Walailak J Sci & Tech 2008; 5(1): 111-123.

Utilization of a Self Organizing Map as a Tool to Study and Predict the Success of Engineering Students at Walailak University

Wattanapong KURDTHONGMEE

School of Engineering and Resources Management, Walailak University, Nakhon Si Thammarat 80161, Thailand

(E-mail: kwattana@wu.ac.th)

ABSTRACT

Many factors have an influence on the success of undergraduate students particularly in engineering programs. Some students have to drop out as a result of obtaining very poor GPA (grade point average) and/or GPAX (accumulated grade point average) after only their first year of studying. It would be helpful for students if they know how their current GPA/GPAX could be improved in order to successfully graduate. In addition, what would be the expected outcome of their study, if their current GPAs of compulsory subjects are not fairly good? In this paper, the Self Organizing Map (SOM) neural network is utilized as a tool to cluster engineering student data into different groups by means of their study results. The results are then used to produce the weight maps. The maps reflect the correlation between GPA/GPAX of the compulsory subjects and the educational status of students. The result from the SOM with some adaptations to its matching phase is also used to create a predictor which is capable of producing a fairly high degree of correctness. The meaningful results are intended to be used as a guideline for students to prepare and improve themselves. In addition, it might be useful for student advisors and counselors to give appropriate advice to students whose GPAX are critically low. This can be accomplished by advising students to register less or withdraw some subjects in order to leverage their GPAX. In addition, some students should be advised to change their field of study if they perform fairly poorly in all compulsory subjects. The approach utilized in this paper is a novel one with respect to this application domain.

Keywords: Self organizing map, engineering education, education prediction

INTRODUCTION

A recent study [1] has reported that while the number of jobs in the USA and many parts of the world requiring training in science and engineering is on the rise, the number of students receiving training in these fields is declining at a disturbing rate. Taking a step toward increasing interest in the fields and boosting student enrollments in related programs is not enough. Retention rates must also be taken into consideration and leveraged. It is pointed out that retention rates in engineering fields are much too low and problematic [2]. That is to say they rely on many factors. Estimates of the loss in undergraduate students who begin engineering and either switch to another major or drop out altogether range from 40 % to 70 %. On average only 42 % of students complete the major with which they began their undergraduate studies. Studies of why students migrate out of engineering have identified several factors at work. They include poor academic performance, inadequate preparation, lack of self-efficacy and unwillingness to work.

Many factors influence the success or failure of undergraduate students particularly in engineering programs which are the focus group of this paper [3-5]. Results from previous publications [2,6] have concluded that family and educational backgrounds, attitudes and student performance in the introductory courses are all related to the success or failure of students. Research to study the impact of these factors was previously carried out at North Carolina State University. Data on backgrounds, attitudes, personality types, study skills, and freshman-year performance records of 124 students in an introductory chemical engineering course were collected. These were then statistically analyzed to determine factors in a student's background that might significantly act as predictors of success or failure in the course, and by extension, in the chemical engineering curriculum. The results revealed that the probability of passing the course with a grade of C or better depends on the type of home community, time devoted to an outside job, father's educational level, SAT (Scholastic Aptitude Test) scores on mathematics and language, freshman-year grade point average, and also grades in selected freshman mathematics, physics, chemistry and English courses. The retention of students based on gender and minority difference was also evaluated [2,6]. It was reported that although women began their pursuit of an undergraduate engineering degree with equal or better credentials than their male peers, by the end of their sophomore year they were more likely to drop out as compared to male students.

In another publication [7], the evaluation to predict the academic success of engineering students at Mercer University was performed. The academic success is measured by GPA after completion of the first 2 semesters. The freshman student scores from 1 non-technical assignment are used as the predictor. The dominant point of this research relies on the fact that a single, non-technical, variable is used as a predictor. This research also employed statistical analysis as a tool for evaluation. The result indicates that there was a strong relationship between the student notebook scores and GPA with a 99 % confidence level.

Despite the fact that artificial neural networks (ANN) have become popular in many disciplines, their adoption in the education field is relatively rare [8]. Numerous publications have continuously reported the results of studying the relationships between the factors described earlier to the retention rate (and also success rate) of engineering students in different institutions with an emphasis on using statistics as an analytical tool. With the large volume of education data and influential factors, there is a need for efficient tools to analyze the data and make a prediction model. In this paper, we propose a technique for structuring, arranging, visualizing, and predicting the success based on the information from a large academic database of engineering students. Our procedure is based on self-organizing maps [6,7,9,10], an ANN commonly used for information extraction from large databases in the form of 2 dimensional maps. The resulting maps along with our proposed modifications will provide easy to understand visual information that can help engineering students to prepare and improve themselves. In addition, it might be useful for student advisors and counselors to give appropriate advice to their students whose GPAX and/or GPA of some compulsory subjects are critical. The ultimate outcome of this result is that it may increase the retention rate of engineering students and also elevate the number of successful engineers.

MATERIALS AND METHODS

In this section, we provide a description of the Kohonen self organizing map technique for data analysis and visualization. Input data preparation, details of the software tool employed during the course of our experiments and the adaptations to the standard Kohonen self organizing approach are also given.

Kohonen self organizing map for data analysis and visualization

ANNs are tools for non-linear data processing and pattern recognition loosely based on the structural arrangement of neurons in the brain. ANNs are now applied to many real-world problems such as computer vision, speech processing, non-linear control, etc. The dominant points of ANNs are their non-linear operation, learning from examples, and noise tolerance. In practice, ANNs are usually implemented as a computer program. However, for critical processing time applications, they can be implemented in hardware as a set of parallel processing elements [11].

The SOM is a competitive unsupervised ANN typically used for pattern recognition [9,10,12,13], exploratory data analysis and for visualizing large databases. A database can be considered as a set of input vectors (or patterns), each consisting of several input variables. In our application, a record of engineering student data is represented by an input vector (or pattern), consisting of variables such as his/her obtained GPAs for any subjects (only common engineering compulsory subjects in the first year), status (only 2 statuses: retired and successfully graduate statuses) and the GPAX. The SOM has the special property of effectively creating spatially organized "internal representation" of various features of input signals and their abstractions. One successful result is that the self-organizing process can also discover semantic relationships in sentences [10]. In this respect, the resulting maps very closely resemble the topographically organized maps found in the cortices of the more developed animal brains. After supervised fine tuning of its weight vectors, the SOM is particularly successful in various pattern recognition tasks involving very noisy signal levels.

Let us briefly describe the algorithmic detail of the SOM [12,13]. Assume that some data sets have to be mapped onto the array; the set of input samples is described by a real vector $x(t) \in R^n$ where t is the index of the sample, or the discrete-time coordinate. Each node i in the map contains a model vector $m_i(t) \in R^n$, which has the same number of elements as the input vector $x(t)$.

The stochastic SOM algorithm performs a regression process. Thereby, the initial values of the components of the model vector, $m_i(t)$, may even be selected at random. In practical applications, however, the model vectors are more profitably initialized in some orderly fashion, e.g., along a 2 dimensional subspace spanned by the 2 principle eigenvectors of the input data vectors. Moreover, a batch version of the SOM algorithm may also be used.

Any input item is thought to be mapped into the location, the $m_i(t)$ which matches best with $x(t)$ in some metric. The SOM algorithm creates the ordered mapping as a repetition of the following basic tasks:

1. An input vector $x(t)$ is compared with all the model vectors $m_i(t)$. The best-matching unit (node) on the map, i.e., the node where the model vector is most similar to the input vector in some metric (e.g. Euclidean) is identified. This best matching unit (BMU) is often called the winner.
2. The model vectors of the winner and a number of its neighboring nodes in the array are changed towards the input vector according to the learning principle specified below.

The basic idea in the SOM learning phase is that, for each sample input vector $x(t)$, the winner and the nodes in its neighborhood are changed closer to $x(t)$ in the input data space. During the learning phase, individual changes may be contradictory, but the net outcome in the process is that ordered values for $m_i(t)$ emerge over the array. If the number of available input samples is restricted, the samples must be presented repeatedly to the SOM algorithm.

Adaptation of the model vectors in the learning phase may take place according to the following equations:

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)] \quad \text{for each } i \in N_C(t)$$
$$m_i(t+1) = m_i(t) \text{ otherwise,}$$

where t is the discrete-time index of the variables, the factor $\alpha(t) \in [0,1]$ is a scalar that defines the relative size of the learning step, and $N_c(t)$ specifies the neighborhood around the winner in the map array.

At the beginning of the learning phase the radius of the neighborhood is fairly large, but it is made to shrink during learning phase. This ensures that the global order is obtained at the beginning, whereas towards the end, as the radius gets smaller, the local corrections of the model vectors in the map will be more specific. The factor $\alpha(t)$ also decreases during learning.

One method of evaluating the quality of the resulting map is to calculate the average quantization error over the input samples, defined as $E\{\|x - m_C(x)\|\}$ where c indicates the BMU for x. After training, for each input sample vector the BMU in the map is searched for, and the average of the respective quantization errors is returned.

To visualize the final SOM, the unified distance matrix method (U-matrix) will be used. The U-matrix method can be used to visualize structure in a high-dimensional data space. It also makes it possible to classify data sets into clusters of similar values. The simplest U-matrix method is to calculate the distances between neighboring neurons, and store them in a matrix, that is the output map, which then can be interpreted. If there are ‘walls’ between the neurons, the neighboring weights are distant, that is, the values differ significantly. The distance values are also displayed in color when the U-matrix is visualized. Hence, dark colors represent great distances while brighter colors indicate similarities among the neurons.

By viewing the individual feature planes, it is possible to visualize the values of a single vector column, that is, in this study, the maps for one study parameter. These feature planes can be analyzed in order to discover how well the students have been doing in relation to a single parameter. Thus, using the feature planes, it is simple to see where the successful graduate students with good GPA/GPAX are located on the map and, in the same fashion, where the retired students with poor GPA/GPAX are located. By using both the U-matrix and the individual feature planes together, we can thus finally identify the clusters on the map, as well as the characteristics of these clusters.

With the nature and many benefits of the SOM in mind, we are quite sure that the approach can be applied to our problem domain, even though previously published papers have not applied SOM to this specific area. In the next section, the details of input data preparation, software tool selection, and our proposed adaptations to the standard SOM algorithm are given.

Input data and software tool

Our aim is to show the interest of SOM as a tool for making the relationships between engineering students’ GPAs, their GPAX and their study status. In addition, we expect that the outcome may be used to create a model to predict the success of engineering students based on such influential parameters. The postulation behind our idea is that students with high GPAs in all compulsory subjects during the first year would also obtain a high GPAX for the whole period of their study. This comes from the fact that compulsory subjects are believed to be the foundation of all advanced engineering subjects. This group of students would successfully graduate without any obstacles. For some students, however, they might obtain a high GPA in some subjects but fail in other subjects. What would be the studying outcome of this group of students? How much effort is necessary in order to survive in the engineering programs? We expect that the SOM weight maps can reveal this information after training it with our history of engineering student database.

In our study, we made use of the engineering student database during the period of 1998 - 2005 from the Center for Educational Services, (CES) of Walailak University. CES is mainly responsible for processing, recording and reporting student grades. The data includes a wide range of engineering student records, extending from retired first year students with very low GPAX to highly successful first class honor students with high GPAX. The average GPAX for this group of students is 2.39. The total number of engineering student records is 720. This is quite a small number compared to long established universities. This makes it inappropriate to employ the very popular data mining approach for analysis. From several compulsory subjects in the first year, we are interested in only 3 main groups of subjects which are English, Mathematics and Physics. These are, in our opinion, 3 main subjects which are the foundation for all engineering students. The GPAs for these groups of subjects are carried out from CES's raw data. For each subject i, its alphabetical value, V_i, is substituted by the numeric counterpart, G_i; i.e. $F = 0, D = 0.5, D+ = 1,$ …and $A = 3.5$. It is noted for clarity that the descriptions of these grades are as follows: fail, very poor, poor, fair, fairly good, good, very good, and excellent, respectively. The numeric value for each grade is then multiplied by the credits for that subject C_i. The intermediate result is then summed up with all subjects belonging to the same group; for example Calculus I, Calculus II, Advanced Mathematics I, Advanced Mathematics II, and Engineering Statistics which all belong to Mathematics. Along with the sum of the multiplication between the numeric value of grades and the credits for the subject, the sum of all credits belonging to this group is also performed. Then, the GPA of this group of subjects is computed from the division between the former and the latter sums. The following equation summarizes the process of GPA or GPAX calculation:

$$GPA, GPAX = \frac{\sum G_i C_i}{\sum C_i}$$

For GPA, i is an index of all subjects belonging to the same group up to the time it is calculated. For GPAX, i covers all the subjects up to the time it is calculated.

There is no need to transform the GPAX as it is in a usable form. After processing, these 2 parameters, GPAs and GPAX, have equal ranges which are between 0 and 4.00. In order to prove the validity/accuracy of prediction of the SOM algorithm, 80 % of the input data in which the study outcome was known was used for training and the rest was used for testing during the course of experimentations.

Having finished input data preparation for training the SOM, we then need to select the appropriate software tool to carry out an experiment. The software used for training and creating the SOM is called "The Self-Organizing Map Program Package" version 3.1 (SOM_PAK), and is based on the Kohonen self-organizing algorithm. The software package was developed by the SOM programming team at the Helsinki University of Technology. During the training process, a number of maps were created in order to determine the most suitable parameters. The maps were constantly monitored in terms of both quantization error and visual cluster structure. The hexagonal lattice type was preferred for the visualization of the output map. In addition, the map ought to be rectangular, rather than square, in order to achieve a stable orientation in the data space as recommended by Kohonen. Commonly, the x-axis should be about 30 % greater than the y-axis, thus forming a rectangular output map. Another recommendation is that the training length of the second part should be at least 500 times the number of network units, in order to reach statistical accuracy.

The final trained map is of the size 10 × 8 neurons. On the final U-matrix map (**Figure 1a**), it is fairly simple to define the different clusters by looking at the color shades of the borders between the hexagons. The brighter colors of the borders between the hexagons imply similar characteristics, while darker borders represent greater distances.

The software we have used to visualize the final constructed self-organizing maps and the feature planes in this research is a program called Nenet version 1.1a. This software was developed by the Nenet team also at the Helsinki University of Technology. Nenet is a user-friendly program designed to illustrate the use of SOMs, and provides an easy way to visualize the output maps with not only the U-matrix method but also the interpolated 2D U-matrix method as individual parameter level maps.

With respect to **Figure 1**, it can be seen that the SOM is capable of grouping input data into different clusters. The relationships between clusters from different planes can be preliminarily summarized. For example, consider the clusters of successful student whose labeled status is '1' in the bottom-right of **Figure 1f**, it is observable from the colours of the maps that these group of students obtained high GPAs in both Mathematics and Physics. As a result, these led them to achieve a high level of GPAX. However, they obtained a moderate level of GPA in English. In contrast, consider the group of retired students whose labeled status is '0' within the third row and the first column in **Figure 1f**. The maps indicate that they got very poor GPA in both Mathematics (the same cluster in **Figure 1b**) and Physics (the same cluster in **Figure 1d**) which are about 0.00. They also got a poor GPA in English (the same cluster in **Figure 1c**) which is about 0.28.

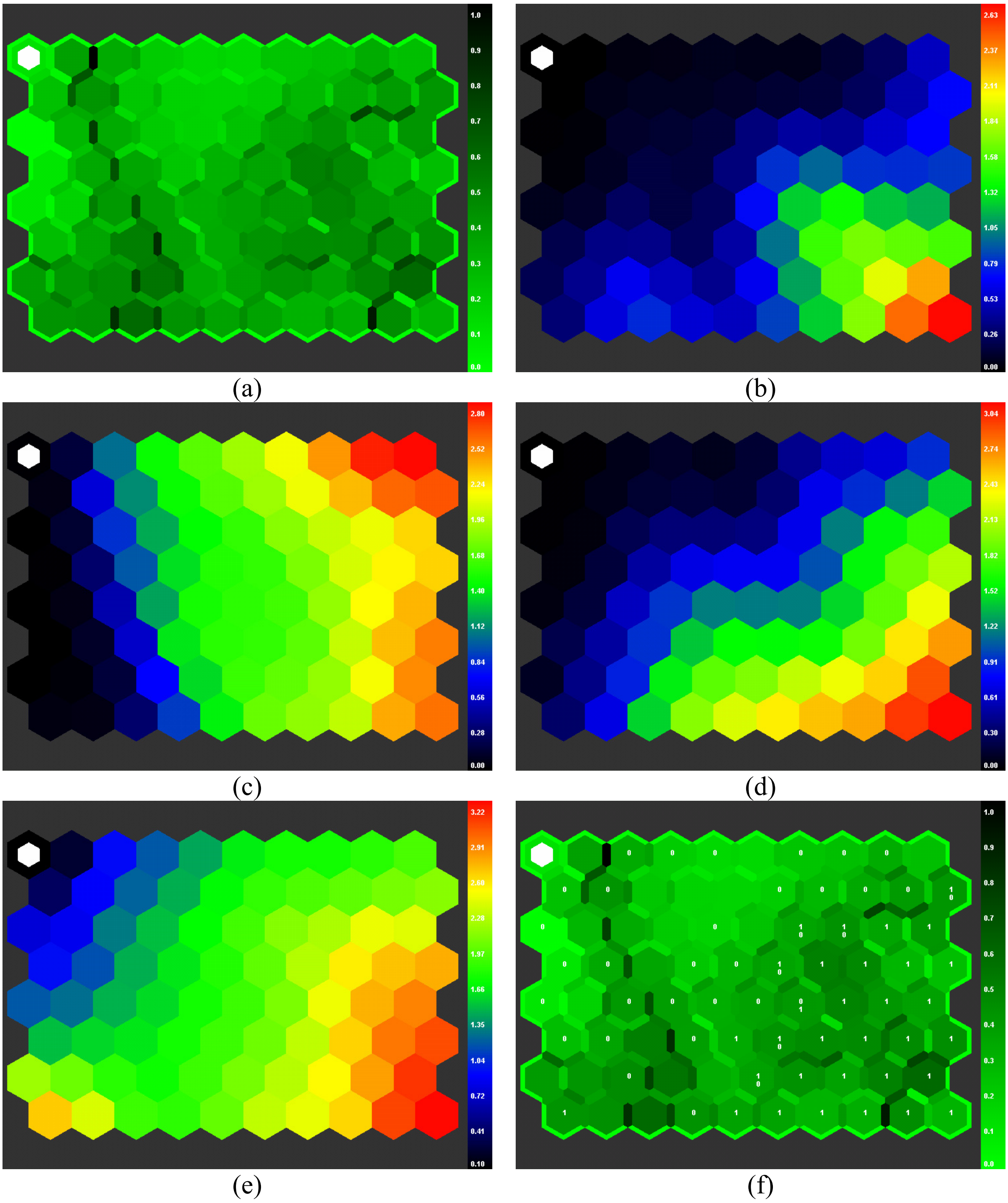

Figure 1 (a) The final U-matrix map and (b-e) feature planes maps of GPA of Mathematics, GPA of English, GPA of Physics, and GPAX, respectively and (f) the final U-matrix map with labeled student status.

The adaptations to the standard Kohonen self organizing approach

The standard SOM algorithm is capable of presenting the correlations between the GPAX, and GPA of compulsory subjects. As a prediction tool, it is, however, able to forecast the study results with only moderate and fairly low probabilities for a successful graduate and a retired student, respectively. These limitations required us to propose adaptations to the standard SOM algorithm especially during the matching phase in order to leverage its correctness of predictions. With respect to the standard SOM algorithm, the matching phase works by finding the BMU or node on the map whose distance (i.e. Euclidean distance) with respect to the input vector is the smallest. This is then followed by labeling the node to indicate to which vector the BMU is matched. Some nodes may be matched by several test vectors with some common properties. According to our application domain and experimental results, many nodes were matched by the vectors of students with both successfully graduated and retired statuses. These are indicated as nodes and labeled (0,1) in **Figure 1f** where 0 and 1 indicate retired and successful graduate statuses, respectively. During the querying process, if a vector of the student record is matched to such nodes, the standard SOM produces the prediction result “unsure” (unpredictable). This contributed to as high as 62.5 % of unpredictable results and reduced the overall validity of our model.

In order to leverage the correctness of predictions, we propose modifications to the data structure of the map’s node and the algorithm within the matching phase of the standard SOM algorithm. The data structure of a node is extended to include multiple counters. Each counter indicates the frequency of being matched by a corresponding status of input vectors during the matching phase. The counters are all reset to zero before performing the matching phase. During the matching phase, after the BMU on the map with respect to an input vector has been found, apart from only labeling the node with the matched vector, the frequency of being matched with respect to the matched vector’s status is also updated. Utilizing this method, although some nodes are possibly matched by several input vectors with different states, the frequency of being matched can be used to justify the final classification status. It is noted that the proposed adaptations can be applied to the outcome of the SOM model after the training phase.

RESULTS AND DISCUSSION

Several hundred maps were trained during the course of the experiment. The best maps, rated according to quantization error and ease of readability, were then selected and used as a basis when training further maps. In order to achieve statistical accuracy, the initial phase includes 3,150 steps and the final phase 31,500 steps. The learning rate factor was set to 0.5 in the first phase and 0.06 in the second, which are near commonly used starting points. The neighborhood radius was set to 11 for the first phase and 1 for the second. The initial network radius was very large, but seemed to provide the best overall maps. Decreasing the radius only resulted in poorer maps. As Kohonen noted [13], the selection of parameters appears to make little difference in the outcome when training small maps. With different selections of parameters, the changes in the quantization error were very small, usually as little as 0.001.

As mentioned earlier, the clusters which indicate the correlations among the input parameter were identified by studying the borders between the neurons on the final U-matrix map as shown in **Figure 1a**. Dark borders identify large differences, that is, the edges of the clusters, whereas light borders identify similarities. The characteristics of the clusters are identified by studying the underlying feature planes of the map. The identified 2 clusters are illustrated in **Figure 1f** which consists of the successful graduate student and the retired student clusters. It is noticeable that many nodes of the map enclose both the study statuses.

The resulting SOM model after clustering was then used for validating. The Matlab®-script was specially developed to automate the query process of the SOM model with respect to the student data previously preserved to use as test data (the latter 20 % group). The predicted results are shown in the middle column of **Table 1**. From the table and with respect to the standard SOM model, it can be interpreted that the model is capable of making correct prediction with 67.3 and 28.6 % for the successful graduate and the retired students, respectively. This means that for any student's record if the standard SOM model is queried and it predicts that the student is a successful graduate student, the probability of being a successful graduate student is 0.67. Similarly, if the standard SOM model is again queried with a student data and it predicts that the student is a retired student, the probability of being a retired student is 0.28. From the results, it is a surprise to observe that the standard SOM model can not determine the result in as high as 62.5 % of the test data.

Table 1 The predicted results obtained by querying the SOM model, with and without adaptations to the matching phase, with preserved student data with known study results.

Classification results	Standard SOM	Modified SOM
Successful student	67.3	82.1
Retired student	28.6	57.1
Unsure classification	62.5	7.7

With respect to our proposed adaptations to the standard SOM algorithm, the Matlab®-script was also coded to separately create the necessary data structure and perform the modified matching phase algorithm. This was then used to query for the study status with respect to the same test set of the student data. The predicted results are shown in the rightmost column of **Table 1**. The proposed approach proved capable of leveraging the probability of correct predictions to 0.82 and 0.57 for the successful graduate and retired students, respectively. In addition, the probability of unpredictable results is reduced significantly to only 0.08.

It is obvious from our study that the SOM in the standard form can reveal the influence of the study results of some compulsory subjects and GPAX to the success of engineering students. This was confirmed by the final map, which indicates the clusters of student data whose properties are almost similar together, as illustrated in **Figure 1**. The cluster of data indicates that all the input parameters or the components of the input vectors used in this study have a correlation. Without any requirement to draw an explicit correlation equation which is the prominent benefit of utilizing the neural network approach, the SOM model can be used as a predictor for an unknown set of input data. By directly following the standard algorithm for the SOM matching phase, the predicted results gave quite a low probability of correctness and a high probability of unpredictable results. The proposed adaptations to the standard SOM algorithm indicated that the predicted results were improved significantly and made the predictor more reliable to be used in practice.

CONCLUSIONS

In this paper, the self organizing map neural network was proposed to be utilized as a tool to automatically cluster engineering student data by means of their study results. The tool proved to be successfully used to produce the feature maps which reflected the correlations between the GPA of major compulsory subjects and the educational status of students. The results confirmed that the success in engineering disciplines relies heavily on compulsory subjects: Mathematics, Physics and English. The result from the SOM could then be used together with the adaptations to some parts of its algorithm to make a better predictor of the study status, compared to the standard approach, given the GPA of compulsory subjects and GPAX.

The results from this study are intended to be used as a guideline for students to prepare and improve themselves during the period of their study. In addition, it might be useful for student advisors and counselors to give appropriate advice to students who are in critical conditions. In practice, this can be accomplished by advising students to register less or withdraw some subjects in order to leverage their GPAX. In addition, some students should be advised to change their field of study in case they perform poorly in all compulsory subjects.

The contribution of this paper is that it is the first time that an SOM has been proposed and successfully used as a tool for extracting valuable and qualitative information, visualizing in a simple format, and used for predicting success in engineering education. The dominant point of the SOM to this application domain is that it is easy to use for data classification, analysis and interpretation. The output from the SOM is in a ready-to-present and easy-to-understand format. In addition, as a prediction tool, the SOM with our proposed adaptations is capable of making fairly accurate results.

ACKNOWLEDGEMENTS

The author would like to thank Mr. Numchai Sae-li of the Center of Educational Services (CES), Walailak University, for providing engineering student raw data. Also, Asst. Prof. Dr. Supit Rittikaew, director of the CES, is thanked for his very kind permission for the author to use the data.

REFERENCES

[1] H Hartman and M Hartman. Leaving engineering: Lessons from Rowan University's College of Engineering. *J. Eng. Educ.* 2006; **1**, 49-61.

[2] RM Felder, KD Forrest, L Baker-Ward, EJ Dietz and PH Mohr. A longitudinal study of engineering student performance and retention: I. Success and failure in the introductory course. *J. Eng. Educ.* 1993; **1**, 15-21.

[3] DC Dickenson. Prediction and success of freshmen Engineers. *Pers. Guid. J.* 1969; **6**, 1008-14.

[4] RL Elkins and JF Leutkemeyer. Characteristics of successful freshmen engineering students. *J. Eng. Educ.* 1974; **4**, 189-91.

[5] J Levin and J Wycokoff. Prediction persistence and success in baccalaureate engineering. *J. Eng. Educ.* 1991; **4**, 461-8.

[6] RM Felder, KD Forrest, L Baker-Ward, EJ Dietz and PH Mohr. A longitudinal study of engineering student performance and retention: III. Gender difference in student performance and attitudes. *J. Eng. Educ.* 1995; **2**, 151-63.

[7] LW Lackey, WJ Lackey, HM Grady and MT Davis. Efficacy of using a single, non-technical variable, to predict the academic success of freshmen engineering students. *J. Eng. Educ.* 2003; **1**, 41-8.

[8] CWT Chiu, P Pashley, M Seastrom and P Carr. Visualizing large-scale data in educational, behavioral, psychometrical and social sciences: utilities and design patterns of the SEER computational and graphical statistics. *Journal of Information Visualization* 2005; **4**, 276-89.

[9] T Eklund, B Back, H Vanharanta and A Visa. Using the self-organizing map as a visualization tool in financial benchmarking. *Journal of Information Visualization* 2003; **2**, 171-81.

[10] T Honkela. Self-Organizing Maps in Natural Language Processing, Available at: http://www.cis.hut.fi/~tho/thesis, accessed February 2007.

[11] W Kurdthongmee. A novel Kohonen SOM-based image compression architecture suitable for moderate density FPGAs. *Image Vision Comput.* **26**; 2008, 1094-105.

[12] AH Dekker. Kohonen neural networks for optimal colour quantization. *Network-Comp. Neural* 1994; **1**: 354-67.

[13] T Kohonen. The self-organizing map. *Proceedings of the IEEE* 1990; **9**, 1464-80.

บทคัดย่อ

วัฒนพงศ์ เกิดทองมี

การประยุกต์ใช้โครงข่ายประสาทเทียมในกลุ่มเซลฟ์ออแกไนซิ่งแมปเพื่อเป็นเครื่องมือในการศึกษาและพยากรณ์การสำเร็จการศึกษาของนักศึกษาวิศวกรรมศาสตร์ มหาวิทยาลัยวลัยลักษณ์

การสำเร็จการศึกษาของนักศึกษาโดยเฉพาะอย่างยิ่งในสาขาวิศวกรรมศาสตร์ขึ้นอยู่กับหลายปัจจัย นักศึกษาส่วนหนึ่งจำเป็นต้องหยุดการศึกษาทั้งๆที่กำลังศึกษาในชั้นปีที่ 1 อันเป็นผลมาจากผลการเรียนต่ำในบางรายวิชาหรือทุกรายวิชารวมกัน การที่นักศึกษาทราบข้อมูลล่วงหน้าในด้านความสัมพันธ์ระหว่างเกรดรายวิชาใดๆ และเกรดเฉลี่ยสะสมที่มีต่อการจบการศึกษาจะเป็นประโยชน์อย่างยิ่งในการวางแผนการเรียน นอกจากนั้นหากมีรูปแบบความสัมพันธ์ที่น่าเชื่อถือได้ระหว่างเกรดเฉลี่ยของรายวิชาบังคับที่มีต่อการสำเร็จการศึกษาน่าจะเกิดประโยชน์ต่อนักศึกษาในการตัดสินใจเพื่อเปลี่ยนหลักสูตรการศึกษาหรือเพิ่มความพยายามมากขึ้นเพื่อให้สามารถสำเร็จการศึกษาได้ ในบทความนี้ผู้เขียนได้ประยุกต์ใช้โครงข่ายประสาทเทียมในกลุ่ม Self Organizing Map เพื่อจัดกลุ่มข้อมูลผลสัมฤทธิ์ทางการศึกษาของบัณฑิตสำนักวิชาวิศวกรรมศาสตร์และทรัพยากร มหาวิทยาลัยวลัยลักษณ์ โดยผู้เขียนได้แยกกลุ่มข้อมูลของบัณฑิตส่วนหนึ่งเพื่อใช้ในขั้นตอนการเรียนรู้ของโครงข่ายประสาทเทียม ผลที่ได้คือ แผนภาพที่แสดงให้เห็นถึงความสัมพันธ์ระหว่างเกรดเฉลี่ยของรายวิชาบังคับในชั้นปีที่ 1 (เฉพาะรายวิชาในกลุ่มคณิตศาสตร์ ฟิสิกส์ และภาษาอังกฤษ) เกรดเฉลี่ยสะสมและผลสัมฤทธิ์ทางการศึกษา จากผลที่ได้นี้ผู้เขียนได้นำเสนอกระบวนการในการวิเคราะห์เพิ่มเติมและทดสอบโดยการป้อนเกรดเฉลี่ยของของรายวิชาบังคับ (ดังกล่าวข้างต้น) และเกรดเฉลี่ยสะสมของบัณฑิตที่อยู่นอกเหนือจากกลุ่มที่ใช้ไปในข้างต้นในการทดสอบการทำนายผลสัมฤทธิ์ทางการศึกษา ผู้เขียนพบว่าโครงข่ายประสาทเทียมที่ใช้รวมทั้งกระบวนการในการวิเคราะห์เพิ่มเติมที่นำเสนอสามารถทำนายได้ผลถูกต้องด้วยเปอร์เซ็นต์สูง ผลที่ได้ทั้งในด้านความสัมพันธ์ระหว่างเกรดเฉลี่ยของรายวิชาบังคับในชั้นปีที่ 1 เกรดเฉลี่ยสะสมและผลสัมฤทธิ์ทางการศึกษาและการทำนายมีประโยชน์ต่อการนำไปใช้เพื่อการเตรียมตัวและเพิ่มความพยายามของนักศึกษาเพื่อให้สามารถสำเร็จการศึกษาได้ นอกจากนั้นยังอาจมีประโยชน์ต่ออาจารย์ที่ปรึกษาในการให้คำปรึกษาที่เหมาะสมต่อกลุ่มนักศึกษาซึ่งเกรดเฉลี่ยของรายวิชาบังคับในชั้นปีที่ 1 และเกรดเฉลี่ยสะสมอยู่ในภาวะวิกฤต

สำนักวิชาวิศวกรรมศาสตร์และทรัพยากร มหาวิทยาลัยวลัยลักษณ์ อำเภอท่าศาลา จังหวัดนครศรีธรรมราช 80161