

**PREDICTION OF UNIVERSITY ADMISSION ACHIEVEMENT  
WITH DISCRIMINANT MODEL: CASE STUDY OF  
PHANATPITTAYAKARN SCHOOL, CHONBURI, THAILAND**

Kidakan Saithanu<sup>1</sup>, Manunya Jeeradid<sup>2</sup>, Jatupat Mekpariyup<sup>3</sup> §

<sup>1,2,3</sup>Department of Mathematics

Burapha University

169, Tambon Saensook, Amphur Muang, Chonburi, 20131, THAILAND

**Abstract:** This study was conducted to classify two achievement groups of university admission (failure and success). High school students studied in Phanatpittayakarn School during the academic year 2006-2011 participated. Gender, high school program, each of eight O-NET scores (Thai, Social, English, Mathematics, Science, Hygiene, Art and Work) as well as GPA of each subject (Social, English, Mathematics, Science, Hygiene, Art, Work) and high school GPA were investigated as predictors of university admission achievement. Discriminant analysis was performed to build a predictive model for achievement of university admission. The correct classification rate (CCR) is utilized to evaluate a model performance. The result presents the discriminant classification function obtained rather highly accurate of the average CCR (80.79%) for prediction of university admission achievement in the validation data set.

**AMS Subject Classification:** 62H30

**Key Words:** university admission system, O-NET, discriminant analysis

## 1. Introduction

Education in Thailand is mainly provided by the Ministry of Education. The constitution certifies 12 years of free basic education. The formal education is divided into 4 stages. The first level (Prathomsuksa 1-3) is 3 years in the elementary school. The

---

Received: July 19, 2013

© 2014 Academic Publications, Ltd.

§Correspondence author

second level (Prathomsuksa 4-6) is another 3 years from grade 4 to 6. The third level (Mattayomsuksa 1-3) includes the first three year in high school. The last level (Mattayomsuksa 4-6) contains the last three year in upper level high school. Attending 6 years in the elementary school and at least the first 3 years in high school are obliged. Thai students are also required to take the National Education Test (NET) after each level. The Ordinary National Education Test (O-NET) and the Advanced National Education Test (A-NET) are also needed to be taken to continue tertiary education. Every student has to pass the Central University Admission System (CUAS) before getting into any university. Initially, the CUAS composed the results of O-NET and A-NET (50%) and grade point average (GPA) from the fourth level (50%) [1]. According to this system, it depended on the standard of school so it led to many arguments. Later, university admission system still have been developed and changed all the time according to time and situation, for example, period of the test to be held or percentage of each test to be measured. Presently, the criterion of university admission consists four components: GPA from the fourth level (20%), the eight O-NET scores (30%), the General Aptitude Test (GAT: 10-50%) and the Professional Aptitude Test (PAT: 0-40%) [2]. Due to the criterion of university admission is uncertainly at present, few papers in Thailand involved finding of influential factors caused to getting the low O-NET score [3], [4]. However, many researchers developed and validated model to predict student achievement in the university rather than in the high school [5], [6], [7], [8], [9]. While some papers applied statistical method including discriminant analysis for model building to predict the student achievement [8], [9], [10], [11].

The purpose of this study is to build discriminant model for prediction of university achievement based on eight O-NET scores, GPA of eight subjects and the high school GPA using case study of Phanatpittayakarn School which is located in Chonburi province.

## 2. Data Description

The observational data analyzed here was collected from the 1,713 high school students of Phanatpittayakarn School, Chonburi, Thailand during the academic year 2006-2011. The achievement of university admission was considered as the dependent variable which was categorized into two groups: failure and success. The following 19 independent variables were as follows: Gender, High School Program (Science-Math, Math-Art and Language-Art), Thai O-NET Score (ONet1), Social O-NET Score (ONet2), English O-NET Score (ONet3), Mathematics O-NET Score (ONet4), Science O-NET Score (ONet5), Hygiene O-NET Score (ONet6), Art O-NET Score (ONet7), Work O-NET Score (ONet8), Thai GPA (Thai), Social GPA (Social), English GPA (Eng), Mathematics GPA (Math), Science GPA (Sci), Hygiene GPA (Hyg), Art GPA (Art), Work GPA (Work) and High School GPA (GPAX)

[12].

For discriminant analysis, two mutually exclusive and distinct data sets were firstly created to train and validate. The training data set composed of 1,260 high school students of Phanatpittayakarn School during the academic year 2006-2010 using for model training. The validation data set contained 453 high school students of Phanatpittayakarn School during the academic year 2009-2011 utilizing to validate the suitability of a data model.

### 3. Methodology

#### 3.1. Investigating of Variable Association

Chi-square test was obtained to measure degree of relationship between the achievement of university admission and each of 2 categorical independent variables (Gender and High School Program). In addition, degree of the linear relationship between GPAX and each of 16 quantity independent variables (ONet1, ONet2, ONet3, ONet4, ONet5, ONet6, ONet7, ONet8, Thai, Social, Eng, Math, Sci, Hyg, Art and Work) were tested by the correlation analysis.

If any of P-value of chi-square test or P-value of Pearson correlation coefficient is less than the chosen significance level, then there is a significant association between these two variables.

#### 3.2. Discriminant Analysis

Discriminant analysis is a multivariate classification technique concerned with distinguishing separate sets of observations and with designating observations to previously defined groups. The principle of this analysis is to find the combination of variables that best predicts the class or group to which an observation belongs. A classification function is the combination of predictor variables which can be used to categorize new observations whose group membership is unknown. Two discriminant scores are criteria usually employed of making analysis. Linear discriminant score is assumed equality of covariance matrix of independent variables for all groups. An estimate of linear discriminant score is defined as of Equation 1 [13].

$$\hat{d}_i(\mathbf{x}) = \bar{\mathbf{x}}_i' \mathbf{S}_{pooled}^{-1} \mathbf{x}_0 - \frac{1}{2} \bar{\mathbf{x}}_i' \mathbf{S}_{pooled}^{-1} \mu_i + \ln(p_i), \quad (1)$$

where  $\bar{\mathbf{x}}_i$  is the vector of mean for group  $i$ ;  $i = 1, 2, \dots, g$ ,  $\mathbf{S}_{pooled}$  is the pooled variance matrix and  $p_i$  is the prior probability. Then  $\mathbf{x}$  would be assigned to population  $\pi_k$  if  $\hat{d}_i(\mathbf{x})$  is the largest of  $\hat{d}_1(\mathbf{x}), \hat{d}_2(\mathbf{x}), \dots, \hat{d}_g(\mathbf{x})$ ;  $k \neq i$ .

Contradictory, quadratic discriminant score does not rely on this assumption.

An estimate of quadratic discriminant score is defined as of Equation 2 [13].

$$\hat{d}_i^Q(\mathbf{x}) = -\frac{1}{2} \ln(|\mathbf{S}_i|) - \frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}}_i)' \sum_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i) + \ln(p_i), \quad (2)$$

where  $\mathbf{S}_i$  is the covariance matrix for group  $i$ ;  $i = 1, 2, \dots, g$ . Then  $\mathbf{x}$  would be assigned to population  $\pi_k$  if  $\hat{d}_i^Q(\mathbf{x})$  is the largest of  $\hat{d}_1^Q(\mathbf{x}), \hat{d}_2^Q(\mathbf{x}), \dots, \hat{d}_g^Q(\mathbf{x})$ ;  $k \neq j$ .

To determine using linear or quadratic discriminant score, Box's M test statistic is used for homogeneity of covariance matrices. If P-value of Box's M test is larger than the chosen significance level, the linear discriminant score is then utilized.

### 3.3. Performance Criterion

To verify how well the discriminant model would predict the achievement of university admission in the validation data set, a correct classification rate (CCR) is the most one frequently used in any classification procedure [14]. It is defined as of Equation 3.

$$CCR = \frac{\sum_{k=0}^{C-1} CC_k}{n}, \quad (3)$$

where  $CC_k$  is the number of correctly classified observations and  $n$  is the number of observation in the group. The model with the largest correct classification rate is the one with better performance.

## 4. Results

### 4.1. Determination of Important Variables Associated to the Achievement of University Admission

Since the large P-value of chi-square test between the achievement of university admission and gender (P-value=0.096) led to conclude that the achievement of university admission was not depend on gender. Each of the Pearson correlation coefficients (between GPAX and each of 16 quantity independent variables (ONet1, ONet2, ONet3, ONet4, ONet5, ONet6, ONet7, ONet8, Thai, Social, Eng, Math, Sci, Hyg, Art and Work)) obtained positive value as well with P-value really closed to zero. It led to conclude that there were positive significant association between GPAX and each of these variables. Thus, the achievement of university admission could be predicted with almost independent variables (18 variables) except only gender.

#### 4.2. Discriminant Model

The linear discriminant technique was applied since the covariance matrices of these independent variables between the two achievement groups were equal as seeing of the large P-value of Box's M statistic (0.069). Once the 18 significant variables were evaluated by the discriminant analysis, the linear discriminant classification function for each of achievement groups was obtained as follows:

**For failure group:**

$$\begin{aligned} \hat{Y}_0 = & -329.24 + 0.16ONet1 + 0.46ONet2 + 0.02ONet3 - 0.03ONet4 + 0.12ONet5 \\ & + 0.81ONet6 + 0.70ONet7 - 0.02ONet8 - 11.45Thai - 16.53Social - 1.22Eng \\ & - 4.92Math - 4.32Sci + 58.01Hyg + 106.90Art + 20.20Work - 30.20Plan1 \\ & - 30.14Plan2 - 4.14GPAX. \end{aligned}$$

**For success group:**

$$\begin{aligned} \hat{Y}_1 = & -331.55 + 0.14ONet1 + 0.50ONet2 - 0.02ONet3 - 0.05ONet4 + 0.13ONet5 \\ & + 0.79ONet6 + 0.67ONet7 - 0.03ONet8 - 12.67Thai - 15.81Social - 0.05Eng \\ & - 4.68Math - 4.20Sci + 57.23Hyg + 107.44Art + 21.00Work - 28.89Plan1 \\ & - 28.38Plan2 - 3.72GPAX, \end{aligned}$$

where any student in Science-Math program was replaced with  $Plan1 = 1$  and  $Plan2 = 0$ , in Math-Art program was replaced with  $Plan1 = 0$  and  $Plan2 = 1$ , and in Language-Art program was replaced with  $Plan1 = 0$  and  $Plan2 = 0$ .

#### 4.3. Performance of Discriminant Model

The predictive discriminant model displayed the CCR and the average CCR for both of training and validation data sets as of Table 1.

### 5. Conclusion and Discussion

As of applying discriminant analysis, it shows that high school program, eight O-NET scores, eight subjects of GPA are the influential factors to be used for making a prediction of university admission achievement. The group achievement was correctly predicted with rather large value of the average CCR (0.807947) in the validation data set. Also, the predictive discriminant model was able to classify students into the success group more than the failure group for both of training and validation data sets with the CCR 0.796740 and 0.809735, respectively. An academic affair of Phanatpittayakarn School can use these findings to predict the achievement

<i>Training Data Set</i>			
Put into Group	True Group		Average of CCR
	Failure	Success	
Failure	155	212	0.782540
Success	62	831	
Total of achievement group	217	1,043	
CCR	0.714286	0.796740	
<i>Validation Data Set</i>			
Put into Group	True Group		Average of CCR
	Failure	Success	
Failure	0	86	0.807947
Success	1	366	
Total of achievement group	1	452	
CCR	0	0.809735	

Table 1: Classification results of discriminant model

of university admission of students who will be admitted to the university in the next academic year. If some students are predicted to be in the failure group, the administrators of an academic affair of Phanatpittayakarn School should arrange tutorial classes for these students.

### Acknowledgments

The authors are also grateful to the academic affair of Phanatpittayakarn School, Chonburi, Thailand for kindly providing all data.

### References

- [1] International study and exchange, *Education and study in Thailand*, Retrieved April 20, 2012, from <http://www.spainexchange.com/schools/study-thailand/>
- [2] Association of University Presidents of Thailand, *Association of University Presidents of Thailand, Central University Admissions System (CUAS)*, Retrieved April 20, 2012, from <http://www.cuas.or.th/info.html/>
- [3] National Institute of Education Testing Service, *A Study of Causes of O-NET Low Achievement in Grade 12 School Student: Case Study of Suksawad Wittaya School*, Retrieved April 20, 2012, from [http://www.niets.or.th/index.php/research\\_th/view/15/](http://www.niets.or.th/index.php/research_th/view/15/)

- [4] National Institute of Education Testing Service, *A Causal factor affecting to the low O-NET test scores of students in grade 6 and grade 12*, Retrieved April 20, 2012, from [http://www.niets.or.th/index.php/research\\_th/view/15/](http://www.niets.or.th/index.php/research_th/view/15/)
- [5] M. Krungsanmuang, *A prediction of learning achievement of Chulalongkorn university students admitted through the central university admission system*, Master's thesis, Department of Educational Research and Psychology, Faculty of Education, Chulalongkorn University (2006).
- [6] D. Sanguanrungsirikul, *A Prediction of Learning Achievement of King Mongkut's University of Technology Thonburi Students Admitted through Central University Admission System*, *KMUTT Research and Development Journal*, **34(4)** (2011), 431-442.
- [7] M. Pinyopornpanish, L. Wongsawasdi, N. Panjaisee, V. Buayuen, Comparison of the academic achievement of Chiang Mai graduate medical students which selected by quota, entrance and rural project, *Chiang Mai Med Bull*, **43(2)** (2004), 77-86.
- [8] K. Saithanu, J. Mekpanyup, Modeling to Predict Decision in Continuing Study for Master's Degree, *Burapha Science Journal*, **16(1)** (2011), 3-11.
- [9] R. Nochai, T. Nochai, Classification of the Student Learning Ability Using Discriminant Analysis: A Case Study of Administration and Management College, King Mongkut's Institute of Technology Ladkrabang, Thailand, *The Third Asian Conference on Education* (2011).
- [10] B. L. Garton, T. Kitchel, University Admission Criteria and Learning Style: Predictors of Academic Success?, *NACTA Journal*, **49(2)** (2005), 10-15.
- [11] N. A. Ismail, A. Othman, Comparing university academic performances of HSC students at the three art-based faculties, *International Educational Journal*, **7(5)** (2006), 668-675.
- [12] An academic affair of Phanatpittayakarn School, *Information of school*, Retrieved January 20, 2012, from <http://www.panatp.ac.th/pphis.html/>
- [13] R. A. Johnson, D. W. Wichern, *Applied Multivariate Statistical Analysis, 6-th Edition*, Prentice-Hall Press, N.J, USA (2007).
- [14] C. Oh, S. G. Ritchie, Recognizing vehicle classification information from blade sensor signature, *Pattern Recognition Letters*, **28(9)** (2007), 1041-1049.

