An Application of the Probabilistic Model to the Prediction of Student Graduation Using Bayesian Belief Networks

Jiraporn Yingkuachat[1],
Prasong Praneetpolgrang[1], and **Boonserm Kijsirikul**[2], Non-members

ABSTRACT

This paper proposes an alternative to the prediction of education accomplishment. It employs a data mining technique, the Bayesian belief network (Bayes net). The technique is used to analyze the independent variables that affect the education accomplishment result of vocational students, undergraduate students and graduate students. The machine learning tool, WEKA, is used to construct the prediction model that is accurate for the prediction based on k-fold cross-validation.

The experimental result shows that the Bayes net technique is able to determine important variables for the prediction of the result of education accomplishment and this technique provides high prediction accuracy. From the models constructed by WEKA, we find that the important variables that affect the education accomplishment are the previous GPA., mother and or fathers career, the total income of the family and the grade point average when they enter the first year in bachelor study. The obtained result is consistent with the result analyzed by multiple regression analysis.

Keywords: Bayesian Belief Networks, Data Mining, Multiple Regression Analysis

1. INTRODUCTION

The capability to generate and collect data has been increasing rapidly in the last several decades. Contributing factors include the computerization of many business, scientific, and government transactions, and the advance in data collection[1],[3].

The educational institutions are organizations which have large amounts of student data. This research has the objective to find the significant variables that affect the prediction of student graduation by using a data mining technique. The data mining can bring to create the prediction model by using the classification of student data[2],[13].

The prediction model is used to predict the possibility of graduation based on the existing student database.Here, we employ the Bayesian belief network, developed from the principle of Bayes theorem [3], for the model construction.

2. THEORIES AND RELATED WORKS

This section briefly describes concepts and theories related to this research, i.e. data mining[3],[4],[6] Bayesian belief networks [5], [11],[21]as well as related works as follows.

2.1 Data Mining

Data mining refers to extraction or “mining” knowledge from large amounts of data [6] and the process of data mining includes the following steps.

1. Data Selection: This step is to identify the data sources for the mining process.
2. Data Pre-processing: This step is for data preparation by using several methods, such as screening out non-valued data, uncorrected data, redundant data and inconsistent data, collecting data from many databases and examining the quality of selected data.
3. Transformation: This step is to transform the selected data into appropriated format for compatibility with the data mining algorithm.
4. Data Mining: This is the main process which uses data mining techniques for discovering the model. The mining techniques can be grouped into the following categories.
4.1 Predictive Data Mining: This kind of techniques constructs models to anticipate or estimate distinct values of data from historical data.
4.2 Descriptive Data Mining: This kind of techniques outputs models to describe explain some characteristics of existing data, or classifies data into several clusters by the characteristics of the data.
5. Interpretation and Evaluation: This step is to interpret and evaluate the obtained result; tools for visualization of the result are also helpful in interpreting the result.

Manuscript received on December 16, 2006; revised on March 16, 2007.
[1]The authors are with the school of Information Technology, Graduate School, Sripatum University, Bangkok, Thailand; E-mail: yjiraporn@gmail.com., prasong@spu.ac.th
[2]The author is with the Department of Computer Engineering, Chulalongkorn University, Bangkok, Thailand; E-mail: boonserm.k@chula.ac.th.

2.2 Bayesian Belief Networks

A Bayesian belief network (Bayes net) [7],[9] is a graphical model for describing relationship among variables. A Bayes net is composed of nodes representing variables. Each node is asserted to be conditionally independent of its non-descendents, given its immediate parents. Associated with each node is a conditional probability table (CPT), which specifies the conditional distribution for the variable given its immediate parents in the network. Therefore, we can use a Bayes net to depict the conditional independency among variables. The Bayes net facilitates the combination of prior domain knowledge and data. If the prior knowledge about the dependency of variables is available, we can use the knowledge to draw the structure of the network, and we can use data to train the probability values in the CPT. If we have no prior knowledge, the structure of the network can also be learned from the data .

2.3 Related Works

In the last decade, there are many researches that employ data mining techniques for analyzing educational data. Waiyamai et al.[8] analyzed the student database system by using data mining techniques to improve quality in education for engineering faculty. They studied and analyzed the student database by using the knowledge engineering to solve some problems of student graduation.

Hendricks [2] analyzed student graduation trend of Texas technical college. The sample data was data warehouse of three technical colleges in Texas. The Knowledge SEEKER IV TM program was used to analyze the data. The results of this research indicated important variables including independent variables which affect the graduation of students.

3. RESEARCH METHODOLOGY

The important feature of the Bayes net is the ability to explain the causal relationship among variables and show such relationship in a graphical model. In this paper, we employ WEKA [4] for creating the prediction model. We also use 10-fold cross-validation for evaluating the performance of the model(see Fig.1). The obtained results will be compared to statistical analysis methods.

All variables appearing in the student database are shown in Table 1. First, some of these variables, i.e. Id and Sex, are removed, as they are irrelevant to the classification of student graduation. Eighteen variables are remained.

The example groups of students used in this research are as follows.

Vocational education degree

1. Student data from the Thai administration and commerce school contains 408 records.

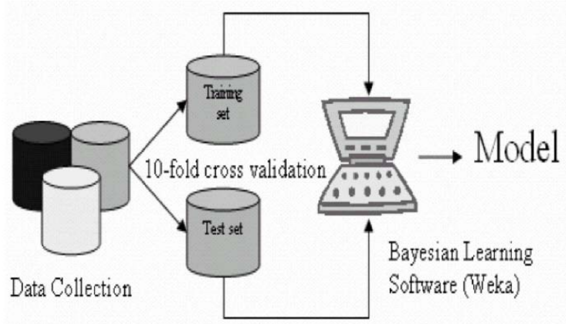

Fig.1: *Research Methodology*

Table 1: *All variables in the student database*

Variable Name	Explanation
Id	Running number
Sex	Gender of student
Age	Age of student
Add	Student's address
Status	Marital status of student's parents
Occfa	Occupation of student's father
Occmo	Occupation of student's mother
Income	Total income of whole family
Expenses	Student's expenditure
Numson	Number of sons in family
ChildRank	Birth rank of children of student in family
Olded	Previous academic degree
Oldsch	Previous school/institute
Oldsch2	Type of school/institute
GPA1	Previous GPA
GPA2	GPA in the first semester
Getin	How to get in the school/institute
GrdE	Grade in English subject
Com	Having computer or not
Finish	Education Status

Bachelor degree

1. Student data from the Thepsatri Rajabhat University, Lopburee province, contains 10,980 records.
2. Student data from the Phetburi Rajabhat University has 10,466 records.
3. Student data from the Ubon Rajathani Rajabhat University has information of 5,170 records.
4. Student data from the Sripatum University contains 308 records.
5. Student data from the Phanakorn Rajabhat University contains 1,124 records.

Master degree

1. Student data from the Sripatum University has 305 records.

An example of the student records are shown in Table2.

Table 2: *An example of student records*

Id	sex	age	...	GPA1	GPA2	Finish
1	f	23	...	2.78	2.40	1
2	f	22	...	3.04	2.72	1
3	f	23	...	2.71	2.21	1
4	m	22	...	3.70	3.50	0
..	....	...	...	...	...	...

4. EXPERIMENTAL RESULTS

We run experiments to evaluate the prediction models using 10-fold cross-validation which is a standard performance evaluation method for machine learning algorithms [9].In 10-fold cross-validation, the student data was partitioned into ten disjoint subsets. Each subset was used as a test set once, and the remaining subsets were used as the training set.

For vocational student data, the Bayes net technique discovered the relationship among variables as shown in Fig.2.

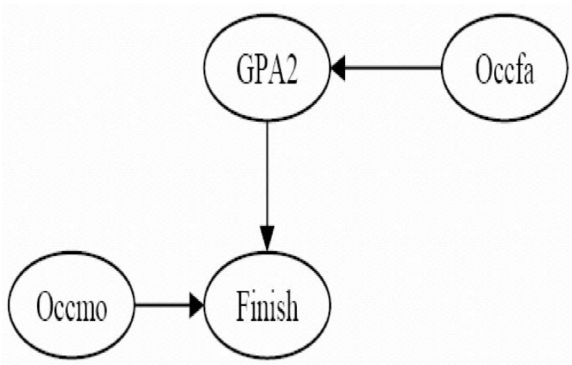

Fig.2: *Prediction Model of Student Graduation for Vocational Level*

For student data of the Ubon Rajathani Rajabhat University with 5,170 records, using WEKA, we found that the important variables which affect the graduation were the grade point average when they entered the first year, mothers career and total income of the family as shown in Fig. 3.

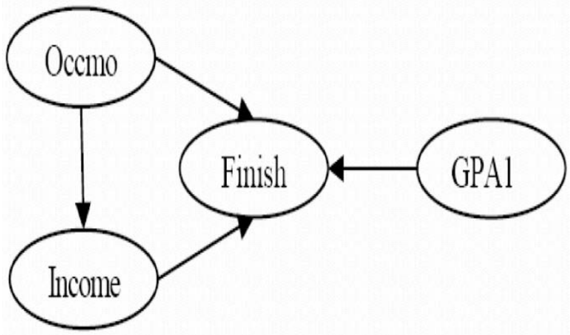

Fig.3: *Prediction Model of Student Graduation for the Ubon Rajathani Rajabhat University*

For student data of the Sripatum University in Undergraduate level containing 308 records, we found the relationship among variables as shown in Fig. 4.

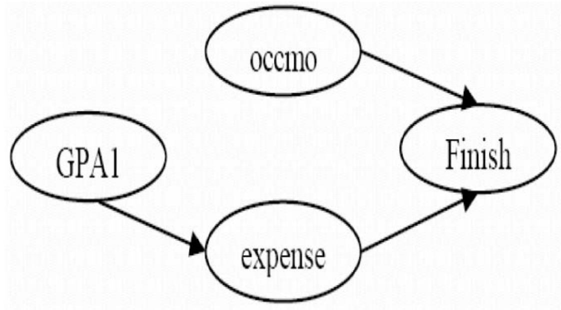

Fig.4: *Prediction Model of Student Graduation for Sripatum University in Undergraduate Level*

The model constructed by WEKA for student data of the Phanakorn Rajabhat University containing 1,124 records is shown in Fig 5.

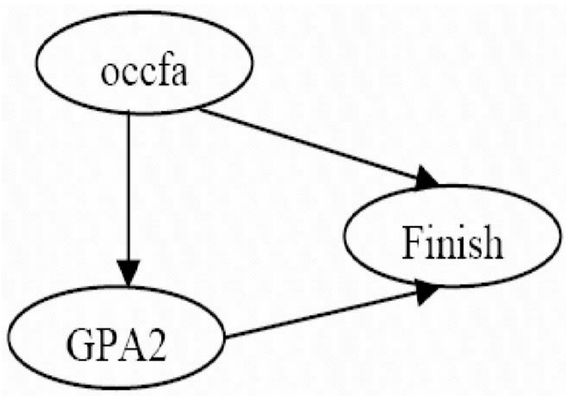

Fig.5: *Prediction Model of Student Graduation for Phanakorn Rajabhat University*

The model obtained from student data of the Thepsatri Rajabhat University is shown in Fig 6.

The model obtained from student data of the Phetburi Rajabhat University is shown in Fig 7.

Finally, the model obtained from student data of the Sripatum University in graduate level has variable relationship as shown in Fig. 8.

In the Bayes net, there is a conditional probability table (CPT) attached to each node, such as one shown in Table 3. CPT in the table is for the finish node, and the first row of CPT indicates that if a student has the family whose income is below 6,000 baths per month (shown by code I801), grade point average in the first semester of the student is below 1.976 and the students mother has agricultural occupation, then the student will graduate with probability equal to 0.007.

As the result of model testing for the Thai administration and commerce school with vocational level, the obtained accuracy is 93.25% as shown in Fig.9.

The model constructed for the Ubonrajathani Rajabhat with undergraduate level has accuracy of 91.26%as shown in Fig.10.

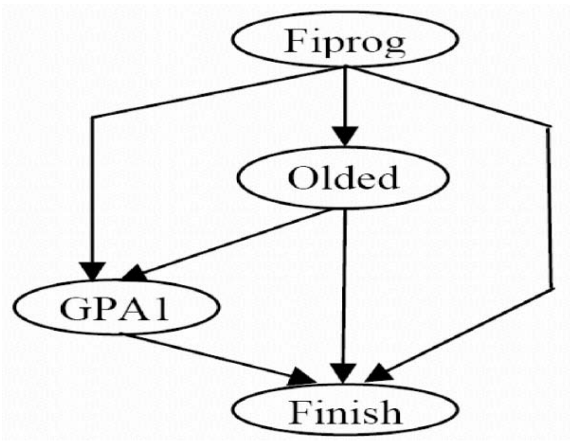

Fig.6: *Prediction Model of Student Graduation for Thepsatri Rajabhat University*

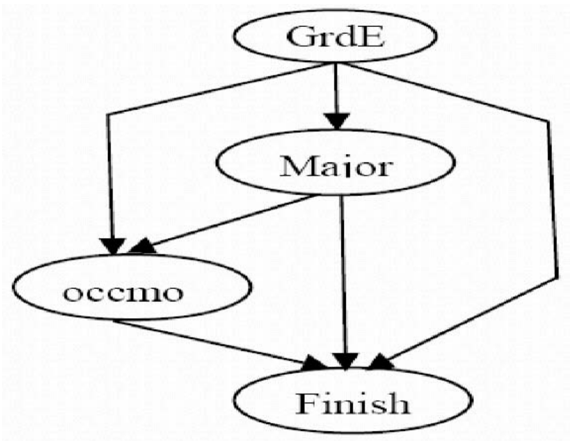

Fig.7: *Prediction Model of Student Graduation for Phetburi Rajabhat University*

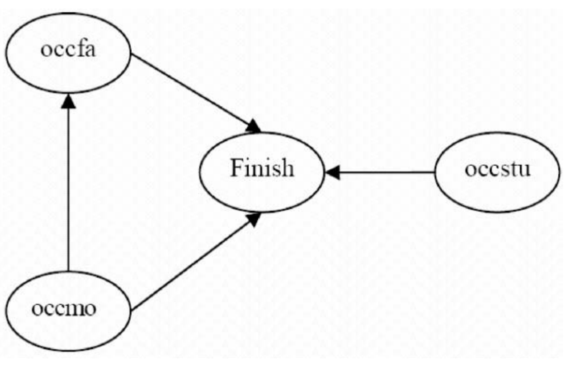

Fig.8: *Prediction Model of Student Graduation for Sripatum University in graduate level*

Table 3: *An Example of Conditional Probability Table*

Probability Distribution Table For finish

income total	first grade	occ mo	no	yes
I801	'(-inf-1.976]'	O604	0.993	0.007
I801	'(-inf-1.976]'	O602	0.5	0.5
I801	'(-inf-1.976]'	O601	0.5	0.5
I801	'(-inf-1.976]'	O606	0.958	0.042
I801	'(-inf-1.976]'	O603	0.5	0.5
I801	'(-inf-1.976]'	O605	0.5	0.5
I801	'(1.976-2.482]'	O604	0.48	0.52
I801	'(1.976-2.482]'	O602	0.828	0.172
I801	'(1.976-2.482]'	O601	0.5	0.5
I801	'(1.976-2.482]'	O606	0.662	0.338
I801	'(1.976-2.482]'	O603	0.5	0.5
I801	'(1.976-2.482]'	O605	0.5	0.5
I801	'(2.482-2.988]'	O604	0.085	0.915
I801	'(2.482-2.988]'	O602	0.253	0.747
I801	'(2.482-2.988]'	O601	0.5	0.5
I801	'(2.482-2.988]'	O606	0.005	0.995
I801	'(2.482-2.988]'	O603	0.5	0.5
I801	'(2.482-2.988]'	O605	0.5	0.5
I801	'(2.988-3.494]'	O604	0.026	0.974
I801	'(2.988-3.494]'	O602	0.022	0.978
I801	'(2.988-3.494]'	O601	0.042	0.958
I801	'(2.988-3.494]'	O606	0.009	0.991

=== Run information ===
Scheme: weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TabuSearch -- -L 5
-U 10 -R -N -P 3 -S BAYES -E
Instances: 408
Attributes: 5
Test mode: 10-fold cross-validation
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	304	93.2515 %
Incorrectly Classified Instances	22	6.7485 %
Kappa statistic	0.9313	
Mean absolute error	0.016	
Root mean squared error	0.0728	
Relative absolute error	57.7039 %	
Root relative squared error	61.8305 %	
Total Number of Instances	326	

Fig.9: *Result of Model Testing for Thai administration and commerce school*

= == Run information ===
Scheme: weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net. -
Instances: 5170
Test mode: 10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier not using ADTree
=== Summary ===

Correctly Classified Instances	4718	91.2573 %
Incorrectly Classified Instances	52	8.7427 %
Total Number of Instances	5170	

Fig.10: *Result of Model Testing for Ubonrajathani Rajabhat*

For the Sripatum University with undergraduate level, the model has accuracy of 94.16%as shown in Fig.11.

===Run information ===
Scheme:weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TabuSearch -- -L 5
-U 10 -R -N -P 3 -S MDL -E
Instances: 308
Attributes: 8
Test mode: 10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier
not using ADTree
Test mode: 10-fold cross-validation
=== Stratified cross-validation ===

Correctly Classified Instances	290	94.1558 %
Incorrectly Classified Instances	18	5.8442 %
Kappa statistic	0.7922	
Mean absolute error	0.0965	
Root mean squared error	0.1969	
Relative absolute error	31.8446 %	
Root relative squared error	50.7054 %	
Total Number of Instances	308	

Fig.11: *Result of Model Testing for Sripatum University with undergraduate level occmo Finish occstu*

For the Phanakorn Rajabhat University which has student database with 1,124 records, the obtained accuracy is 94.13%as shown in Fig.12.

The model constructed for the Thepsatri Rajabhat University has 77.97% accuracy as shown in Fig.13.

The model constructed for the Phetburi Rajabhat University has 78.12%accuracy as shown in Fig.14.

Finally, the model for the Sripatum University has accuracy of 98.53%as shown in Fig.15.

Next, the important or significant variables that affect student graduation are compared to multiple regression analysis as described in the following section.

=== Run information ===
Scheme: weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TabuSearch -- -L 5
-U 10 -R -N -P 3 -S BAYES -E
Weka.classifiers.bayes.net.estimate.SimpleEstimator -- -
A 0.5
Instances: 1124
Attributes: 6
Test mode: 10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier
not using ADTree
#attributes=6 #classindex=5
Time taken to build model: 5.91 seconds
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	1058	94.1281 %
Incorrectly Classified Instances	66	5.8719 %
Kappa statistic	0.2528	
Mean absolute error	0.1107	
Root mean squared error	0.2279	
Relative absolute error	85.2496 %	
Root relative squared error	89.6716 %	
Total Number of Instances	1124	

Fig.12: *Result of Model Testing for Phanakorn Rajabhat University*

Table 4: *Model Summary*

Probability Distribution Table For finish

income total	first grade	occ mo	no	yes
I801	'(-inf-1.976]'	O604	0.993	0.007
I801	'(-inf-1.976]'	O602	0.5	0.5
I801	'(-inf-1.976]'	O601	0.5	0.5
I801	'(-inf-1.976]'	O606	0.958	0.042
I801	'(-inf-1.976]'	O603	0.5	0.5
I801	'(-inf-1.976]'	O605	0.5	0.5
I801	'(1.976-2.482]'	O604	0.48	0.52
I801	'(1.976-2.482]'	O602	0.828	0.172
I801	'(1.976-2.482]'	O601	0.5	0.5
I801	'(1.976-2.482]'	O606	0.662	0.338
I801	'(1.976-2.482]'	O603	0.5	0.5
I801	'(1.976-2.482]'	O605	0.5	0.5
I801	'(2.482-2.988]'	O604	0.085	0.915
I801	'(2.482-2.988]'	O602	0.253	0.747
I801	'(2.482-2.988]'	O601	0.5	0.5
I801	'(2.482-2.988]'	O606	0.005	0.995
I801	'(2.482-2.988]'	O603	0.5	0.5
I801	'(2.482-2.988]'	O605	0.5	0.5
I801	'(2.988-3.494]'	O604	0.026	0.974
I801	'(2.988-3.494]'	O602	0.022	0.978
I801	'(2.988-3.494]'	O601	0.042	0.958
I801	'(2.988-3.494]'	O606	0.009	0.991

5. STATISTICAL ANALYSIS

We tested the constructed models by multiple regression analysis, a technique for statistical analysis as shown in Table 4.

As shown in Table 4, the coefficient of determination (R Square) is 38.10 %. This indicates that the total income of the family (income-total), occupation of students mother (occmo), and GPA in the first semester (first-grade) can explain the variation of the dependent variable by 38.10%.

We then consider another statistical analysis technique, i.e. ANOVA (see Table 5).

As shown in the table, the significant value is less than 0.001, indicating that all the independent variables in the model (income-total, occmo and first-grade) have directly affected the dependent variable. The relationships of each variable can be seen in Table 6.

From the result in Table 6, we can conclude that the independent variables are significant; the independent variables in the model have the relationship to the dependent variable. Therefore, the result of statistical analysis confirms that the independent variables used in the model constructed by the Bayes net are reliable for the prediction of the dependent variable.

6. CONCLUSION

In this paper, we have demonstrated that the model constructed by the Bayes net technique is reliable with high prediction accuracy. We then further examined the variables occurring in the model

```
=== Run information ===
Scheme:    weka.classifiers.bayes.BayesNet   -D   -Q
weka.classifiers.bayes.net.search.local.TabuSearch  --  -L
10 -U 30 -R -N -P 3 -S ENTROPY –E
Instances:  10979
Attributes:  5
Test mode:  10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier
not using ADTree
#attributes=5 #classindex=4
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      8560      77.967  %
Incorrectly Classified Instances    2419      22.033  %
Kappa statistic                     0.5091
Mean absolute error                 0.2845
Root mean squared error             0.3912
Relative absolute error             60.6794 %
Root relative squared error         80.7991 %
Total Number of Instances           10979
```

Fig.13: *Result of Model Testing for Thepsatri Rajabhat University*

```
=== Run information ===
Scheme:    weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TabuSearch -- -L 5
-U 10 -R -N -P 3 -S ENTROPY -E
Instances:  10465
Attributes:  8
Test mode:  10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier
not using ADTree
#attributes=8 #classindex=7
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      8238      78.7195 %
Incorrectly Classified Instances    2227      21.2805 %
Kappa statistic                     0.2898
Mean absolute error                 0.3012
Root mean squared error             0.3899
Relative absolute error             81.8033 %
Root relative squared error         90.8825 %
Total Number of Instances           10465
```

Fig.14: *Result of Model Testing for Phetburi Rajabhat University*

Table 5: *Anova* [b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.617[a]	.381	.381	.310

a. Predictors: (constant), income_total, occmo, first_grade

```
=== Run information ===
Scheme:    weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TabuSearch -- -L 5
-U 10 -R -N -P 3 -S BDeu –E
Instances:  305
Attributes:  6
Test mode:  10-fold cross-validation
=== Classifier model (full training set) ===
Bayes Network Classifier
=== Summary ===
Correctly Classified Instances      268      98.5294 %
Incorrectly Classified Instances      4       1.4706 %
Kappa statistic                     0.9846
Mean absolute error                 0.0112
Root mean squared error             0.0411
Relative absolute error             20.3276 %
Root relative squared error         24.8471 %
Total Number of Instances           272
```

Fig.15: *Result of Model Testing for Sripatum University in Graduate Level*

Table 6: *Coefficients* [a]

Model	Unstandardized coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	-4.347	.555		-7.829	.000
Occmo	.006	.001	.076	6.477	.000
First_grade	.521	.010	.614	52.460	.000
Income_total	.000	.000	.060	5.127	.000

a. Dependent Variable: finish

by multiple regression analysis, and found that the independent variables in the model can be used to predict the value of the dependent variable. This statistical analysis confirmed that the variables discovered by the Bayes net are reliable for the prediction of the student graduation.

References

[1] J. Han and M. Kamber, *Data Mining Concepts and Techniques*, Morgan Kaufmann Publishers, 2001.

[2] G. J. Hendricks, *An Analysis of Student Graduation Trends in Texas State Technical Colleges Utilizing Data Mining and Other Statistical Techniques*, Doctoral Dissertation of Educational Administration, Baylor University, Texas, U.S.A., 2000.

[3] U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, 1996.

[4] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations* , Morgan Kaufmann Publishers, New York, 2000.

[5] R. R. Bouckaert, *Bayesian Network Classifiers in Weka* , Department of Computing Science, University of Waikato, New Zealand, 2005.

[6] G. Piatetsky-Shapiro and W. J. Frawley, *Knowledge Discovery in Databases* , MIT Press, 1991.

[7] B. Kijsirikul,*Artificial Intelligence* , Department of Comuter Engineering, Faculty of Engineering, Chulalongkorn University, 2003 (in Thai).

[8] K. Waiyamai, T. Rakthanmanon and C. Songsiri, "Data Mining Techniques for Developing Education in Engineering Faculty," *NECTEC Technical Journal* , vol. III, no. 11., pp. 134-142, 2001.

[9] J. Polpinij, *The Probabilistic Models Approach for Analysis the Factors Affecting of Car Insurance Risk* , Master Thesis, Department of Computer Science, Kasetsart University, 2002.

[10] B. Paulo and J. M. Silva, *Mining On-line Newspaper Web Access logs* , Departamento de Informatica Faculdade de Ciencias, Universidade de Lisboa, 2001.

[11] R. Kirkby and E. Frank, *Weka Explorer User Guide* , University of Waikato, New Zealand, 2005.

[12] J. R. Roiger and W. M. Geatz, *Data Mining: A Tutorial based Primer* , Addison Wesley Publishing Company, 2003.

[13] M. S. Babaoye, *Input and Environmental Characteristics in Student Success: First-Term GPA and Predicting Retention at an Historically Black University* , Ph.D. Dissertation, Louisiana University and the Agricultural & Mechanical College, 2000.

[14] D. Olson and Y. Shi, *Introduction to Business Data Mining* , McGraw-Hill, International Edition, 2007.

[15] S. Mitra and T. Acharya,*Data Mining Multimedia, Soft Computing, and Bioinformatics*, John Wiley & Sons, Inc., New Jersey. 2003.

[16] M. Kantardzic,*Data Mining Concepts, Models, Methods, and Algorithms* , John Wiley & Sons, Inc., 2003.

[17] A. Berson and J. S. Smith, *Data Warehousing, Data Mining, and OLAP* , McGraw-Hill, Inc., 1997.

[18] M. M. Koker, *A Predictive Model of Bachelor's Degree Completion for Transfer Students at an Urban, Research I University* , Ph.D Dissertation, University of Minnesota, U.S.A., 2000.

[19] H. Almuallin and T. G. Dietterich, "Efficient Algorithms for Identifying Relevant Features," *In Proceedings of the Ninth Canadian Conference on Artificial Intelligence* , Vancouver, BC. San Francisco: Morgan Kaufmann, pp. 38-45. 1992.

[20] R. R. Bouckaert, *Bayesian Belief Networks: From Construction to Inference* , Ph.D Dissertation, Computer Science Department, University of Utrecht, The Netherlands, 1995.

[21] R. R. Bouckaert, "Bayesian Network Classifiers in Weka," Working Paper 14/2004, Department of Computer Science, University of Waikato, New Zealand, 2004.

[22] P.Cheeseman, and J. Stutz,*Bayesian Classification (AutoClass): Theory and Results* , In U. M. Fayyad, G. Piatetsky-Shapiro, P Smyth, and R. Uthurusamy, editors, Advances in Knowledge Discovery and Data Mining. Menlo Park, CA: AAAI Press, 1995.

[23] R. Kohavi, "A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection," *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* , Montreal, Canada, San Francisco: Morgan Kaufmann, pp. 1137-1143, 1995a.

[24] P. Langley, W. Iba, and K. Thompson, "An Analysis of Bayesian Classifiers," In W. Swartout, editor, *Proceedings of the Tenth National Conference on Artificial Intelligence* , San Jose, CA. Menlo Park, CA: AAAI Press, pp. 223-228, 1992.

[25] P. Langley, and S. Sage, "Induction of Selective Bayesian Classifiers,"In R. L. de Mantaras and D. Poole, editors, *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence* , Seattle, WA. San Francisco: Morgan Kaufmann, pp. 399-406, 1994.

[26] Y.Wang, and I. H. Witten, "Modeling for Optimal Probability Prediction," In C. Sammut and [28] A.Hoffmann, editors, *Proceedings of the Nineteenth International Conference on Machine Learning* , Sydney, Australia. San Francisco: Morgan Kaufmann, pp. 650-65, 2002.

[27] Yang, Y., and G. I. Webb,"Proportional k-interval Discretization for Naive Bayes Classifiers," In L. de Raedt and P Flach, editors, *Proceedings of the Twelfth European Conference on Machine Learning* , Freiburg, Germany, Berlin: Springer-Verlag, pp. 564-575, 2001.

[28] D. Heckerman, D. Geiger, and D. M. "Chickering, Learning Bayesian Networks: The Combination of Knowledge and Statistical Data," Machine Learning 20 (3):197-243, 1995.

[29] Zheng, Z., and G. Webb, "Lazy Learning of Bayesian Rules," Machine Learning 41(1):53-84, 2000.

[30] Devroye, L., L. Gyrfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*, New York: Springer-Verlag, 1996.

[31] D.Grossman, and P. Domingos,"Learning Bayesian Network Classifiers by Maximizing Conditional Likelihood," In R. Greiner and D. Schuurmans, editors, *Proceedings of the Twenty-First International Conference on Machine Learning*, Banff, Alberta, Canada. New York: ACM, pp. 361-368, 2004.

[32] P.Adriaans, and D. Zantige,*Data Mining*, Harlow, England: Addison Wesley, 1996.

[33] M. J. A. Berry, and G. Linoff,*Data Mining Tech-*

niques for Marketing, Sales, and Customer Support, New York: John Wiley, 1997.

[34] J. P. Bigus,*Data Mining with Neural Networks* , New York: McGraw Hill, 1996.

[35] S. M.Weiss, and N. Indurkhya, *Predictive Data Mining: A Practical Guide* , San Francisco: Morgan Kaufmann, 1998.

[36] D. Pyle,*Data preparation for Data Mining* , San Francisco: Morgan Kaufmann.

[37] Quinlan, J. R. 1986.*Induction of Decision Trees* , Machine Learning 1(1):81-106. 1999.

[38] John, G. H., and P Langley, "Estimating Continuous Distributions in Bayesian Classifiers," In P. Besnard and S. Hanks, editors, *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* , Montreal, Canada. San Francisco: Morgan Kaufmann, pp. 338-345. 1995.

[39] R. Groth, *Data Mining: A Hands-on Approach for Business Professionals* , Upper Saddle River, NJ: Prentice Hall, 1998.

[40] D. J.Hand, H. Mannila, and P Smyth, *Principles of Data Mining* , Cambridge, MA: MIT Press, 2001.

[41] T. Hastie, R. Tibshirani, and J. Friedman.*The Elements of Statistical Learning* , New York: Springer-Verlag, 2001.

[42] G. H. John, *Enhancements to the Data Mining Process* , Ph.D Dissertation, Computer Science Department, Stanford University, Stanford, CA. 1997.

[43] Z. Zheng and G. Webb, "Lazy learning of Bayesian rules," Machine Learning 41(1):53-84, 2000.

[44] C. J. Wild and G. A. F. Seber,*Introduction to Probability and Statistics* , Department of Statistics, University of Auckland, New Zealand, 1995.

[45] V.Vapnik, *The Nature of Statistical Learning Theory* , second edition. New York: Springer-Verlag. Wang, Y., and I. H. 1999.

[46] J. Swets, "Measuring the Accuracy of Diagnostic Systems," Science, 240:1285-1293. 1988.

[47] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz, "A Bayesian Approach to Filtering Junk e-Mail," *Proceedings of the AAAI-98 Workshop on Learning for Text Categorization*, Madison, WI. Menlo Park, CA: AAAI Press, pp. 55-62, 1998.

[48] N. H. Nie, C. H.Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, *Statistical Package for the Social Sciences* , New York: McGraw Hill, 1970.

[49] R. Kass, and L. Wasserman,"A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwarz Criterion,"*Journal of the American Statistical Association* , 90:928-934. 1995.

[50] M. V. Johns, *An Empirical Bayes Approach to Nonparametric Two-way Classification*, H. Solomon, editor, Studies in item analysis and prediction, Palo Alto, CA: Stanford University Press, 1961.

[51] G. H. John, and P Langley, "Estimating Continuous Distributions in Bayesian Classifiers," In P Besnard and S. Hanks, editors,*Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* , Montreal, Canada. San Francisco: Morgan Kaufmann, pp. 338-345. 1995.

[52] D. E. Appelt, Introduction to Information Extraction Technology, Tutorial, *Int. Joint Conf. on Artificial Intelligence IJCAI99* . Morgan Kaufmann, San Mateo, 1999. Tutorial notes available at www.ai.sri.com/ appelt/ie-tutorial.

[53] R. Agrawal and R. Srikant,"Fast Algorithms for Mining Association Rules in Large Databases," In J. Bocca, M. Jarke, and C. Zaniolo, editors, *Proceedings of the International Conference on Very Large Databases*, Santiago, Chile. San Francisco: Morgan Kaufmann, pp. 478-499. 1994.

[54] R. Agrawal, T. Imielinski and A. Swami, Database Mining: A Performance Perspective, *IEEE Transactions on Knowledge and Data Engineering* 5(6): 914-925. 1993.

Jiraporn Yingkuachat received the B.A. in Business Computer in 2001 from school of Informatics, Sripatum University, the Master of Science in Information Technology in 2006, Information Technology Program, Graduate School, Sripatum University, Jatujak, Bangkok, Thailand. Her research interests are in the areas of Information Technology which focus on Data Mining Applications in Education and other Business.

Prasong Praneetpolgrangreceived the B.Sc.(1st Hons) in Electrical Engineering from the Royal Thai Air Force Academy, Bangkok, THAILAND, in 1987, the M.S. in Computer Engineering, 1989, the M.S. in Electrical Engineering, 1993, and the Ph.D degree in Computer Engineering from Florida Institute of Technology, Florida, USA, in 1994. He currently has the rank of associate professor and on the Director of Information Technology program at Sripatum University, Bangkok, Thailand. His research interests are in the areas of Information Technology Management, Information Security, Knowledge Management, and e-Learning. Dr. Prasong Praneetpolgrang has several published articles in these areas. He has served on program committees of both international and national conference on Computer Science and Engineering, Information Technology and e-Business such as served on program committee for the international conference on Electronic Business (INCEB), during 2002-Present. He is also a member of ACM, and IEEE. He has recorded in Whos Who in the world in Information Technology.

Boonserm Kijsirikul received the B.Eng.in Electronic and Electrical Engineering in 1988,the M.Eng.in Computer Engineering in 1990 and the D.Eng.in Computer Engineering in 1993,from Tokyo Institute of Technology, Japan.He is currently working as associate professor at Chulalongkorn University,Bangkok,Thailand.His research interests include Machine Learning, Artificial Intelligence,Speech Processing,etc.He is a member of AAAI.