GA-optimized Support Vector Regression for an Improved Emotional State Estimation Model

Hyunchul Ahn[1], Seongjin Kim[1], and Jae Kyeong Kim[2]
[1]Graduate School of Business IT, Kookmin University
Seoul, 136-702, Republic of Korea
[e-mail: hcahn@kookmin.ac.kr, scott_kim@kookmin.ac.kr]
[2]School of Management, Kyunghee University
Seoul, 130-701, Republic of Korea
[e-mail: jaek@khu.ac.kr]
*Corresponding author: Jae Kyeong Kim

Received March 14, 2014; revised May 2, 2014; accepted May 30, 2014; published June 27, 2014

Abstract

In order to implement interactive and personalized Web services properly, it is necessary to understand the tangible and intangible responses of the users and to recognize their emotional states. Recently, some studies have attempted to build emotional state estimation models based on facial expressions. Most of these studies have applied multiple regression analysis (MRA), artificial neural network (ANN), and support vector regression (SVR) as the prediction algorithm, but the prediction accuracies have been relatively low. In order to improve the prediction performance of the emotion prediction model, we propose a novel SVR model that is optimized using a genetic algorithm (GA). Our proposed algorithm—GASVR—is designed to optimize the kernel parameters and the feature subsets of SVRs in order to predict the levels of two aspects—valence and arousal—of the emotions of the users. In order to validate the usefulness of GASVR, we collected a real-world data set of facial responses and emotional states via a survey. We applied GASVR and other algorithms including MRA, ANN, and conventional SVR to the data set. Finally, we found that GASVR outperformed all of the comparative algorithms in the prediction of the valence and arousal levels.

Keywords: Emotional state estimation, Genetic Algorithm, Support Vector Regression

This research was supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2011.
A preliminary version of this paper was presented at ICONI 2013 and was selected as an outstanding paper.

http://dx.doi.org/10.3837/tiis.2014.06.014

1. Introduction

Affective computing has recently been gaining attention from researchers who are studying interactive and personalized Web services. Affective computing technologies deal with the systems, devices, and computers that can recognize, interpret, and express affective and emotional states [1]. These devices can improve human-computer interaction by enabling the communication of the users' emotional states [2].

In order to implement affective computing, it is necessary to understand the tangible and intangible responses of the users such as their speech, gestures, and facial expressions. From this wide array of responses, facial expressions have often been selected as the major source of information for estimating the users' emotional states [3][4][5].

Facial affect detection using emotional state estimation models may enhance user experiences by providing intelligent interaction with the users. For example, Thompson and McGill [2] proposed affective tutoring systems that enable e-learning applications to have the ability to detect and respond to the emotions that are experienced by the learner. Lin et al. [6] also proposed a learning emotional recognition model that enhances the students' understanding during distance learning courses. Jung and Kim [5] and Kim et al. [7] presented emotional state estimation models for implementing interactive exhibitions.

The process of facial affect detection using an emotional state estimation model requires the selection of an appropriate prediction algorithm. Various algorithms have been used for estimating emotional states, including multiple regression analysis (MRA), artificial neural network (ANN), and support vector regression (SVR). The SVR method has been used in many recent studies because of its high level of prediction accuracy [7][8][9].

SVR—a quantitative support vector machine (SVM) model—attempts to minimize the generalized error bound so as to achieve generalized performances [10]. Compared to ANN, it is known that SVR offers good prediction performances, even with a limited number of learning patterns [7]. In spite of these advantages, SVR and SVM are often criticized because their architectures are often determined by heuristic factors, such as the parameters of the kernel function and appropriate feature subsets [11]. However, optimization of these factors may lead to better results for emotional state estimation.

Given this background, our study proposes a novel algorithm, called Genetic Algorithm-based Support Vector Regression (GASVR), as the prediction algorithm for emotional state estimation. Our proposed algorithm uses a genetic algorithm (GA) to optimize the parameters and the feature subset selection for SVR. We use facial responses and the GASVR method to predict the two indicators of emotional state—the levels of valence and arousal.

The rest of this paper is organized as follows. In section 2, we present theoretical background information about emotional state estimation, SVR, and GA. In section 3, we describe the research model that is proposed in this paper and the procedures that are associated with it. Section 4 describes the empirical analysis that we used to validate the effectiveness of the proposed algorithm with a real-world data set. The conclusions and limitations of the study are discussed in the final section.

2. Theoretical Background

2.1 Emotional state estimation

Emotional state estimation plays an important role in affective computing. When implementing affective computing in practical situations, the computer systems should estimate the emotional states of humans in order to output an appropriate response for those emotions. As a result, researchers in the field of affective computing have studied methods for constructing effective emotional state estimation models.

Before building an emotional state estimation model, we first have to adopt a theoretical model that quantifies the levels of emotional states. Various emotional state models have been proposed in the field of psychology. We have adopted the V-A (valence-arousal) model, which is the most popular model for emotional state estimation.

The V-A model was proposed to measure the emotions felt by humans. It uses a two-dimensional approach [12][13]. The dimensions of the V-A model are valence and arousal. The valence dimension (V) represents how positive or negative the emotion is. It ranges from unpleasant to pleasant. The arousal dimension (A) refers to the level of excitement or apathy of the emotion and it ranges from sleepiness to frantic excitement [5][8][12][13]. **Fig. 1** shows the two dimensions of the V-A model and the positions of blended emotions.

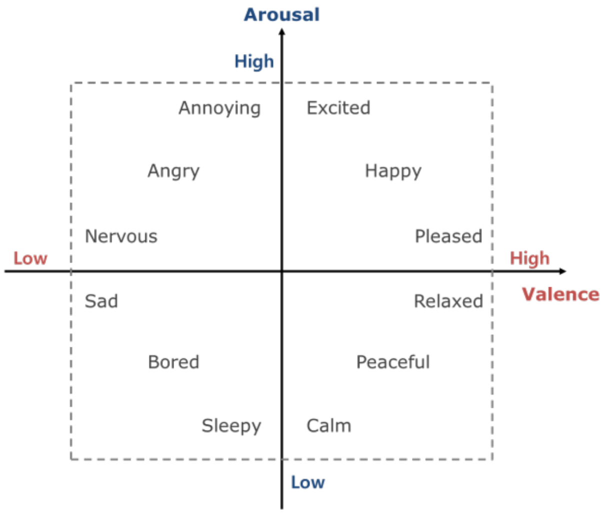

Fig. 1. V-A model.

2.2 Support Vector Regression

SVM is a statistical learning technique that was introduced by Vapnik [14]. SVM is a data mining algorithm that can be applied to classification (referred to as support vector classification—SVC) and prediction (referred to as support vector regression—SVR) [10]. SVM, which includes SVC and SVR, can lead to superior performances in practical applications because of its structural risk minimization principle, which is more general and superior to the empirical risk minimization principle that is adopted by conventional neural networks [15].

SVR, the regression model of SVM, is able to solve nonlinear estimation problems effectively. In order to extend SVM from classification to regression, Vapnik et al. [16]

adopted an ε-insensitivity loss function. In order to illustrate the concept of SVR, a typical regression problem will be formulated [15][17]. Let us assume that there is a set of data $\mathrm{G} = \{(\mathrm{x}_i, q_i)\}_{i=1}^{n}$, where x_i is a vector of the model's input features, q_i is the value of the output variable, and n is the number of data patterns. Then, the goal of the general regression analysis is to find a function $f(\mathrm{x})$ that is able to accurately predict the desired output values (q). A typical linear regression function can be derived as $q = f(\mathrm{x}) + \delta$, where δ is the random error with distribution of $N(0, \sigma^2)$.

In SVR, in order to solve nonlinear regression problems, the input features (x_i) are first mapped into a high-dimensional feature space (F) where they are correlated linearly with the outputs. Thus, SVR derives the following linear estimation function [15][18].

$$f(\mathbf{x}) = \big(\mathbf{v} \cdot \Phi(\mathbf{x})\big) + b \tag{1}$$

where v is the weight vector, b is the constant, and $\Phi(\mathrm{x})$ represents a mapping function in the feature space.

In SVR, the problem of nonlinear regression in the input space (x) is transformed into the problem of linear regression in a higher-dimensional feature space (F). Here, the robust ε-insensitive loss function (L_ε) that is presented in (2) is the most commonly used cost function [15][18].

$$L_\varepsilon(f(\mathbf{x}), q) = \begin{cases} |f(\mathbf{x}) - q| - \varepsilon & \text{if } |f(\mathbf{x}) - q| \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where ε denotes a precision parameter that represents the radius of the tube around the regression function $f(\mathrm{x})$ as shown in **Fig. 2** [15].

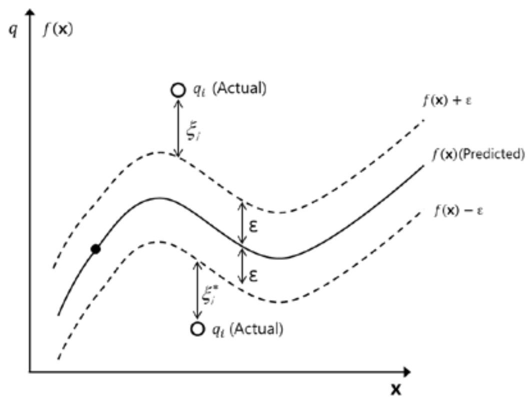

Fig. 2. SVR using ε-insensitive loss function (Adopted from [15]).

The weight vector (v) and the constant (b) in (1) can be estimated by minimizing the following regularized risk function [15][18].

$$\mathrm{R(C)} = \mathrm{C}\frac{1}{n}\sum_{i=1}^{n} L_\varepsilon(f(\mathbf{x}), q) + \frac{1}{2}|\mathbf{w}|^2 \tag{3}$$

where L_ε is the ε-insensitive loss function, $\frac{1}{2}|\mathrm{w}|^2$ is the regularization term that controls the trade-off between the complexity and the approximation accuracy of the regression model, and C is the regularization constant that is used to specify the trade-off between the empirical risk and the regularization term.

By adopting slack variables (ξ_i, ξ_i^*), (3) can be transformed into the following constrained form:

$$\textbf{Min. } R_{\text{reg}}(f) = \frac{1}{2}|\mathbf{w}|^2 + C\frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{4}$$

$$\textit{subject to} \quad \begin{cases} q_i - (\mathbf{w} \cdot \Phi(\mathbf{x}_i)) - b \le \varepsilon + \xi_i \\ (\mathbf{w} \cdot \Phi(\mathbf{x}_i)) + b - q_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \quad \text{for} \quad i = 1, \cdots, n \end{cases}$$

The application of Lagrangian multipliers and Karush-Kuhn-Tucker (KKT) conditions to (4) finally leads to the following general form of the SVR-based regression function [18]:

$$f(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}, \alpha, \alpha^*) = \sum_{i=1}^{n}(\alpha - \alpha^*)K(\mathbf{x}, \mathbf{x}_i) + b \tag{5}$$

where $K(\mathbf{x}, \mathbf{x}_i)$ denotes the kernel function.

Although there are several possible choices for the kernel functions, the Gaussian radial basis function (RBF) that is defined as $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2})$ is the most popular [15][17]. For further information on the theoretical background of SVR, it is recommended to refer to Lu et al. [15] and Vapnik [18].

2.3 GA and its application with SVR

GA is a simple, but effective optimization method that attempts to simulate biological evolution [19][20]. With the application of genetic operations such as selection, crossover, and mutation, it is designed to gradually improve the search results. In particular, there is a mechanism that is similar to mutation that prevents GA from falling into the local optima, and mechanisms like selection and crossover that enable efficient searches. Goldberg [21] provides more information about the evolution process and the genetic operators of GA.

So far, many researchers have applied GA for searching optimal architectural parameters of machine learning algorithms, including case-based reasoning (CBR) [22][23], artificial neural network (ANN) [24][25], and SVM (SVC) [26][27]. Additionally, some recent studies have attempted to use GA to optimize the architectural factors of SVR [28][29][30][31][32][33][34][35][36].

Chen [28], Chen and Wang [29], Ji et al. [30], and Liu et al. [31] used GA as the tool for optimizing the kernel parameters (C, σ, and ε) of SVR in order to forecast the target variable more effectively. Lahiri and Ghanta [32] and Wu et al. [33] tried to use GA to optimize both the kernel functions of SVR and their parameters simultaneously. All of these studies reported that SVR methods with kernel optimizations based on GA led to improved prediction accuracies as compared to conventional SVR methods.

Since Huang and Wang [26] proposed GA for feature selection and kernel parameter optimization for SVC, several recent studies, like He et al. [34], Oliveira et al. [35], and Huang [36], have tried to apply GA for feature selection and kernel parameter optimization for SVR. Huang [36] reported that optimal feature selection and kernel parameter selection both contributed significantly to the prediction accuracy of the algorithm. However, there have been few previous studies that have applied the optimization of the feature subsets and the kernel parameters of SVR to the estimation of emotional states. Thus, this study applies

GASVR, which optimizes both feature subsets and kernel parameters, to emotional state estimation.

3. Research Model

In this study, we propose a novel SVR algorithm that uses GA to optimize the feature selection and the kernel parameter settings simultaneously. Hereafter, we will call our algorithm GASVR (GA-optimized SVR). **Fig. 3** shows how the GASVR process works. The detailed explanation for each step of GASVR is as follows.

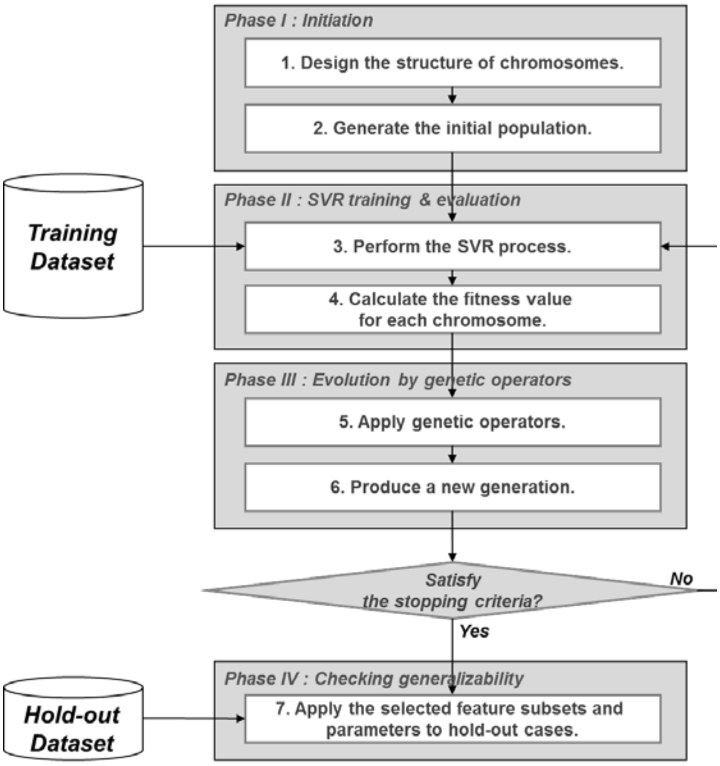

Fig. 3. Process of GASVR

3.1 Phase I: Initiation

In the first phase, the initial population of the GA search is initiated based on the structure of a chromosome. In order to apply the genetic operators of GA, the values that are encoded in a chromosome must be transformed into binary forms. In the case of GASVR, each chromosome should contain information about feature selection and kernel parameter settings. Feature selections can easily be encoded as binary strings since the values of the codes for selection are set to '0' or '1' where '1' signifies that the corresponding feature is selected and '0' signifies that it is not selected. For kernel parameters, the values should be converted to binary numbers. In this study, we assign 14 bits per parameter. Our study uses Gaussian RBF as the kernel function for SVR. Regarding the Gaussian RBF kernel function, Tay and Cao [37] reported that the upper bound C and the kernel parameter σ^2 are critical to the performance of SVM when using Gaussian RBF. Thus, the chromosome of GASVR is designed to optimize the two kernel parameters (C, σ^2) for Gaussian RBF and the precision parameter ε of the ε-insensitive loss function. Finally, the length of each chromosome becomes (m + 52) bits, where m is the number of features.

GASVR generates the initial population based on the chromosome structure that is described above. At this time, the values of the chromosomes in the population are initiated to random values.

3.2 Phase II: SVR training and evaluation

In the second phase, GASVR repeats a typical ε-SVR training process based on the assigned value in the chromosomes. Then it evaluates the fitness value of each chromosome. The main objective of the GA search in GASVR is to find the optimal or near optimal feature subsets and kernel parameters that lead to the most accurate predictions. From this perspective, we use the mean absolute error (MAE) of the training data set as the fitness function for GASVR.

3.3 Phase III: Evolution by genetic operators

In this phase, GASVR applies genetic operators, such as selection, crossover, and mutation, to the current population based on the fitness values that were generated in Phase II. As a result, a new generation of the population is created in this phase.

From this point, Phases II and III are iterated until the stopping conditions are satisfied. When the stopping conditions are satisfied, the chromosome that shows the best fitness value in the last population is selected, and the values of the optimal chromosome (i.e., optimal feature subset and kernel parameters) are finally determined.

3.4 Phase IV: Checking for Generalizability

The optimized feature subset and kernel parameters that are derived from Phase III generally fit well with the training data set because the GA search in the GASVR method is guided by the prediction accuracy of the training data set. However, GA searches may show disappointing performances when they are applied to unknown data sets because of overfitting. In order to avoid this type of danger, we apply the finally selected features and kernel parameters to the hold-out (unknown) data set in the last phase. In this way, we can check the general applicability of GASVR.

4. Empirical Validation

4.1 Data collection and preprocessing

We applied the proposed algorithm to a real-world data set in order to establish its validity. The data set used in this study was collected from the participants at the 2011-2012 Digital Media City (DMC) Culture Open Festival that was held in Seoul, Republic of Korea. At the festival, we ran a booth that displayed video stimuli to the participants via a large digital information display (DID), and captured their facial responses using a high-resolution camera. Funny, sad, and disgusting user created content (UCC) were used as the video stimuli. At the conclusion of the booth experience, we surveyed the participants about the emotional states that they experienced as they were exposed to the video stimuli. Their emotional state was measured in two dimensions—valence and arousal—using the 7-point Likert scale.

We collected 381 responses from a total of 244 participants. This study used 297 valid responses for the empirical validation. In order to quantify the facial responses of the participants, we extracted the coordinates of 64 facial points from their facial response images (refer to **Fig. 4**), and derived 35 facial feature values that were proposed by Pantic and

Rothkrantz [38]. These facial feature values were used as the candidate input variables for the prediction of the participants' emotional states. **Table 1** presents the candidate input variables for the emotional state estimation model in this study.

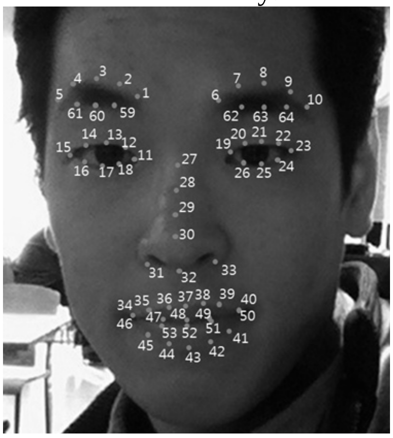

Fig. 4. The positions of sixty-four facial points used in this study

Table 1. Candidate input variables

No	Name	Description of variables
1	f1	angle ∠11,15,59
2	f2	angle ∠19,23,62
3	f3	distance 15-61
4	f4	distance 23-64
5	f5	distance A^{*}-13
6	f6	distance B^{**}-21
7	f7	distance A^{*}-7
8	f8	distance B^{**}-25
9	f9	distance 13-17
10	f10	distance 21-25
11	f11	distance 32-37
12	f12	distance 11-34
13	f13	distance 19-40
14	f14	distance 32-34
15	f15	distance 32-40
16	f16	distance 34-40
17	f17	distance 37-43
18	f18	distance 32-56
19	AU1	increased f1 & f2
20	AU2	increased f1 or f2
21	AU4	decreased f1 & f2
22	AU5	increased f5 & f6
23	AU7	decreased f7 or f8
24	AU10	decreased f11
25	AU12	decreased f12, decreased f13, increased f14, increased f15
26	AU13	decreased f12, decreased f13, decreased f14, decreased f15
27	AU15	increased f12 or f13
28	AU18	decreased f16
29	AU20	increased f16, non-increased f12, non-increased f13
30	AU23	decreased f17, non-decreased f16, non-increased f12, non-increased f13
31	AU24	decreased f17, decreased f16
32	AU25	non-increased f18, increased f17
33	AU26	f18 between two thresholds
34	AU28	f17 = 0
35	AU41	non-decreased f7, decreased f9, decreased f5 or decreased f10, decreased f6, non-decreased f8

* A: the middle point of 11 and 15
** B: the middle point of 19 and 23

In order to build a model for predicting the level of valence and arousal from these candidate input variables, we filtered the variables that did not have statistically significant correlations with the dependent variables (valence or arousal). In order to accomplish this, we used correlation analysis for the ratio input variables (f), and an independent samples t-test for the nominal input variables (AU). As a result, two ratio variables (f1 and f18) and six nominal variables (AU5, AU7, AU10, AU18, AU20, and AU41) were selected as the independent variables for the prediction of valence. Three ratio variable (f4, f11, f14) and seven nominal variables (AU1, AU2, AU13, AU15, AU18, AU20, and AU24) were selected as the independent variables for the prediction of arousal.

238 samples out of 297 samples (80%) were used as the training data set and 59 samples (20%) were used as the hold-out data set.

4.2 Experimental design

The experimental system for GASVR was developed using Microsoft EXCEL VBA (Visual Basic for Applications) and LIBSVM [39]. Palisade Software's Evolver 5.5, a commercial software package for implementing GA, was also used in our experimental system.

As mentioned in 3.1, Gaussian RBF was used as the kernel function for GASVR. The ranges of parameters C, σ^2, ε for GA search were set to [10 100], [1 100], and [0.05 0.5], respectively [7][15][37]. The population size was set to 100 organisms, and the crossover and mutation rates were set to 50% and 10%. As a stopping condition, 50 generations were permitted.

In order to test the effectiveness of GASVR, we also adopted three algorithms for comparison purposes – MRA, ANN, and conventional SVR with a grid search. The detailed descriptions of the settings for these algorithms that were used for comparison purposes were reported in our previous study [7].

4.3 Experimental results

As the criterion for evaluating the prediction performance, we adopted the mean absolute error (MAE) method that is shown in the following equation:

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| \tag{6}$$

where n is the number of samples, f_i is the predicted output, and y_i is the actual output.

Table 2 shows the features and the optimal kernel parameters that were finally selected and **Table 3** describes the prediction performances of GASVR and the other algorithms that were used for comparison purposes. As shown in **Table 2**, GASVR used only six variables for the prediction of valence and nine variables for the prediction of arousal. Although GASVR used fewer input variables than the other algorithms, it achieved the highest level of prediction accuracy for the hold-out data set in the prediction of both valence and arousal, as shown in **Table 3**.

Table 2. Result of GASVR

	Model for valence estimation	Model for arousal estimation
Selected features	f18, AU5, AU7, AU10, AU18, AU41	f11, f14, AU1, AU2, AU13, AU15, AU18, AU20, AU24
Optimal C	80.000	65.000
Optimal σ^2	80.000	71.283
Optimal ε	0.050	0.160

Table 3. MAEs of the algorithms

(a) Model for valence estimation

	MRA	ANN	SVR	GASVR
Optimal parameter values	-	# of nodes in hidden layer = 5	Linear C=78, ε=0.05	RBF C=80, σ^2=80 ε=0.05
Training	1.2034	1.0674	1.1806	1.0359
Test		0.9921		
Hold-out	1.3596	1.3527	1.3067	1.1189

(b) Model for arousal estimation

	MRA	ANN	SVR	GASVR
Optimal parameter values	-	# of nodes in hidden layer = 10	Linear C=1, ε=0.05	RBF C=65, σ^2=71.3 ε=0.16
Training	0.9433	0.9457	0.9323	0.8982
Test		0.9736		
Hold-out	0.8682	0.8298	0.8099	0.8057

A related-samples t-test was used to examine whether the prediction error of GASVR was significantly lower than that of the other algorithms. **Table 4** presents the results of the t-test. As shown in the table, GASVR was better than MRA and SVR at the 1% statistical significance level, and better than ANN at the 5% significance level in the prediction of valence. However, it did not outperform any of the other algorithms with statistical significance for the prediction of arousal. This may be interpreted that the idea of our model fits better to the estimation of valence rather than arousal. However, the experimental results would be understood in a limited context, as the insufficient number of hold-out samples (just 59 samples) might have been the cause of these disappointing results.

Table 4. Results of related-samples t-test

(a) Model for valence estimation

	ANN	SVR	GASVR
MRA	0.093	1.569	3.610**
ANN		0.644	2.533*
SVR			2.655**

(b) Model for arousal estimation

	ANN	SVR	GASVR
MRA	1.367	2.539*	1.544
ANN		0.615	0.524
SVR			0.100

* significant at the 5% level
** significant at the 1% level

5. Conclusion

In this paper, we suggested a new kind of hybrid SVR and GA model, named GASVR, in order to improve the prediction performance of the typical SVR algorithm for emotional state estimation. From the experimental results, we found that GASVR led to better estimation results when predicting valence levels or arousal levels from facial features.

From a practical perspective, the algorithm that was proposed in this study could be applied in various personalized Web services, including distance learning, content recommendations, and personalized advertisements on the Web. In particular, estimations of Web users' arousal levels will enable the system to respond properly when the users feel bored or uninterested.

The future research directions of this study are as follows. First, the empirical validation should be refined. In this study, we validated the proposed algorithm with a limited number of samples. As a result, it was not possible to observe statistically significant differences between GASVR and other algorithms during the predictions of arousal levels. Thus, in the future, we need to validate GASVR using larger samples.

Second, the optimization of SVR by GA can be extended to 'instance selection'. Some prior studies have indicated that proper selection of training instances might play a critical role in improving the prediction performances of support vector-based learning algorithms [11][27]. Thus, we expect that there will be studies about the application of proper instance selection to SVR in the near future.

Lastly, the proposed algorithm—GASVR—can be applied to many types of prediction / estimation problems, though we used it for emotional state estimation. Thus, it is recommended to endeavor to apply GASVR to other prediction domains for resolving business problems in the future.

References

[1] J. Tao, and T. Tan, "Affective Computing: A Review," *Lecture Notes in Computer Science*, vol. 3784, pp. 981-995, 2005. Article (CrossRef Link)

[2] N. Thompson, and T. J. McGill, "Affective Tutoring Systems: Enhancing e-Learning with the Emotional Awareness of a Human Tutor," *International Journal of Information and Communication Technology Education*, vol. 8, no. 4, pp. 75-89, October-December, 2012. Article (CrossRef Link)

[3] P. Ekman, "Universals and Cultural Differences in Facial Expression of Emotion," *Nebraska Symposium on Motivation*, vol. 19, pp. 207-283, 1971. http://www.ekmaninternational.com/ResearchFiles/Universals-And-Cultural-Differences-In-Facial-Expressions-Of.pdf

[4] P. Ekman, and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto, 1978.

[5] M.-K. Jung, and J.-K. Kim, "The Intelligent Determination Model of Audience Emotion for Implementing Personalized Exhibition," *Journal of Intelligence and Information Systems*, vol. 18, no. 1, pp. 39-57, March, 2012. http://jiisonline.evehost.co.kr/index.php?mnu=archive&archiveId=481

[6] K. C. Lin, T.-C. Huang, J. C.-S. Hung, N. Y. Yen, and S.-J. Chen, "Facial Emotion Recognition towards Affective Computing-based Learning," *Library Hi Tech*, vol. 31, no. 2, pp. 294-307, 2013. Article (CrossRef Link)

[7] S. Kim, E. Ryoo, M. K. Jung, J. K. Kim, and H. Ahn, "Application of Support Vector Regression for Improving the Performance of the Emotion Prediction Model," *Journal of Intelligence and Information Systems*, vol. 18, no. 3, pp. 185-202, September, 2012. http://jiisonline.evehost.co.kr/index.php?mnu=archive&archiveId=508

[8] M. Nicolaou, H. Gunes, and M. Pantic, "Continuous Prediction of Spontaneous Affect from

Multiple Cues and Modalities in Valence-Arousal Space," *IEEE Transactions on Affective Computing*, vol. 2, no. 2, pp. 92-105, April-June, 2011. Article (CrossRef Link)

[9] M. Nicolaou, H. Gunes, and M. Pantic, "Output-associative RVM regression for dimensional and continuous emotion prediction," *Image and Vision Computing*, vol. 30, no. 3, pp. 186-196, March, 2012. Article (CrossRef Link)

[10] D. Basak, S. Pal, and D. C. Patranabis, "Support Vector Regression," *Neural Information Processing – Letters and Reviews*, vol. 11, no. 10, pp. 203-224, October, 2007. http://bsrc.kaist.ac.kr/nip-lr/V11N10/V11N10P1-203-224.pdf

[11] H. Ahn, K. Lee, and K.-j. Kim, "Global Optimization of Support Vector Machines Using Genetic Algorithms for Bankruptcy Prediction," *Lecture Notes in Computer Science*, vol. 4234, pp. 420-429, 2006. Article (CrossRef Link)

[12] A. Mehrabian, and J. A. Russell, *An Approach to Environmental Psychology*, MIT Press, 1974.

[13] J. A. Russell, "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161-1178, December, 1980. Article (CrossRef Link)

[14] V. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.

[15] C. J. Lu, T. S. Lee, and C. C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decision Support Systems*, vol. 47, no. 2, pp. 115-125, May, 2009. Article (CrossRef Link)

[16] V. Vapnik, S. Golowich, and A. Smola, "Support vector method for function approximation regression estimation, and signal processing," In M. Mozer, M. Jordan, T. Petsche, editors, *Advances in Neural Information Processing Systems*, vol. 9, MIT Press, December, 1997. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.3139

[17] T. H. Hong, and E. M. Kim, "The Prediction of Purchase Amount of Customers Using Support Vector Regression with Separated Learning Method," *Journal of Intelligence and Information Systems*, vol. 16, no. 4, pp. 213-225, December, 2010. http://jiisonline.evehost.co.kr/index.php?mnu=archive&archiveId=429

[18] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 2000.

[19] Y.-H. Kim, and Y. Yoon, "Effect of Changing the Basis in Genetic Algorithms Using Binary Encoding," *KSII Transactions on Internet and Information Systems*, vol. 2, no. 4, pp. 184-193, August, 2008. Article (CrossRef Link)

[20] Y. Yang, W. Zheng, and S. Huang, "A Novel Automatic Block-based Multi-focus Image Fusion via Genetic Algorithm," *KSII Transactions on Internet and Information Systems*, vol. 7, no. 7, pp. 1671-1689, July, 2013. Article (CrossRef Link)

[21] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, New York, 1989.

[22] H. Ahn, and K.-j. Kim, "Global Optimization of Case-based Reasoning for Breast Cytology Diagnosis," *Expert Systems with Applications*, vol. 36, no. 1, pp. 724-734, January, 2009. Article (CrossRef Link)

[23] H. Ahn, and K.-j. Kim, "Bankruptcy Prediction Modeling with Hybrid Case-Based Reasoning and Genetic Algorithms Approach," *Applied Soft Computing*, vol. 9, no. 2, pp. 599-607, March, 2009. Article (CrossRef Link)

[24] K.-j. Kim, and H. Ahn, "Simultaneous optimization of artificial neural networks for financial forecasting," *Applied Intelligence*, vol. 36, no. 4, pp. 887-898, June, 2012. Article (CrossRef Link)

[25] S.-K. Oh, H.-S. Park, C.-W. Jeong, and S.-C. Joo, "GA-based Feed-forward Self-organizing Neural Network Architecture and Its Applications for Multi-variable Nonlinear Process Systems," *KSII Transactions on Internet and Information Systems*, vol. 3, no. 3, pp. 309-330, June, 2009. Article (CrossRef Link)

[26] C. L. Huang, and C. J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Systems with Applications*, vol. 31, no. 2, pp. 231-240, August, 2006. Article (CrossRef Link)

[27] K.-j. Kim, and H. Ahn, "Optimization of Support Vector Machines for Financial Forecasting," *Journal of Intelligence and Information Systems*, vol. 17, no. 4, pp. 241-254, December, 2011. http://jiisonline.evehost.co.kr/index.php?mnu=archive&archiveId=474

[28] K.-Y. Chen, "Forecasting systems reliability based on support vector regression with genetic algorithms," *Reliability Engineering & System Safety*, vol. 92, no. 4, pp. 423-432, April, 2007. Article (CrossRef Link)

[29] K.-Y. Chen, and C.-H. Wang "Support vector regression with genetic algorithms in forecasting tourism demand," *Tourism Management*, vol. 28, no. 1, pp. 215-226, February, 2007. Article (CrossRef Link)

[30] J. Huang, Y. Bo, and H. Wang, "Electromechanical equipment state forecasting based on genetic algorithm – support vector regression," *Expert Systems with Applications*, vol. 38, no. 7, pp. 8399-8402, July, 2011. Article (CrossRef Link)

[31] S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, and Y. Wei, "A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction," *Mathematical and Computer Modelling*, vol. 58, nos. 3-4, pp. 458-465, August, 2013. Article (CrossRef Link)

[32] S. K. Lahiri, and K. C. Ghanta, "Prediction of pressure drop of slurry flow in pipeline by hybrid support vector regression and genetic algorithm model," *Chinese Journal of Chemical Engineering*, vol. 16, no. 6, pp. 841-848, December, 2008. Article (CrossRef Link)

[33] C.-H. Wu, G.-H. Tzeng, and R.-H. Lin, "A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4725-4735, April, 2009. Article (CrossRef Link)

[34] W. He, Z. Wang, and H. Jiang, "Model optimizing and feature selecting for support vector regression in time series forecasting," *Neurocomputing*, vol. 72, nos. 1-3, pp. 600-611, December, 2008. Article (CrossRef Link)

[35] A. L. I. Oliveira, P. L. Braga, R. M. F. Lima, and M. L. Cornélio, "GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation," *Information and Software Technology*, vol. 52, no. 11, pp. 1155-1166, November, 2010. Article (CrossRef Link)

[36] C.-F. Huang, "A hybrid stock selection model using genetic algorithms and support vector regression," *Applied Soft Computing*, vol. 12, no. 2, pp. 807-818, February, 2012. Article (CrossRef Link)

[37] F. E. H. Tay, and L. Cao, "Application of support vector machines in financial time series forecasting," *Omega*, vol. 29, no. 4, pp. 309-317, August, 2001. Article (CrossRef Link)

[38] M. Pantic, and L. J. M. Rothkrantz, "Expert System for Automatic Analysis of Facial Expressions," *Image and Vision Computing*, vol. 18, no. 11, pp. 881-905, August, 2000. Article (CrossRef Link)

[39] C.-C. Chang, and C.-J. Lin, "LIBSVM : a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm. Article (CrossRef Link)

Hyunchul Ahn is an associate professor in the School of Management Information Systems at Kookmin University, Seoul, Korea. He has a master's degree and a Ph.D. in management engineering from the Korea Advanced Institute of Science and Technology (KAIST) Graduate School of Management. His research interests include technical issues on intelligent information systems for marketing and finance, and behavioral issues on the adoption of information systems. His works has been published in Annals of Operations Research, Applied Soft Computing, Computers & Operations Research, Computers in Human Behavior, Expert Systems with Applications, International Journal of Electronic Commerce, Information & Management, and Technological Forecasting and Social Change.

Seongjin Kim is currently enrolled in a master's program at Graduate School of Business IT, Kookmin University, where he also earned his B.A. degree in Management Information Systems. His primary research interests include data mining for business, Information Technology adoption and use, and Human-Computer Interaction.

Jae Kyeong Kim is a professor at School of Management, Kyunghee University. He obtained his M.S. and Ph.D. in Management Information Systems (MIS) from KAIST (Korea Advanced Institute of Science and Technology), and his B.S. in Industrial Engineering from Seoul National University. His current research interests focus on business intelligence, network management, and green business/IT. He has published numerous papers which have appeared in Artificial Intelligence Review, Electronic Commerce Research and Applications, European Journal of Operational Research, Expert Systems with Applications, Group Decision and Negotiations, IEEE Transactions on Services Computing, International Journal of Human-Computer Studies, International Journal of Information Management, and Technological Forecasting and Social Change. He is also a chief editor of JIIS (Journal of Intelligence and Information Systems), and an AE (associate editor) of Information Technology and Management (SSCI).