

# Automatic Medical Case Study Essay Scoring by Support Vector Machine and Genetic Algorithms

S. Yenaeng, S. Saelee, and W. Samai

**Abstract**—The study of medical various institutions in Thailand found that the success for teaching and learning by means of Problem Based Learning (PBL) depends on several factors including correct attitudes and cognitive learning. Problem of teaching a small group of PBL is consistent with the results of the audit were to evaluate the group's facilitator, diverse and inaccurate. Especially, the grading of a medical case study essay reports (Clinical case summaries). To solve such problems, we proposed automatic medical case study essay scoring for PBL of medical students. SVM with Genetic Algorithms (GA-SVM) was used to assess the quality medical case study essays written by medical students in the subject matter of muscular systems and movement. The medical case study essays written in response to a question were each evaluated by facilitators and assigned a human score. In the experiment, we used raw term frequency vectors of the essays and their corresponding human scores to train the SVM while GA was used for choosing the kernel function type and its parameter values to find a proper solution to an optimization and obtain the machine scores. The experimental results show that the addition of GA-SVM technique improves scoring performance.

**Index Terms**—Essay scoring, medical case study, SVM, genetic algorithms (GA).

## I. INTRODUCTION

Medical institutions in Thailand really realize the importance for teaching and learning to provide medical students with skills in various fields and the potential to accumulate knowledge in order to solve the problems systematically. Problem Based Learning (PBL) is a form of learning that encourages students to learn by themselves. This is a good reason to implement the use of the PBL system for teaching and learning in the faculty of medicine for more than a decade. Chulalongkorn University is the first institution to implement the PBL for teaching in the Faculty of Medicine. Later, Khon Kaen University, Thammasat University and Prince of Songkla University (PSU) have been used the PBL system respectively. However, it seems that PBL is still very little successful in Thailand. The study of various institutions, both in Thailand and abroad, found

Manuscript received September 5, 2013; revised November 13, 2013. This work was supported in part by the department of computer education, faculty of technical education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.

S. Yenaeng is with the Department of Computer Education, Faculty of Education, Bansomdejchaopraya Rajabhat University, Bangkok, Thailand (e-mail: modssk@gmail.com).

S. Saelee is with Department of Computer Education, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.

W. Samai is with the Faculty of Medicine, Prince of Songkla University, Songkhla, Thailand.

that the success for teaching and learning by means of PBL depends on several factors including correct attitudes and cognitive learning. The PBL process is to learn in small groups and focuses on student-centered learning by means of problem solving (Case or Scenario). Cases or scenarios combine the contents of several subjects. Medical students need problems to get to a common set of assumptions and comment on the most important things students need to write learning objectives of each problem. Problem of teaching a small group PBL is consistent with the results of the audit were to evaluate the group's facilitator, diverse and inaccurate. Especially, the grading of a medical case study reports (Clinical case summaries).

Computers and artificial intelligence have been proposed as tools to facilitate the evaluation of student essays. In theory, computer scoring can be faster, reduce costs, increase accuracy, and eliminate concerns about rater consistency and fatigue. Further, the computer can quickly rescore materials should the scoring rubric be redefined. Using different methods, Page [1], [2], Landauer, Holtz, and Laham [3], and Burstein [4] report very high correlations between human judgment and computer generated scores. Page uses a regression model with surface features of the text (document length, word length, and punctuation) as the independent variables and the essay score as the dependent variable. The approach by Landauer *et al.* [3] is a factor-analytic model of word co-occurrences which emphasizes essay content. Burstein [4] uses an eclectic model with different content features.

For this reason, many machine learning for text classification techniques have been recently considered to design Automated Essay Scoring (AES). Kakkonen *et al.* [5] present an automatic essay grading with probabilistic latent semantic analysis. This technique proposed to improve the problems found in Latent Semantic Analysis (LSA). The methods were found to be almost equal in the accuracy measured by Spearman correlation between the grades given by the system and a human. Ramineni and Williamson [6] present an automated essay scoring: Psychometric guidelines and practices which overview of psychometric procedures and guidelines Educational Testing Service (ETS) uses to evaluate automated essay scoring for operational use. They briefly describe the e-rater system, the description of e-rater includes a summary of characteristics of writing covered by e-rater, variations in modeling techniques available, and the regression-based model building procedure. The evaluation procedures cover multiple criteria, including association with human scores. They conclude that the a priori establishment of performance expectations and the evaluation of performance of e-rater against these expectations help to ensure that automated scoring provides a positive

contribution to the large-scale assessment of writing. Rudner and Liang [7] present an automated essay scoring using Bayes' theorem. Bin *et al.* [8] present an automated essay scoring using the KNN algorithm. They presents an approach to essay scoring that builds on the text categorization model which incorporates K-Nearest Neighbor (KNN) algorithm. Transforming the essays into the vector space model (VSM), TF-IDF and IG are applied for feature selection from the feature pool of words, phrases and arguments. After training for the KNN algorithm, a precision over 76% is achieved on the CLEC corpus.

This paper presents an automatic medical case study essay scoring by support vector machine and genetic algorithm, we introduced that the application of GA to optimize the kernel function and all parameters of the SVM based prediction model can be used to assign quality scores of essays.

Support vector machines were first introduced by Vapnik and colleagues [9] and have recently been used for some problems such as bioinformatics and pattern recognition. SVM provides a help to fit a hyperplane surface to the training data by a kernel function. A crucial problem in SVM is how to set the best kernel parameters because it influences the problem performances. The parameters that should be optimized are penalty parameter  $C$  and the kernel function parameters such as the gamma ( $\gamma$ ) for the radial basis function kernel. Therefore, to dominate this challenge several hybrid methods had been proposed by researchers.

The biggest difficulties in setting up the SVM model are to choose the kernel function and its parameter values. If the parameter values are not set properly, then the classification outcomes will be less than optimal ones.

In addition, the proper parameter setting can improve the SVM classification accuracy. The parameters that should be optimized including penalty parameter  $C$  and the kernel function parameters such as the gamma ( $\gamma$ ) for the radial basis function (RBF) kernel. To design a SVM, one must choose a kernel function, set the kernel parameters and determine a soft margin constant  $C$  (penalty parameter).

Genetic algorithms have the potential to generate the optimal SVM parameters at the same time. Our research objective is to optimize the parameters. The proposed method performs parameters setting in an evolutionary way.

This paper is organized as follows: the tools and method are mentioned in Section II. Section III describes the proposed GA-SVM model. Section IV discusses the experimental design. Section V presents the experimental results from using the proposed method to classify several real world datasets and Section VI describes the conclusions.

## II. THE TOOLS AND METHOD

### A. Medical Case Study Essays Dataset

The sample was drawn from a subject matter of "Muscular systems and movement" a course taught in the Faculty of Medicine at PSU in the pre-clinical medicine. Two hundred medical students of the same class level were asked to respond to a question on the subject matter by writing a short medical case study essay. This essay dataset contains 3 fields (see Fig. 1) of which 200 essays were problems and 416

fields were hypotheses and 226 fields were learning objectives. The scores of an essay has obtained by averaging the scores from three facilitators who have been treated as human scores. Next, we performed word segmentation on the essays using the longest and maximal matching algorithm. Raw term frequency vector was then created for each essay.

The dataset was divided into two parts. One was the training set (70% of samples) which was used for training the parameter of SVM. The other was the testing set (30% of samples) which was introduced for testing the ability of SVM.

### B. Support Vector Machines (SVM)

SVM can be briefly described as follows [10]–[12]. Let  $(x_1, y_1), \dots, (x_m, y_m) \in X \times \{\pm 1\}$  denote a set of training data, where  $X$  represents some nonempty set from which the pattern  $x_i, y_i$  are called target  $y_i \in \{\pm 1\}$ . SVM attempts to identify a hyper-plane in a multidimensional space. This hyper-plane function was a separating plane for data classification. Consider the class of hyper-planes in some dot product space  $H$ . The parameters  $w$  and  $b$  are described as follows:

$$(\langle w \cdot x \rangle + b) = 0, \text{ where } w \cdot H, b \in R \quad (1)$$

If a hyper-plane satisfaction Eq. (1) exists, then a linear separation is obtained. In this case,  $w$  and  $b$  can be rewritten as follows: Eq. (1) becomes

$$\max_{w \in H, b \in R} \min_{1 \leq i \leq m} \{ \|x - x_i\| \mid x \in H, \langle w \cdot x \rangle + b = 0, i = 1, \dots, m \} \quad (2)$$

Let the distance from the data point to the hyper-plane be  $1/\|w\|$ .

The separating hyper-planes include one optimal separating hyper-plane (OSH), which has the largest distance between two support vector points on its two sides. Since the distance between two support vector points is  $1/\|w\|^2$ , the minimal distance to OSH,  $\|w\|^2$ , can be derived from Eq. (2).

The margin  $2/\|w\|$  of a separating hyper-plane can be regarded as the hyper-plane's generalization ability, and the OSH has the maximal margin among separating hyper-planes.  $\|w\|^2$  can be minimized with Eq. (2) and Lagrange's polynomial. Let  $\alpha$  a denote  $(\alpha_1, \dots, \alpha_m)$ . Combining Lagrange's polynomial (in the order of  $m$ ) with Eq. (2) produces the following maximization equations:

$$W(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum \alpha_i \alpha_j y_i y_j \langle x_i x_j \rangle \quad (3)$$

$$\text{Subject to } \alpha_i \geq 0 \text{ for all } i = 1, \dots, m \quad (4)$$

$$\text{and } \sum_{i=1}^m \alpha_i y_i = 0.$$

Quadratic programming can be applied to solve this

maximization problem.

Given a vector satisfies Eq. (3) in maximization, the OSH expressed in terms of  $(w, b)$  can be written as follows:

$$w = \sum_{i=1}^m \alpha_i y_i x_i \quad (5)$$

where the support vector points must satisfy  $\alpha_i \geq 0$  and Eq. (2). When considering expanding in constraint Eq. (4), the determinant function of hyper-plane is written as follows:

$$f(x) = \text{sgn} \left( \sum_{i=1}^m y_i \alpha_i \langle x, x_i \rangle + b \right) = 0 \quad (6)$$

In most cases, the data is not linearly separable, and is thus mapped to higher dimensional feature space. This means that if the data cannot be classified explicitly in the current dimensional space, then the SVM maps the data to a higher dimensional space for classification.

The input data are mapped to a higher dimensional feature space by plotting a nonlinear curve. The OSH is built in the feature space. The feature space vectors  $x, x', k(x, x') = \langle x, x' \rangle$ .

The kernel function can be applied because all feature vectors occur only in dot products. The weight vector then becomes an expansion in feature space, and therefore typically no longer corresponds to the F-image of a single input space vector.

The decision function is expressed as follows:

$$\begin{aligned} f(x) &= \text{sgn} \left( \sum_{i=1}^m a_i y_i \langle \Phi(x), \Phi(x_i) + b \rangle \right) \\ &= \text{sgn} \left( \sum_{i=1}^m a_i y_i k(x_i, x_j) + b \right), \end{aligned} \quad (7)$$

The quadratic program is written as

$$\max_{\alpha \in R^m} W(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (8)$$

$$\text{Subject to } \alpha_i \geq 0 \text{ for all } i = 1, \dots, m \text{ and } \sum_{i=1}^m \alpha_i y_i = 0. \quad (9)$$

Several kernel functions help the SVM in finding the optimal solution. The most frequently used such functions are the polynomial kernel, sigmoid kernel and radial basis kernel function (RBF) [13]–[15]. The RBF is used most often in general linear kernel function. Furthermore, the RBF has fewer parameters than a polynomial kernel. Overall, performance of an RBF and other kernel functions are not significantly different. Consequently, the RBF is an effective choice for the kernel function. Therefore, this study employs an RBF kernel function in the SVM to discover the optimal solution.

Two major parameters of the RBF applied in SVM, C and U, have to be set appropriately. Parameter C is the cost of the penalty. The choice of value for parameter C influences the

classification outcome. If C is too large, then the classification accuracy rate is very high in the training stage, but very low in the testing stage. If C is too small, then the classification accuracy rate is unsatisfactory, making the model useless. Parameter U has a much stronger impact than parameter C on classification outcomes, because its value influences the partitioning outcome in the feature space. An excessive value for parameter U leads to over-fitting, while a disproportionately small value results in under-fitting.

Grid search [16], [17] is the simplest way to determine the values for parameters C and U. Sets of values for parameters C and U that produce the highest classification accuracy rate in this interval are found by setting the upper and lower limits (search interval) for parameters C and U and the jumping interval in the search. However, this type of search is a local search and prone to a local optimality. Additionally, setting the search interval is a problem. Too large a search interval wastes computing power, while too small a search interval might render a satisfactory outcome impossible. Methods for obtaining the optimal parameters in the SVM are currently still under development.

### C. Genetic Algorithms for Parameter Tuning of SVM

The genetic algorithm is type of an evolutionary computing. The genetic algorithms are inspired by theory of evolution [15]. In genetic algorithm, problems are solved by an evolutionary process resulting in a fittest solution. Rechenberg presented evolutionary computing in the 1960s. Then, Rechenberg's idea was developed by subsequent researchers. John Holland and his students presented the genetic algorithm (GA). The GA begins with a set of solutions which are represented by chromosomes. The chromosomes from one population are taken and used to form a new population. Each of chromosomes of this new population is called as offspring. The main aim of GA that the new population will be better than the old one. The new offsprings are selected according to their fitness. The chromosomes, whose fitnesses are the more suitable, have more chances for reproduce.

This iterative GA process is repeated until some conditions (for example, number of populations or improvement of the best solution, etc.) are satisfied. Outline of the basic genetic algorithm can be summarized below:

Step of starting: It is generated the random population of  $n$  chromosomes.

Step of fitness: It is evaluated the fitness  $f(x)$  of each chromosome  $X$  in the population

Step of new population: It is created a new population by repeating following steps until the new population is complete.

Step of selection: It is selected two parent chromosomes from a population according to their fitness.

Step of crossover: The crossover is performed between the parents to form new offspring with a crossover probability.

Step of mutation: The mutation is performed for new offspring with a mutation probability.

Step of accepting: It is placed new offspring in the new population.

Step of replace: The new generated population is used for a further run of the algorithm.

Step of test: The GA is stopped, if the end condition is satisfied. It is returned the best solution in current population.  
 Step of loop: It is gone to step of fitness.

### III. THE PROPOSED GA-SVM MODEL

In this section, it is described the design of proposed method based on GA-SVM for the automatic medical case study essay scoring. GA-SVM approach is a genetic algorithm based SVM classification technique. This approach selects of the optimal kernel function type and kernel function parameters of SVM classifiers.

The structure and algorithm of GA-SVM in Fig. 1 and Fig. 2 approach used in this study for the automatic medical case study essay scoring. In GA-SVM approach, it is selected of the optimal kernel function type and optimal kernel function parameter of support vector machine (SVM) classifiers. The structural and operational process of the GA-SVM is given below.

#### A. Data Preprocessing

The training essay data generation is shown by Fig. 1. We can select short medical case study essays into 3 fields of which 200 essays were problem, 416 fields were hypothesis, and 226 fields were learning objective. The essays are graded first by more than one human experts of that subject. The number of human graders may increase for the non-biased system. The average value of the human grades has been treated as training score of a particular training essay. The preprocessing has been done on training essay set.

#### B. Document Representation

We created the raw term frequency vectors of the essays prior to feeding them to the SVM. First, we using LexTo software for word segmentation, missed word filter and using Lexitron software for spell check and then computed frequency of words, compute weight of word, and created vector of training and testing essay scored.

#### C. Processing of the GA-SVM

##### 1) Composing of the initial population

The initial population of the genetic algorithm is consisted of 20 chromosomes. Each chromosome is consisted of 5 segments. The first segment of a chromosome represents kernel function types, which are linear, polynomial, radial basis function (RBF), sigmoid function. The second segment represents RBF kernel parameter. The third segment of a chromosome represents the value of polynomial kernel parameter. The fourth segment of a chromosome represents the value of sigmoid kernel parameter. The fifth segment of a chromosome represents the value of linear kernel parameter.

The kernel function type and values of kernel function parameters represented by this chromosome are sent into SVM classifier for realizing the training and testing of this combination of SVM classifier by using kernel function type and kernel functions parameter related with this chromosome.

##### 2) Calculation of the fitness function for a chromosome

The training of SVM classifier, which consists of 3 SVM

classifiers because of medical case study essays into 3 fields are used in this study, is performed by using training dataset. Second, classification accuracy of this trained SVM classifier is tested by using testing dataset.

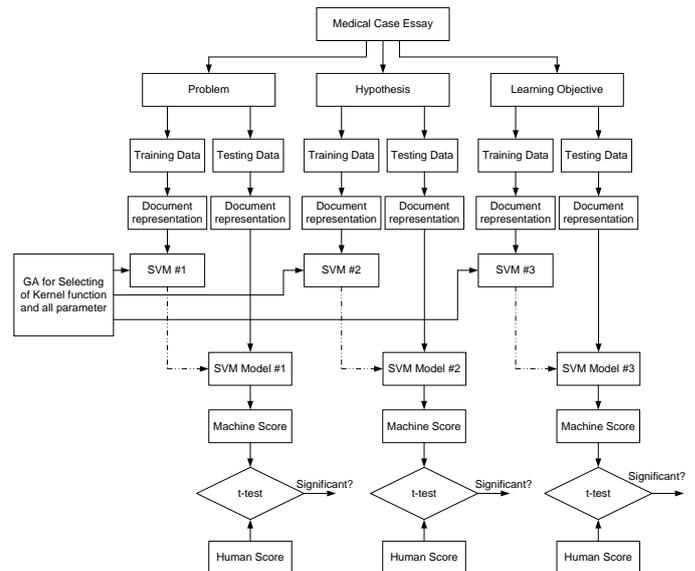


Fig. 1. A framework of medical case study essay scoring by GA-SVM.

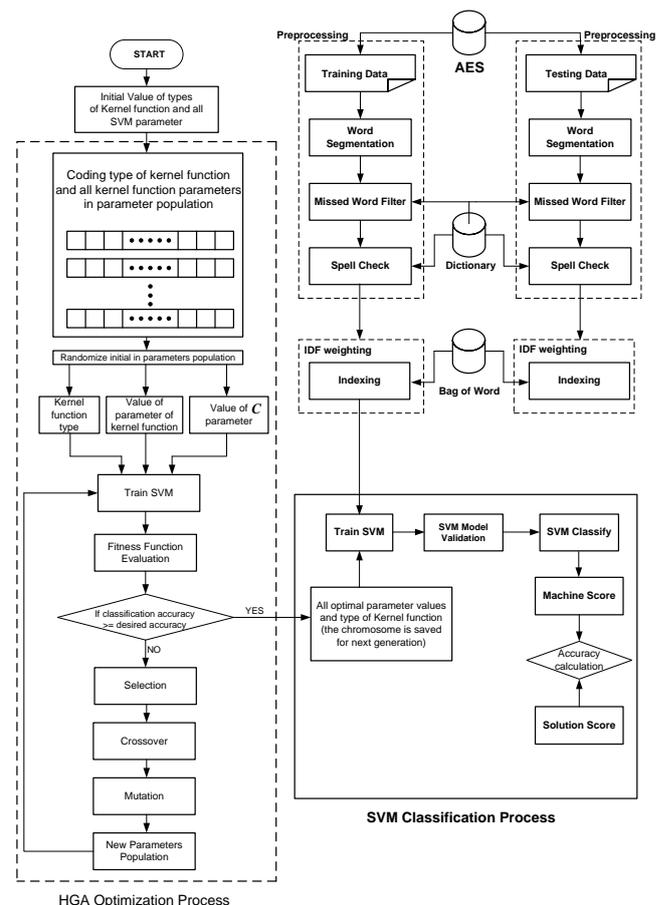


Fig. 2. The structure and algorithm of GA-SVM approach used for automatic medical case study essay scoring.

The fitness function used in GA-SVM is based on classification accuracy of this trained SVM classifier. The classification accuracy of a SVM classifier is calculated by using the classification rate ( $C$ ) denotes the percentage of correctly classified samples, which is computed by the following formula [18].

$$C = \frac{n_c}{n_i} \times 100\% , n_c \leq n_i \quad (12)$$

where  $n_c$ ,  $n_i$  represent the number of correctly classified samples and the total number of the samples, respectively.

3) *Crossover operation*

Two bits of the each of the random 2 chromosomes are randomly selected and replaced each other for crossover operations. At the final of the crossover operations, a new chromosome is obtained.

4) *Mutation operation*

In there, the bit inversion method is used for mutation operator [9]. If a bit is equal to 1, it is changed to 0. If it is equal to 0, it is changed to 1. At the final of the mutation operation, a new chromosome is obtained. There are total 20 chromosomes in population of next generation at the final of previous step and this step.

IV. EXPERIMENTAL DESIGN

The used platform is Linux (Ubuntu One) operating system. The development environment is NetBeans IDE 7.0.1 and Java programming. The software of SVM is LIBSVM [19].

A. *Data and Variables*

The sample was drawn from a subject matter of “Muscular systems and movement” which course in the faculty of medicine at PSU has taught courses in the pre-clinical medicine. 200 medical students of the same class level were asked to respond to a question on the subject matter by writing a short medical case study essay. This essay dataset contains 3 fields (see Fig. 1) of which 200 essays were problem and 416 essays were hypothesis and 226 essays were learning objective. The score of an essay has obtained by averaging the scores from three facilitators which has been treated as human score. Next, we performed word segmentation on the essays using the longest and maximal matching algorithm. Raw term frequency vector was then created for each essay.

B. *Experimental Results and Performance Analysis*

TABLE I: THE ESSAY SETS USED IN THE EXPERIMENTS

Set no.	Field	Training Essay	Testing Essay
1	Problem	200	40
2	Hypothesis	416	116
3	Learning Objective	226	76

We have practised our system by essay training dataset contains 3 fields (see Fig. 1) of which 200 essays were problems, 416 fields were hypotheses, and 226 fields were learning objectives. We have tested our model by essay testing dataset containing 3 fields of which 40 essays were problems and 116 essays were hypotheses and 76 essays were learning objectives. The score of each essay ranged from 1 point to 4 points, where a higher point represented a higher quality. The score of an essay has obtained by averaging the scores from three facilitators, which has been treated as human score. The numbers of essays corresponding to different scores are in the ranges of 0, 1, 2, 3,

and 4. Table I shows the summary of the essay set used in this experiment.

The optimal kernel function types and optimal values of kernel function parameters estimated by GA-SVM algorithm are given in Table II - Table IV.

TABLE II: THE OPTIMAL KERNEL FUNCTION TYPES AND OPTIMAL PARAMETERS OF SVM CLASSIFIER ESTIMATED BY GA-SVM APPROACH (PROBLEM ESSAY)

Optimal kernel function types	Optimal Values of Kernel function parameters				%C
	Cost	Gamma	Degree	Coef.	
Linear	-	-	-	-	
Polynomial	-	-	-	-	
RBF	1.0	0.0156	2.0	0.0156	67.5
Sigmoid	-	-	-	-	

TABLE III: THE OPTIMAL KERNEL FUNCTION TYPES AND OPTIMAL PARAMETERS OF SVM CLASSIFIER ESTIMATED BY GA-SVM APPROACH (HYPOTHESIS ESSAY)

Optimal kernel function types	Optimal Values of Kernel function parameters				%C
	Cost	Gamma	Degree	Coef.	
Linear	0.125	0.0078	4.0	0.063	77.58
Polynomial	-	-	-	-	
RBF	-	-	-	-	
Sigmoid	-	-	-	-	

TABLE IV: THE OPTIMAL KERNEL FUNCTION TYPES AND OPTIMAL PARAMETERS OF SVM CLASSIFIER ESTIMATED BY GA-SVM APPROACH (LEARNING OBJECTIVE ESSAY)

Optimal kernel function types	Optimal Values of Kernel function parameters				%C
	Cost	Gamma	Degree	Coef.	
Linear	-	-	-	-	
Polynomial	-	-	-	-	
RBF	64	0.0156	2.0	0.125	81.58
Sigmoid	-	-	-	-	

According to results in Table II and Table IV, RBF kernel function is the best of kernel function types used for medical essay. Table III shows that Linear kernel function is the best of kernel function types used for medical essay. In many SVM classification applications, the RBF kernel function is commonly preferred to other kernel function types [20]. The selection of the kernel function types and the values of used kernel function parameter are very important topics for classification problem performed by using SVM classifier.

In addition, we also computed the t-test which indicates that the scores by human and machine were significant at .05 level.

V. CONCLUSION

In this research, we proposed automatic medical case study essay scoring by GA-SVM. The optimal kernel function types and optimal values of kernel function parameters were estimated by GA-SVM algorithm.

These experimental results show that GA is used for choosing the kernel function type and its parameter values yield a proper solution to an optimization and obtain the machine scores. The experimental results also show that the addition of GA-SVM technique improves scoring performance.

REFERENCES

- [1] E. B. Page, "Grading essays by computer: Progress report," in *Proc. Notes from the 1966 Invitational Conference on Testing Problems*, 1966, pp. 87–100.
- [2] E. B. Page, "Computer grading of student prose: Using modern concepts and software," *Journal of Experimental Education*, vol. 62, no. 2, pp. 27–42, 1994.
- [3] T. K. Landauer, P. W. Holtz, and D. Laham, "Introduction to Latent Semantic Analysis," *Discourse Processes*, vol. 25, pp. 259–284, 1998.
- [4] J. Burstein. (May 1999). *Quoted in Ott. Essay questions. Salon*. [Online]. Available: <http://www.salonmag.com/tech/feature/1999/05/25/>, 1999.
- [5] T. Kakkonen, N. Myller, E. Sutinen, and J. Timonen, "Automatic Essay Grading with Probabilistic Latent Semantic Analysis," in *Proc. the 2nd Workshop on Building Educational Applications Using NLP*, Ann Arbor, June 2005, pp. 29–36.
- [6] C. Ramineni and D. M. Williamson, "Automated essay scoring: Psychometric guidelines and practices," *Assessing Writing*, vol. 18, no. 1, pp. 25–39, 2012.
- [7] L. M. Rudner and T. Liang, "Automated Essay Scoring Using Bayes Theorem," *The Journal of Technology, Learning, and Assessment*, vol. 1, no. 2, pp. 22, June 2002.
- [8] L. Bin, L. Jun *et al.*, "Automated Essay Scoring Using the KNN Algorithm," *International Conference on Computer Science and Software Engineering*, 2008, pp. 735–738.
- [9] V. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.
- [10] C. J. C. Burgers, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowledge Discov.*, vol. 2, pp. 121–167, 1998.
- [11] B. S. Lkoph and A. J. Smola, *Learning with Kernels*, London: MIT, 2002.
- [12] L. Zhang, L. B. Jack, and A. K. Nandi, "Fault detection using genetic programming," *Mech. Syst. Signal Process*, vol. 19, pp. 271–289, 2005.
- [13] Y. Liao, S.-C. Fang, and H. L. W. Nuttle, "A neural network model with bounded-weights for pattern classification," *Comput. Operat. Res.*, vol. 31, pp. 1411–1426, 2004.
- [14] H. T. Lin and C. J. Lin, "A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods," *Technical Report*, University of National Taiwan, Department of Computer Science and Information Engineering, pp. 1–32, March 2003.
- [15] K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, pp. 181–201, 2001.
- [16] J. Wang, X. Wu, and C. Zhang, "Support vector machines based on k-means clustering for real-time business intelligence systems," *Int. J. Business Intell. Data Mining*, vol. 1, pp. 54–64, 2005.
- [17] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," *Technical Report*, University of National Taiwan, Department of Computer Science and Information Engineering, pp. 1–12, July 2003.
- [18] Y. Huang, W. Li, and X. Ye, "A Study of Genetic Neural Network as Classifiers and its Application in Breast Cancer Diagnosis," *Journal of Computers*, vol. 6, no. 7, pp. 1319–1324, 2011.
- [19] C. C. Chang and C. J. Lin. (2001). LIBSVM: A library for support vector machines. [Online]. Available <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [20] J. A. F. Pierna, V. Baeten, A. M. Renier, R. P. Cogdill, and P. Dardenne, "Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds," *Journal of Chemometrics*, vol. 18, no. 7–8, pp. 341–349, 2004.



S. Yenaeng received the MS degree in computer education from King Mongkut's University of Technology North Bangkok (KMUTNB), Bangkok, Thailand in 2006. Currently, she is a computer education PhD candidate at KMUTNB. Her research interests include artificial intelligence, pattern recognition and machine learning. Currently, she is with the department of computer education, faculty of education, Bansomdejchaopraya Rajabhat University, Bangkok, Thailand.



S. Saelee received the Ph.D. degree in computer education from KMITNB, Thailand. His research interests include artificial intelligence, pattern recognition and machine learning. Currently, he is with the department of computer education, faculty of technical education, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.



W. Samai received the degree of medicine (MD), Diploma Thai Board of Forensic Medicine (DTBForensicMed). His research interests include artificial intelligence for medicine, pattern recognition and machine learning. Currently, he is with the faculty of medicine, Prince of Songkla University, Songkhla, Thailand.