Journal of Marine Science and Technology

Volume 21 | Issue 3 Article 10

# APPLICATION OF A SUPPORT VECTOR MACHINE FOR LIQUEFACTION ASSESSMENT

Ching-Yinn Lee
*Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, R.O.C.*

Shuh-Gi Chern
*Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, R.O.C.*, sgchern@mail.ntou.edu.tw

Journal of Marine Science and Technology, Vol. 21, No. 3, pp. 318-324 (2013)
DOI: 10.6119/JMST-012-0518-3

# APPLICATION OF A SUPPORT VECTOR MACHINE FOR LIQUEFACTION ASSESSMENT

Ching-Yinn Lee and Shuh-Gi Chern

Key words: ANN, CPT, liquefaction, SVM.

## ABSTRACT

This study presents a support vector machine (SVM)-based approach for predicting earthquake liquefaction. The SVM model database includes five indexes: earthquake magnitude, total overburden pressure, effective overburden pressure, $q_c$ values from cone penetration tests (CPT), and peak ground acceleration. The proposed model was trained and tested on a dataset comprising 466 field liquefaction performance records and CPT measurements. A grid search method with *k*-fold cross-validation was also used to verify the feasibility. Compared with an artificial neural network (ANN)–based method, the SVM-based method has the advantage of increased accuracy and simpler operation. Experimental results show that the proposed SVM approach can increase the classification accuracy rate to a standard of 98.71%.

## I. INTRODUCTION

Liquefaction is one of the most destructive phenomena caused by earthquakes, and often occurs in loose, saturated soil deposits. Examples of liquefaction include the earthquakes in Niigata, 1964; Alaska, 1964; Tangshan, 1979; Loma Prieta, 1989; Kobe, 1995; Turkey, 1998; Chi-Chi, Taiwan, 1999; and Honshu, Japan, 2011. In view of the serious damage caused by earthquake-induced liquefaction, geotechnical engineers are actively engaged in the study of the soil liquefaction caused by earthquakes, and have developed many assessment methods for evaluating soil liquefaction. However, the high uncertainty in earthquake environments and soil characteristics makes it difficult to choose a suitable empirical equation for regression analysis. Therefore, many scholars and experts have attempted to develop scientifically derived analytical models that are simpler, easier to implement, and more accurate than traditional empirical equations for soil liquefaction analysis.

Many of the existing assessment methods were developed from observations of the behavior of sites during earthquakes. Geotechnical engineers have often used the simple liquefaction analytical model developed by SPT-N, because of its computational speed and analytical ability. Based on recent improvements in data processing and analytical ability, the cone penetration test (CPT) offers the advantage of fast, continuous, and accurate soil parameter measurements. Related testing data has also continued to accumulate. Therefore, the potential of applying CPT to liquefaction research has grown significantly.

This study presents a relatively new soft computing method known as a support vector machine (SVM) [1, 5]. SVMs have been widely used in recent years in areas such as image identification and facial recognition. An identification model that adopts SVM analysis is an effective method for accurately predicting liquefaction, and can be used in practical applications.

Previous studies have shown that the SVM method is a powerful and effective tool for dealing with liquefaction problems, and is more accurate and reliable than conventional methods [8, 10].

## II. OVER VIEW OF SVM

This section presents the basic SVM concepts for typical binary classification problems.

### 1. Linear SVM

A support vector machine, as presented by Vapnik [12], is a machine learning algorithm based on the statistical learning theory. The diagram in Fig. 1 shows the basic concepts of this approach. The circles and the diamonds in this figure represent two samples, and H is a labeling line separating the two samples. The H1 and H2 dashed lines pass through the nearest samples to the labeling line. The nearest data points used to define the margin are called support vectors (SV), and the distance between H1 and H2 is called the margin. The

*Paper Submitted 11/11/11; revised 04/14/12; accepted 05/18/12. Author for correspondence: Shuh-Gi Chern (e-mail: sgchern@mail.ntou.edu.tw).*
*Department of Harbor and River Engineering, National Taiwan Ocean University, Keelung, Taiwan, R.O.C.*

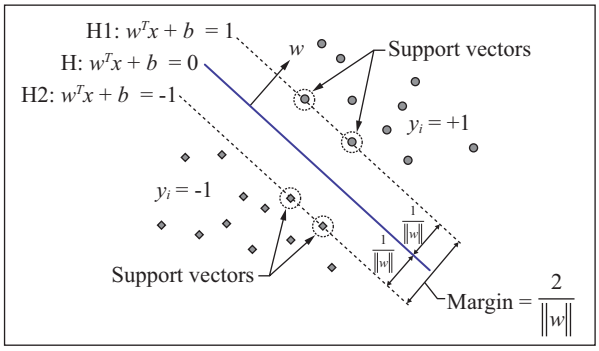


**Fig. 1. Optimal hyperplane for a linear SVM.**

separating hyperplane H that has the maximum distance between the nearest data (i.e., the maximum margin) is called the optimal separating hyperplane.

As Fig. 1 shows, the data patterns can be shown as $\{x_i, y_i\}$, $i$ = 1, 2, ……, $k$, where $x_i \in R^N$ is an N-dimensional data vector with each sample belonging to either of the two classes labeled as $y_i \in \{-1, +1\}$, and the decision function (hyperplane) can be expressed as

$$w^T x + b = 0 \tag{1}$$

where $x$ is an input vector, $w$ is an adaptive weight vector, $b$ is a bias, and $w^Tx$ is an inner product of $w$ and $x$. For the linearly separable class, a separating hyperplane for the two classes can be defined as

$$w^T x_i + b \geq 1, \quad \forall y_i = +1 \tag{2}$$

$$w^T x_i + b \leq -1, \quad \forall\ y_i = -1 \tag{3}$$

Eqs. (2) and (3) can be combined into

$$y_i(w^T x_i + b) - 1 \geq 0 \tag{4}$$

The goal of the SVM is to find $w$ and $b$ for the optimal separating hyperplane to maximize the margin $2/\|w\|$ (Fig. 1). Hence, the hyperplane that optimally separates the data is the one that minimizes $\|w\|$. The optimal separating hyperplane can be obtained by solving the following convex quadratic optimization problem [12]:

$$\text{Minimize} \qquad \frac{1}{2} w^T w = \frac{1}{2}\|w\|^2 \tag{5}$$

$$\text{subject to} \qquad y_i(w^T x_i) + b \geq 1, \forall i \tag{6}$$

The above equation can be transformed into the equivalent Lagrangian dual problem as

$$L_P(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i \left[ y_i(w^T x_i + b) - 1 \right]$$

$$= \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i y_i w^T x_i - b \sum_{i=1}^{N} \alpha_i y_i + \sum_{i=1}^{N} \alpha_i \tag{7}$$

where $\alpha_i \geq 0$ ($i$ = 1, 2, …, $N$) are the Lagrangian multipliers. The goal here is to find $w$ and $b$ which minimizes, and the $\alpha$ which maximizes Eq. (7). This can be done by differentiating $L_p$ with respect to $w$ and $b$ and setting the derivatives to zero

$$\frac{\partial L(w, b, \alpha)}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{N} \alpha_i y_i x_i \tag{8}$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = 0 \Rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{9}$$

Based on Eq. (9), the third term on the right hand side of Eq. (7) is zero. Multiplying Eq. (8) by $w^T$ leads to

$$w^T w = \sum_{i=1}^{N} \alpha_i y_i w^T x_i = \sum_{i=1}^{N} \sum_{j-1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_j \tag{10}$$

Eq. (7) can then be reformulated as

$$L_D(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j-1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_j \tag{11}$$

$$\text{subject to} \quad (1) \quad \sum_{i=1}^{N} \alpha_i y_i = 0$$

$$(2) \quad \alpha_i \geq 0, \forall i \tag{12}$$

The problem is now re-cast as finding the optimum Lagrangian multipliers that maximize the objective function Eq. (11) subject to Eq. (12). This is a convex quadratic optimization problem, and requires a quadratic program (QP) solver that returns $\alpha_t$. The solution $\alpha_t$ for the dual optimization problem determines the parameter $w^*$ and $b^*$ of the optimal hyperplane. Thus, the optimal hyperplane decision function can be written as

$$f(x) = sign(w^{*T} x_i + b^*) = sign\left( \sum_{i=1}^{N} \alpha_i^* y_i x_i^T x_j + b^* \right) \tag{13}$$

where ***sign*** is the signum function. If the result is positive, then it is classified $x$ as class 1, and classified as class 2 otherwise.

## 2. Linearly Inseparable SVM

The soft margin method, which introduces an additional cost function associated with misclassification, is an appro-

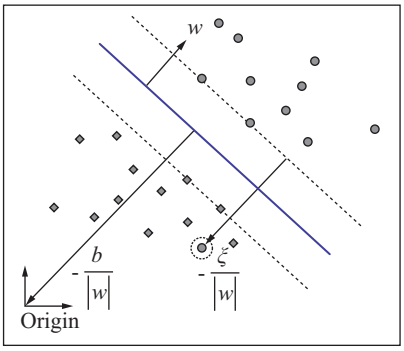


**Fig. 2. Hyperplane through two linearly inseparab classes.**

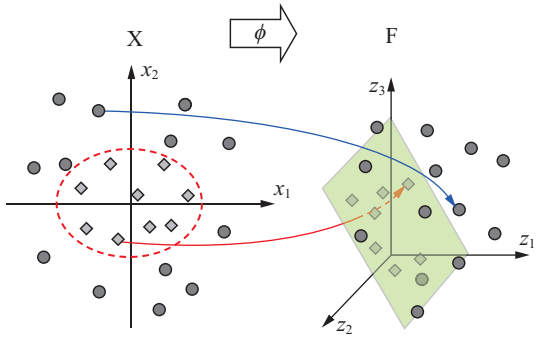


**Fig. 3. Mapping from the data space X to the feature space F.**

priate way to extend the SVM methodology to data that is not linearly separable. Cortes and Vapnik [5] introduced positive slack variables $\xi$ and a penalty factor $C$.

As Fig. 2 shows, data points on the incorrect side of the margin boundary have a penalty that increases with distance. To reduce the number of misclassifications, modify the constraints of Eq. (5) for the non-separable case as follows:

$$\text{minimize} \qquad \frac{1}{2}w^T w + C\sum_{i=1}^{l}\xi_i \tag{14}$$

$$\text{subject to} \qquad y_i(w^T x_i) + b - 1 + \xi_i \geq 0\,,\ \forall i \tag{15}$$

where $\xi$ is called a slack variable used to account for the effects of misclassification. $C$ is called a penalty factor, a parameter defines the trade-off between the number of misclassification in the training data and margin maximization. As before, reformulating this as a Lagrangian requires the minimization of $w$, $b$, and $\xi$, and the maximization of $\alpha$ (where $\alpha_i \geq 0$):

$$L(w, b, \xi, \alpha, \beta) = \frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left[y_i(w^T x_i + b) - 1 + \xi_i\right] + \sum_{i=1}^{N}\beta\xi_i \tag{16}$$

$$\text{subject to} \qquad \alpha_i, \beta_i \geq 0\ (i = 1, 2, ..., N) \tag{17}$$

Differentiating $L$ with $w$, $b$, and $\xi$, and setting the derivatives to zero leads to

$$\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{N}\alpha_i y_i x_i \tag{18}$$

$$\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial b} = 0 \Rightarrow \sum_{i=1}^{N}\alpha_i y_i = 0 \tag{19}$$

$$\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial \xi_i} = 0 \Rightarrow C - \alpha_i - \beta_i = 0 \tag{20}$$

After substituting these values in, $L_D$ has the same form as Eq. (10), Eq. (11). Again, maximize

$$L_D(w, b, \xi, \alpha, \beta) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j-1}^{N}\alpha_i\alpha_j y_i y_j x_i^T x_j \tag{21}$$

$$\text{subject to} \quad (1) \qquad \sum_{i=1}^{N}\alpha_i y_i = 0$$

$$(2) \qquad 0 \leq \alpha_i \leq C, \forall i \tag{22}$$

The equations are almost the same dual problem as before, with a slight difference being that the multipliers $\alpha_i$ have an extra constraint.

## 3. Nonlinear Separable SVM

The concepts can also be extended to the case of a nonlinear separating hyperplane by mapping the input space onto a high dimensional space, $x \to \phi(x)$, where the data can be linearly classified (Fig. 3). The key property of this mapping is that the function $\phi$ must be subject to the condition that the dot product of the two functions $\phi(x_i) \cdot \phi(x_j)$ can be written as a kernel function $K(x_i, x_j)$ The decision function then becomes

$$f(x) = \sum_{i=1}^{N} y_i\alpha_i K(x_i, x_j) + b \tag{23}$$

Different kernel functions can construct various learning machines. Some typical kernel functions are as follows:

$$\text{Linear kernel:} \qquad K(x_i, x_j) = x_i^T x_j \tag{24}$$

$$\text{Polynomial kernel: } K(x_i, x_j) = (\gamma x_i^T x_j + r)^d\,,\ \gamma > 0 \tag{25}$$

Radial basis function (RBF):

$$K(x_i, x_j) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right),\ \gamma > 0 \tag{26}$$

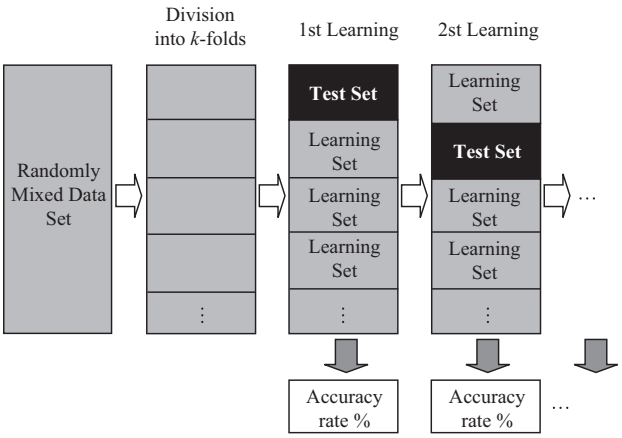


**Fig. 4. A *k*-fold cross-validation procedure.**

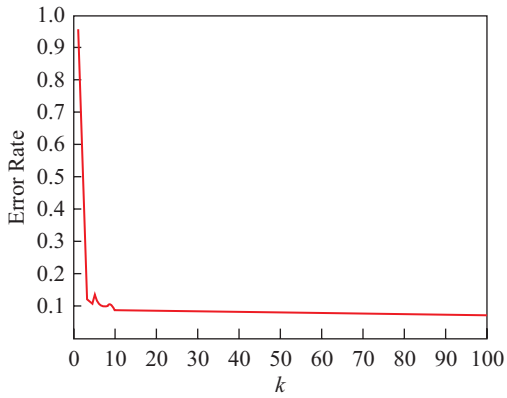


**Fig. 5. A plot of *k*-fold cross-validation error *vs.* *k* [11].**

Sigmoid kernel: $K(x_i, x_j) = \tan(\gamma x_i^T x_j + r)$ (27)

In the questions above, $\gamma$, $r$ and $d$ are kernel parameters.

## III. CROSS-VALIDATION

Cross-validation is a technique for assessing how the results of statistical analysis can be generalized to an independent dataset. This technique is mainly used in situations where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice.

This study adopts a *k*-fold cross-validation technique that randomly partitions the original sample into *k* subsamples. A single subsample is retained as validation data for testing the model, and the remaining $k - 1$ subsamples are used as training data. The cross-validation process is repeated *k* times (the folds), with each of the *k* subsamples used exactly once as the validation data. The k results from the folds can be averaged (or otherwise combined) to produce a single estimation. Fig. 4 provides an example of a *k*-fold cross-validation procedure.

The advantage of this method over repeated random subsampling is that all observations are used for both training and validation, and each observation is used for validation exactly once. The main drawback of this method is that it requires intense computation. Fig. 5 shows the *k*-fold cross-validation error versus *k* for a big data set, and indicates that a *k* value between 4 and 10 is a good trade-off: increasing this value significantly increases computation time and does not significantly improve results [11]. Thus, this study adopts 5-fold cross-validation. This approach may not be useful in achieving high training accuracy, but it can prevent the over-fitting problem.

## IV. GRID SEARCH

The grid search algorithm performs an exhaustive search through the parameter space of a learning algorithm to solve the problem of model selection (i.e., finding the optimal parameters for a dataset).

Researchers have proposed four basic kernel functions for SVM models. First, decide which one to use, and then choose the penalty $C$ and kernel parameters. For example, there are two parameters for an RBF kernel: $C$ and $\gamma$. Various pairs of ($C$, $\gamma$) values are tried with a grid search procedure and the one with the best cross-validation accuracy is chosen. Testing exponentially growing sequences of $C$ and $\gamma$ is a practical method for identifying good parameters (e.g., $C = 2^{-4}$, $2^{-3.5}$, …, $2^4$; $\gamma = 2^{-4}$, $2^{-3.5}$, …, $2^4$).

## V. APPLICATIONS OF SVM CLASSIFICATION

The case records in this study were evaluated using the MATLAB (R2010a) program and tool box [2, 6]. Fig. 6 shows the flowchart of the proposed SVM system.

The database includes 466 CPT-based field liquefaction records from over 11 major earthquakes between 1964 and 1999. The data consists of 21 case records from Japan, 85 from China, 7 from Canada, 219 from the USA, and 134 from Taiwan. This represents 250 sites that liquefied and 216 sites that did not liquefy. Five parameters that were recorded in all 466 sites are (1) earthquake magnitude, $M$; (2) total overburden pressure, $\sigma_0$; (3) effective overburden pressure, $\sigma_0'$; (4) $q_c$ values from CPTs; and (5) peak acceleration, $a_{max}$ Table 1 summarizes the maximum and minimum values of each parameter. The parameter values for all 466 case records are presented in a paper written by Chern *et al.* [3]. The input representing the liquefaction potential is given a binary value of 1 for liquefied sites and a value of −1 for non-liquefied sites.

Before the datasets were used to train the SVM model, they were preprocessed using Eq. (28). Each parameter is normalized between 0 and 1, with

$$y = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (28)$$

in which $y$ is a normalized input parameter, $x$ is the original input parameter, and $x_{max}$ and $x_{min}$ are the maximum and minimum parameters, respectively.

**Table 1. The maximum and minimum values of the reference datasets.**

| Parameter | M | $\sigma_0$(kpa) | $\sigma_0$(kpa) | $q_c$(kpa) | $a_{max}$(kpa) |
|---|---|---|---|---|---|
| Max. | 7.8 | 364.5 | 227.5 | 25.0 | 0.8 |
| Min. | 5.9 | 16.7 | 16.7 | 0.18 | 0.08 |

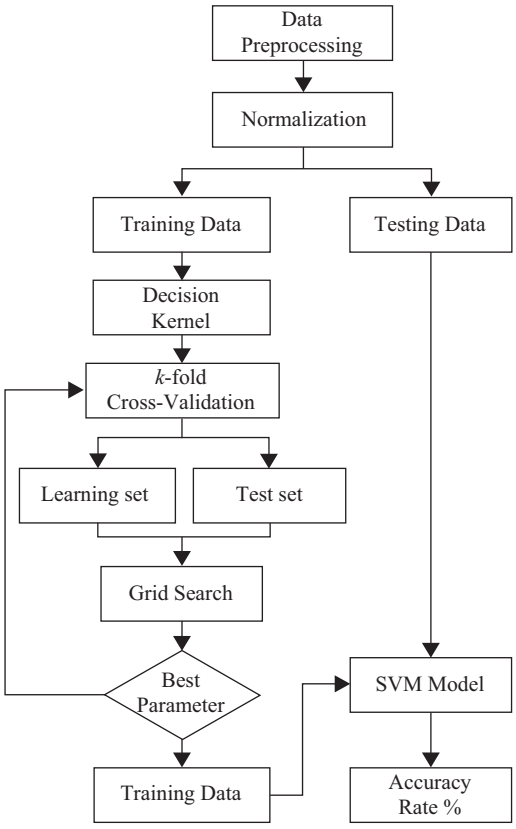


**Fig. 6. Flowchart of the proposed SVM system.**

The main advantage of scaling is to avoid attributes with greater numeric ranges dominating those with smaller numeric ranges. Another advantage of this method is to avoid numerical difficulties during calculation.

A normal SVM model randomly selects kernel parameters using a trial-and-error method [8, 10]. The grid search approach in this study is an alternative way to find the best parameters for the SVM classifier. This approach avoids the over-fitting problem of the SVM model occurring because of the improper determination of these parameters.

The RBF kernel is a reasonable first choice for an SVM model [9]. Hence, the proposed SVM model was first constructed by a radial basis function (RBF) kernel. There are two parameters, $C$ and $\gamma$, to be determined. After the grid search procedure, the optimal parameters with maximal classification accuracy were selected. As shown in Fig. 7, the best ($C$, $\gamma$) is ($2^{2.5}$, $2^5$) with a cross-validation rate of 95.279%. In this result, the optimal parameters are used to train the SVM model to generate the final classifier.

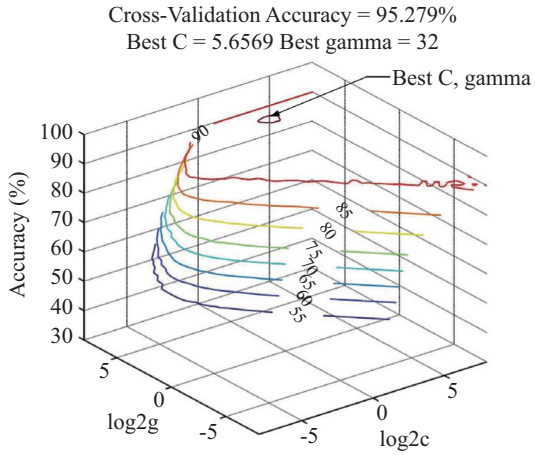


**Fig. 7. Parameters $C$ and $\gamma$ versus the accuracy rate.**

To improve SVM classification accuracy, the grid search procedure plays an important role in the performance of the SVM. Fig. 7 also shows that the parameters $C$ and $\gamma$ greatly affect the classification accuracy of the SVM.

## VI. RESULTS AND DISCUSSION

The procedure for using the SVM is described below:

1. Transform data to the format of the SVM package.
2. Conduct simple scaling on the data.
3. Consider the RBF kernel.
4. Use cross-validation and grid searching to find the best parameters $C$ and $\gamma$.
5. Use the best $C$ and $\gamma$ to train the whole training set.
6. Test.
7. Find the best accuracy rate.

After the training procedure, the best ($C$, $\gamma$) is ($2^{2.5}$, $2^5$) with a cross-validation rate of 95.279%. Out of the 466 datasets used, only 6 cases were misclassified, achieving an overall classification accuracy rate of 98.71%.

In addition to verifying the effectiveness of the proposed method, this study compares it with an ANN method in the reference [3, 4]. The ANN model proposed in that paper combines fuzzy theory with a subtractive clustering algorithm to form a fuzzy-neural network system. To verify the feasibility of the ANN model, this study compares that ANN model with the B5 model employed in Goh [7] using the same 109 data groups, including 74 training data groups and 35 test data groups. The results of this analysis are presented in Table 2. The ANN-G5 model [3] performs better than Goh's optimal B5 model in both the training and testing segments. Therefore, the 466 collected CPT datasets are used in this study to compare the SVM model with the ANN (C4, C4H6, C5, and C5N) models [3]. Results are listed in Table 3, it shows that the SVM model achieves better results than the ANN models because of its lower total error rate of 1.29%.

**Table 2. Result of the G5 model and B5 model [7].**

| Model | Input variables | No. of elements in every training cluster | No. of hidden neurons | No. of Error | | Total error rate (%) |
|---|---|---|---|---|---|---|
| | | | | Training | Testing | |
| B5 | $M$, $D_{50}$, $\sigma_0'$, $q_c$, $a_{max}$ | – | 5 | 1 | 2 | 2.75 |
| G5 | $M$, $D_{50}$, $\sigma_0'$, $q_c$, $a_{max}$ | 53, 23, 11 | 5 | 0 | 1 | 0.92 |

**Table 3. Comparison between SVM and ANN models [3].**

| Model | Input variables | No. of elements in every training cluster | No. of hidden neurons | No. of Misclassified | Total error rate (%) |
|---|---|---|---|---|---|
| C4 | $M$, $\sigma_0'$, $q_c$, $a_{max}$ | 217,116,82,79 | 5 | 20 | 4.29 |
| C4H6 | $M$, $\sigma_0'$, $q_c$, $a_{max}$ | 217,116,82,79 | 6 | 19 | 4.08 |
| C5 | $M$, $\sigma_0$, $\sigma_0'$, $q_c$, $a_{max}$ | 190,93,114,89 | 5 | 12 | 2.58 |
| C5N | $M$, $\sigma_0$, $\sigma_0'$, $q_{c1N}$, $a_{max}$ | 190,91,113,91 | 5 | 16 | 3.43 |
| SVM | $M$, $\sigma_0$, $\sigma_0'$, $q_c$, $a_{max}$ | – | – | 6 | 1.29 |

**Table 4. The classification accuracies versus $C$ for different kernel functions.**

| C | $10^{-2}$ | $10^{-1}$ | 2 | $2^{2.25}$ | $2^5$ | 100 | 200 |
|---|---|---|---|---|---|---|---|
| Linear | 53.65 | 82.40 | 90.99 | 90.99 | 91.20 | 90.99 | 91.20 |
| Poly | 53.65 | 53.65 | 72.10 | 76.61 | 85.84 | 87.98 | 88.84 |
| RBF | 53.65 | 93.99 | 98.07 | 98.71 | 97.65 | 96.53 | 96.53 |
| Sigmoid | 53.65 | 58.80 | 89.06 | 90.77 | 89.70 | 86.70 | 84.55 |

As indicated previously, there are four types of basic kernel functions: linear, RBF, second order polynomial, and sigmoid. This study employs the accuracy rate as a criterion to find the optimal kernel function. Table 4 shows the accuracy rate versus the $C$ parameter from 0.01 to 200 for different kernel functions. The RBF kernel function with parameter $C = 2^{2.25}$ provides the best performance for the SVM model.

The excellent classification accuracy of an SVM suggests its practicality for engineering applications. Therefore, this study develops a liquefaction assessment algorithm based on SVM theory, called LA-SVM. The graphical user interface (GUI) of this algorithm was implemented in a MATLAB/GUI. This interface provides an intuitive and user-friendly means of interaction. Users do not need any diagrams, formulae, or manuals. By simply using a mouse cursor to select options and input training data and parameter ranges, they can receive the classification results and accuracies of the testing data in a short CPU runtime. LA-SVM greatly simplifies the liquefaction assessment process and produces extremely accurate results. The operation steps are listed as follows:

1. Launch LA-SVM program (Fig. 8).
2. Select the input button, and input the training data and testing data.
3. Select the data's normalization and its ranges (specified by users).
4. Select the grid search method and specify the parameter ranges.

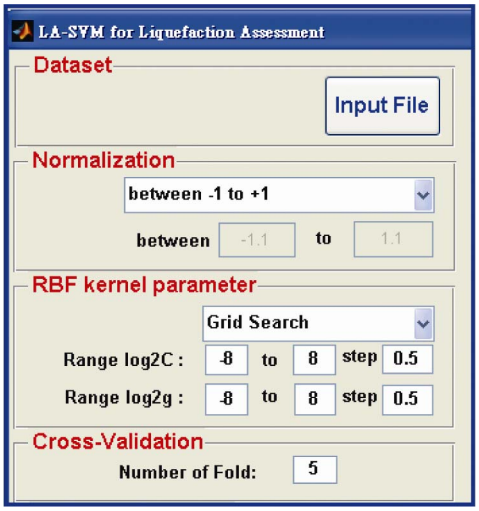


**Fig. 8. Easy to use LA-SVM/GUI interface.**

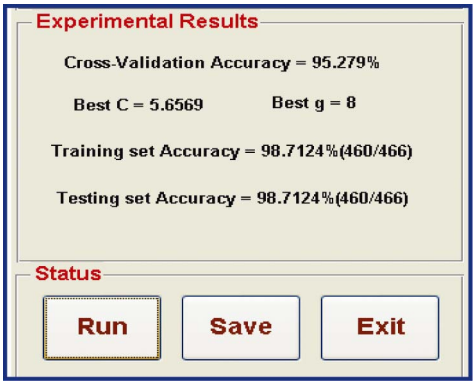


**Fig. 9. Experimental Result of LA-SVM.**

5. Specify the number of folds for cross-validation.
6. Select the run button, and start the analysis.
7. When analysis is complete, obtain the optimal kernel function parameters and the classification accuracy (Fig. 9).

## VII. CONCLUSION

SVM has been successfully applied in many applications, but it is less widely applied in the geotechnical field. The results in this study show that SVM is a powerful computational tool that can be used to analyze the complex relationship between soil and seismic parameters in liquefaction assessment.

The experimental results in this study indicate that an SVM achieves greater classification accuracy than an ANN. In addition to its higher accuracy rate, the SVM model requires only two parameters, as compared to the ANN, which requires multiple parameters. In conclusion, the SVM model is more effective and feasible than the conventional ANN. An SVM not only has a solid foundation in statistical learning theory, but can also effectively handle nonlinear classification. Therefore, it is regarded as one of the most effective classification methods.

The experimental results and discussion above show that the proposed LA-SVM can be effectively applied to liquefaction assessment. The LA-SVM program has an intuitive interface that is easily understandable. As new liquefaction assessment cases are collected to expand the database, the classification accuracy of LA-SVM can be further increased. Thus, LA-SVM is a novel liquefaction assessment tool worthy of promotion and support.

## REFERENCES

1. Boser, B. E., Guyon, I. M., and Vapnik, V. N., "A training algorithm for optimal margin classiers," *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ACM Press, pp. 144-152 (1992).
2. Chang, C. C. and Lin, C. J., *LIBSVM: A Library for Support Vector Machines*, http://www.csie.ntu.edu.tw/~cjlin/libsvm (2001).
3. Chern, S. G., Lee, C. Y., and Wang, C. C., "CPT-based liquefaction assessment by using fuzzy-neural network," *Journal of Marine Science and Technology*, Vol. 16, No. 2, pp. 139-148 (2008).
4. Chern, S. G., Lee, C. Y., and Wang, C. C., "CPT-based simplified liquefaction assessment by using fuzzy-neural network," *Journal of Marine Science and Technology*, Vol. 17, No. 4, pp. 326-331 (2009).
5. Cortes, C. and Vapnik, V., "Support-vector network," *Machine Learning*, Vol. 20, pp. 273-297 (1995).
6. Faruto and Liyang, LIBSVM-faruto Uliti- mate Version, a toolbox with implement for support vector machine based on libsvm, http://www.matlabsky.com (2011).
7. Goh, A. T. C., "Neural network modeling of CPT seismic liquefaction data," *Journal of Geotechnical Engineering*, ASCE, Vol. 129, No. 1, pp. 70-73 (1996).
8. Goh, A. T. C. and Goh, S. H., "Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data," *Computer and Geotechnics*, Vol. 34, pp. 410-421 (2007).
9. Hsu, C. C., Chang, C. C., and Lin, C. J., *A Practical Guide to Support Vector Classification*, Taipei (2010).
10. Samui, P. and Sitharam, T. G., "Machine learning modeling for predicting soil liquefaction susceptibility," *Journal of Natural Hazard and Earth System Science*, Vol. 11, pp. 1-9 (2011).
11. Sterlin, P., Overfitting Prevention with Cross-Validation, Paris (2007).
12. Vapnik, V. N., *The Nature of Statistical Learning Theory*, Springer, New York (1995).