm to compute the probabilities but it can be only applied to particular structure [20]. In the early of 1990, BN became famous tools as the uncertainty is solved mathematically in a simple way. This is due to the adoption of effective algorithm to calculate the probabilities without structural limitation ([21] and [22]) and the introduction of modeling software [23]. From the studies of [24] and [25], the dramatical development of machine learning technique enables its application widely used in many field. Few years later, BN was showed as a competitive tool which is used to recognize or classifies pattern [26].

Bayesian networks are widely applied in different areas such as engineering, education, medical, biology and others. Bayesian network equipped selection of best tests' technique to carry out on patients in order to decrease the uncertainty of diseases [27]. Shannon entropy is a good measure to identify the fee of the tests needed and can diagnose the side effect of the patient. This application can help physicians to overcome the obstacle of diagnostic reasoning. Besides, prognostic reasoning can predict the future in the field of biomedicine and health care. This is considered as a new development in the medical field. For instance, traditional survival analysis implement in BN.

Garcia and his friend used BN to detect students' learning style in web based system [28]. There are comparison between the approach of Bayesian networks and result of Index of Learning Style (ILS) questionnaire which are used to evaluate students' learning style. Bayesian networks help to combine qualitative and quantitative data from experts and database. Thus, it can be used to predict the expected and unexpected loses in financial areas [19].

III. DATA DESCRIPTION AND METHODOLOGY

A. Research Data

In this study, we aim to investigate the factors that influence student's higher level study decision in private institution. We targeted our respondents to be Year 1 students from the private university in Malaysia. The purpose of Year 1 students are chosen as they are freshmen and have a clearer picture in

their mind about the university selection compared to others who are less favorable in our study.

The data were collected from random sample survey by distributing the questionnaire to the year one students from all the faculties in one of the private university in Perak, Malaysia. A total of 200 questionnaires were disseminated, however only 159 copies of respondents were used as a complete dataset. Incomplete dataset were ignored since missing value is not in the consideration of this study. The description of the data is shown in Table 1. The 159 respondents' demographic profile shows 42.1% and 57.9% are male and female respectively. Majority of the respondents are from art and social science students which contribute to 30.2%, followed by information technology students at 20.1%, business and finance students at 18.2%, science stream students at 17% and engineering students at 14.5%.

Table 1
Demographic Profile of Respondents

Demographic	Category	Number	Percentage (%)
Gender	Male	67	42.1
	Female	92	57.9
Study Field	Art and social science	48	30.2
	Business and finance	29	18.2
	Engineering	23	14.5
	Information Technology	32	20.1
	Science	27	17.0

Table 2
List of Variable Description

Variable	Possible value	Description
Qualification	4	Qualification of enroll in current university: STPM, Diploma, Foundation and UEC
Quality	4	Quality of lecturer on average: worst, fair, good, very good
Fee	3	Total course fee: RM35000–RM40000, RM40001–RM45000 and RM45001 and above
Financial_Aids	5	Type of financial aids receive: university scholarship, external scholarship, external loan, PTPTN and none
Reputation	2	Prefer the reputation of the private university: yes or no
Facilities	2	Enjoy facilities in the private university: yes or no
Course Offer	2	Get desire offer from public university: yes or no
Employment	2	Graduate employment rate is high: yes or no
Family	2	Follow family recommendation: yes or no
Career	2	Career opportunity: yes or no
Location	4	Current study university location: strategic and near hometown, strategic but far from hometown, not strategic but near hometown, not strategic and far from hometown
Parent_Ed	3	Parent highest education level: primary school level, secondary school level and college or university level
Private Uni	2	Made the right choice to study in a private university?: yes or no

As shown in Table 2, there are more than 10 possible factors listed in questionnaire as the potential factors that have been adapted from previous studies. These factors, for example, include location, reputation and facilities of the university, course fees, recommendation from family, quality of lecturers, financial aids, the graduate employment rate, career opportunity, course offered by public university is unwanted and etc.

B. Research Method

Bayesian networks consists of qualitative (structural learning) and quantitative (parameter learning). The structural learning can be represented by graphical view whereas parameter learning can be showed by the probability table after the structure is obtained [29]. The algorithms used to learn the structure of the pattern include K2 algorithm, grow and shrink algorithm, hill climbing (HC) algorithm, simulated annealing algorithm and others [30]. K2 algorithm is applied in [24] while [31] applied HC algorithm in their paper. In this study, we focus on the structural learning of BN in identifying the factors that affect students' decision on higher level study.

Structural learning of BN consists of score-based and constraint-based [32]. Score-based algorithm is the popular approach that induces BN from data. According to [33], Bayesian Information Criterion (BIC) score is useful when the prior information is unknown or difficult to find. The score for BIC for a graph *G* which is formed based on data *D* is defined as follow:

$$BICscore(G, D) = \log \Pr(D \mid \hat{p}, G) - \frac{d}{2} \log N$$

where $\hat{p} = $ the set of maximum-likelihood estimates of parameter **p** of the BNs, $d = $ the number of the free parameters of the multivariate Gaussian, and *N* is the sample size.

In this study, we use hill climbing (HC) algorithm which is one of the most popular choices in score-based method. It is started with an empty network and followed by computation of the scores. The algorithm's main loop consists of attempting every possible addition, removal or reversal and making the network increasing the score in the most current applicant. The iteration will stop if there are no changes in increasing the score. Hence, the network with best score is obtained. Figure 1 as shown best describes the search procedure of using HC algorithm when learning BN structure.

The software used in this study is Rstudio version 0.99.489 with bnlearn and Rgraphviz packages installed to form the BN structure. The default score used for HC algorithm in Rstudio software is the BIC score.

IV. RESULTS AND DISCUSSION

The output of the graph structure obtain from Rstudio based on hill climbing algorithm is shown in Figure 2. It is reasonable to assume that the quality of lecturer is not in the factors list of BN structure. This is due to that the students do not know the lecturers before enrolling in private institution so they would not have a clear picture about the quality of lecturers. As a result, we conclude that the quality of lecturers does not influence much in the decision of student's selection to pursue Degree in private institution based on our findings. At the same time, educational level of parent does not have any influence on student selection on their tertiary education.
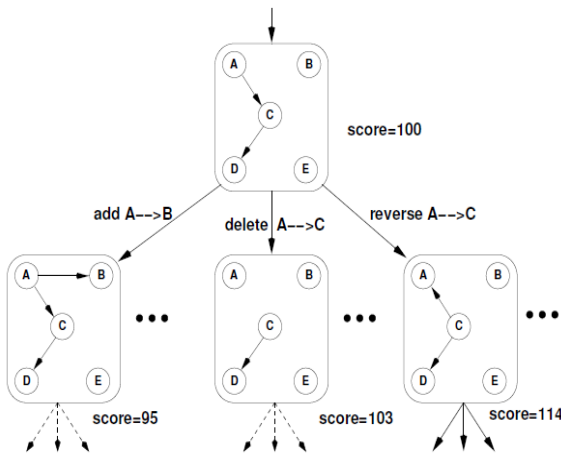
Figure 1: Illustration of a BN structure using HC search procedure [32]

Other variables such as qualification, financial aid and course fees, also do not strongly affect the decision of students' university selection. The qualification of students do not affect mainly in enrolling private institution. This is understandable as some private institutions implement first come first serve basis. Thus different qualifications have the same chance to enroll in a private institution. Although some private institutions provide financial aids to students but it does not affect much in decision of students selecting private institution since other universities are also providing financial aids to them. The course fees can be covered by the financial aids provided so it would not become the concern of the student to enroll in a private institution.
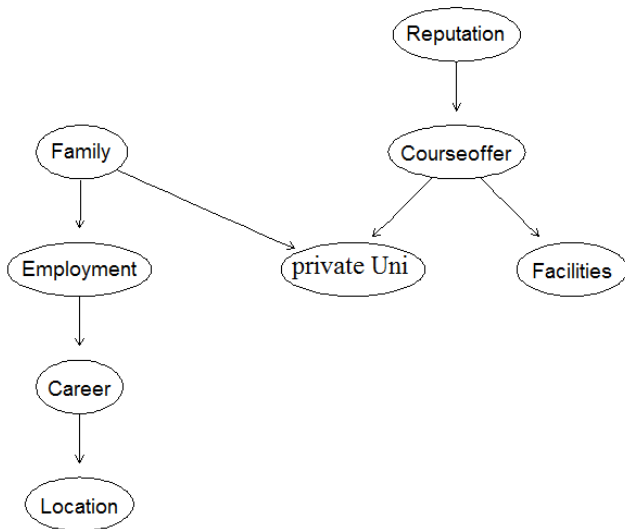

Figure 2: Graph structure using hill climbing algorithm

Based on the results obtained, we may clearly see that one of the main reasons students had chosen a private university as their tertiary educational institution is because they did not get the desired course that offered by the government. Secondly, students take the recommendation from family and surprisingly most of the parents suggest their children to enroll in a private institution. However, we notice that, knowing some private institutions have good reputation; students did not directly choose that institution until the course offered by government is undesirable.

There is an indirect relationship between "Private Uni" and "Facilities". We found that students enjoy the facilities provided by the private institution as their university selection after they get unwanted course offered and enrolled in the private institution. On the other hand, there is an association between "Family" and "Employment" which indicates that family recommended their children to pursue in a certain university and by considering some private universities that ensure a good graduate employment rate. This can persuade more students to enroll in private institution. There is high future career opportunity given that the good graduate employment rate from the private institution graduated students. Besides, there is interrelationship between "Career" and "Location". Some students may prefer pursue their tertiary education near from their hometown and convenient for them to go back home. If there is high future career opportunity in this strategic location then it will encourage more students to enroll in private institution. Nevertheless, the missing arrow encodes the information that the dependency between the nodes does not exist. For instance, "Family" is unrelated with "Courseoffer" because the unwanted course offered by public university has no relationship with the recommendation of family to their children.

V. CONCLUSION

Throughout this study, the usefulness of Bayesian networks enables us to identify the main factors that affect the decision of students to pursue their Degree in private institution. The graph structure using hill climbing algorithm illustrated the main factors that influence students' university decision. Although there are many algorithms available to plot the graph, hill climbing algorithm is the main choice to carry out in this study. Other algorithms can be carried out in the future studies. It is recommended that different algorithms could be used for the future studies. Different networks can be compared and to search the best graph to solve the problem.

This study highlighted two main factors that are recommendation from family and unwanted course offered by public university are regarded as extremely important in the decision to pursue studies in private institution. The failure of getting interested courses in public university enables students to find alternative way that is studied in a private institution.

Last but not least, this study is carried out hopefully be useful for not only private institution but also parents or students themselves in understanding the actual needs in making decision to pursue a degree and therefore improve the civilization in Malaysia in this 21st century. In particular researchers can consider other aspects of the student decision-making process such as personal factors, family background, academic achievements, and others possible factors.

REFERENCES

[1] J. P. Ivy, "Higher education institution image: A correspondence analysis approach", *The International Journal of Educational Management*, vol. 15, no. 6, pp. 276–282, 2001.

[2] J. L. Fernandez, "An exploratory study of factors influencing the decision of students to study at UniversitiSains Malaysia", *Kajian Malaysia*, vol. 28, no. 2, pp. 107-136, 2010.

[3] T. W. Mazzarol, "Critical success factors for international education marketing", *International Journal of Education Management*, vol. 12. No. 4, pp. 163-175, 1998.

[4] B. A. Bourke, "A Model of the determinants of international trade in higher education", *The Service Industries Journal*, vol. 20, no. 1, pp. 110-138, 2000.

[5] J. Gutman, and G. Miaoulis, "Communicating a quality position in service delivery: An application in higher education", *Managing Service Quality*, vol. 13, no. 2, pp. 105–111, 2003.

[6] R. Paramewaran, and A. E. Glowacka, "University image: An information processing perspective", *Journal of Marketing for Higher Education*, vol. 6, no. 2, pp. 41–56, 1995.

[7] A. F. Cabrera, and S. M. La Nasa, "Understanding the college-choice process", *New Directions for Institutional Research*, vol. 107, pp. 5–22, 2000.

[8] D. Hossler, J. Schmit, and N. Vesper, *Going to college: How social, economic and educational factors influence the decisions students make*. Baltimore: John Hopkins University Press, 1999.

[9] N. Foskett, F. Maringe, and D. Roberts, "Changing fee regimes and their impact on student attitude to higher education", *Higher Education Academy UK*, vol. 22, no. 2, pp. 23–31, 2006.

[10] I. L. Price, F. Matzdorf, F. Smith and H. Agahi, "The impact of facilities on student choice of university", *International Journal of Educational Management*, vol. 21, no. 10, pp. 212–222, 2003.

[11] M. Joseph, and B. Joseph, "Identifying needs of potential students in tertiary education of strategy development", *Quality Assurance in Education*, vol. 6, no. 2, pp. 90–96, 1998.

[12] T. Shanka, V. Quintal, and R. Taylor, "Factors influencing international students' choice of education destination: A correspondence analysis", *Journal of Marketing for Higher Education*, vol. 15, no. 2, pp. 31–46, 2005.

[13] G. A. Jackson, "Public efficiency and private choice in higher education", *Educational Evaluation and Policy Analysis*, vol. 4, no. 2, pp. 237–247, 1982.

[14] C.B. Kandiko, and M. Mawer, *Student Expectations and Perceptions of Higher Education*. London: King's Learning Institute, 2013.

[15] C. Avery, and C. M. Hoxby. *Do and should financial aid packages affect students' college choices? In College choices: The economics of where to go, when to go, and how to pay for it*. Chicago: University of Chicago Press: 239–302, 2004.

[16] Samsinar Md. Sidin, SitiRahayuHussin, and H. S. Tan, "An exploratory study of factors influencing the college choice decision of undergraduate students in Malaysia", *Asia Pacific Management Review*, vol. 8, no. 3, pp. 259-280, 2003.

[17] D. T. Ellwood, and T. J. Kane, Who is getting a college education? Family background and growing gaps in enrollment. In *Securing the future: Investing in children from birth to college*, eds. S. Danziger and J. Waldfogel. New York: Russell Sage Foundation, pp. 283–324, 2000.

[18] F. V. Jensen, and S. K. Andersen, "Approximations in Bayesian belief universes for knowledge-based systems", In *Proceedings of the 6th Conference on Uncertainty in Artificial Intelligence*. March 27, 2013, pp. 162-169.

[19] M. Neil, N. Fenton, and M. Tailor, "Using Bayesian network to model expected and unexpected operational losses", *Risk Analysis*, vol. 25, no. 4, pp. 1-10, 2005.

[20] J. Pearl, "A constraint propagation approach to probabilistic reasoning", in *Proceedings of the First Conference on Uncertainty in Artificial Intelligence*. July 10-12, 1985, pp. 31-42.

[21] F. V. Jensen, S. L. Lauritzen, and K. G. Olesen, "Bayesian updating in causal probabilistic networks by local computation", *Computational Statistics Quarterly*, vol. 4, pp. 269-282, 1990.

[22] P. P. Shenoy, and G. Shafer, "Axioms for probability and belief functions propagation", *Artificial Intelligence*, vol. 4, pp. 169-198, 1990.

[23] S. K. Andersen, K. G. Olesen, F. V. Jensen, and F. Jensen. HUGIN: a shell for building Bayesian belief universes for expert systems. *Readings in Uncertain Reasoning, Kaufmann*,San Mateo, pp. 332-337, 1990.

[24] G. F. Cooper, and E. Herskovitz, "A Bayesian method for the induction of probabilistic networks from data", *Machine Learning*, pp. 309–347, 1992.

[25] P. Spirtes, C. Glymour, and R. Sheines, Causation, Prediction and Search. *MIT Press*: 41-86, 1993.

[26] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers", *Machine Learning*, vol. 29, pp. 131-163, 1997.

[27] P. J. F. Lucas, C. Linda, and A. Abu-Hanna, "Bayesian networks in biomedicine and health care", *Artificial Intelligence in Medicine*, vol. 30, pp. 201-214, 2004.

[28] P. Garcia, A. Amandi, S. Schiaffino, and M. Campo, "Using Bayesian networks to detect students' learning styles in a web-based education system:, In *Proceeding of the Argentina Symposium on Artificial Intelligence*. August 29-30, 2005, pp. 115-126

[29] I. Ben-Gal, "Bayesian networks", *Encyclopedia of Statistics in Quality & Reliability*. Wiley & Sons, pp. 1-6, 2007.

[30] M. Scutari, "Learning Bayesian networks with the bnlearn R package", *Journal of Statistical Software*, pp. 1-22, 2010.

[31] D. M. G. Taborda, and L. Zdravkovic, "Application of a hill climbing technique to the formulation of a new cyclic nonlinear elastic constitutive model", *Computers and Geotechnics*, vol. 43, pp. 80-91, 2012.

[32] D. Margaristis, *Learning Bayesian network model structure from data*. PhD thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, 2003.

[33] G. Schwartz, "Estimating the dimension of a model", *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, 1978.