The SFA-LSSVM as a Decision Support System for Mitigating Liquefaction Disasters

Julian Pratama Putra Thedja[1], Jui-Sheng Chou[1], and Tri Joko Wahyu Adi[2]

Abstract-Advanced data mining techniques are potential tools for solving civil engineering problems. This study proposes a novel classification system that integrates smart firefly algorithm (SFA) with least squares support vector machine (LSSVM). SFA is an optimization algorithm which combines firefly algorithm (FA) with smart components, namely chaotic logistic map, chaotic gauss/mouse map, adaptive inertia weight and Lévy flight to enhance optimization solutions. The least squares support vector machine (LSSVM) was adopted in this study for its superior performance of solving real-world problems. Based on the provided engineering data, the analytical results confirm that the SFA-LSSVM has 95.18% prediction accuracy.

Index Terms - Data mining, optimization, firefly algorithm, support vector machines, liquefaction.

INTRODUCTION

Existence of soil liquefaction when earthquake happens is one of the critical issues in geotechnical engineering. Liquefaction can be defined as the transformation of a granular material from a solid to a liquefied state because of increased pore-water pressure and reduced effective stress. For example, in water saturated sand, the sand grain packed together. However, between each of sand grain, there is a body of water known as pore water. As the sand vibrates, it shifts. The water under pressure then pushes the sand grains apart. Therefore, sand grains are no longer wrestling together and no longer stable. This phenomenon is usually caused by earthquake and greatly reducing soil effective stress that leads to losses bearing capacity of a foundation.

A least squares support vector machine (LSSVM) is an AI algorithm based on Statistical Learning Theory. The LSSVM is now recognized as an excellent AI algorithm and has been widely used due to its advantages in many fields. However, the performance of LSSVM depends on the selection of penalty parameter (C) and kernel parameter (σ). Both of LSSVM parameters known as LSSVM hyperparameters. Optimization of LSSVM hyperparameters avoids over-fitting, avoids local minima problems, and improves prediction accuracy. Some researcher has proven modified firefly algorithm combined with LSSVM is better than other hybrid algorithms. A chaotic firefly algorithm for optimizing the LSSVM hyper-parameters performs better than other algorithms [1]. Thus, the chaotic firefly algorithm is further improved by combining it with new smart components, namely adaptive inertia weight and Lévy flights in this study.

[1]Julian Pratama Putra Thedja and Jui-Sheng Chou are with National Taiwan University of Science and Technology, Taiwan. E-mail: Julian.thedja@hotmail.com; jschou@mail.ntust.edu.tw.
[2]Tri Joko Wahyu Adi is with Departement of Civil Engineering, Faculty of Civil Engineering and Planning, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia. E-mail: trijoko_w@yahoo.com.

METHOD

A. *Least Squares Support Vector Machine*

The support vector machine (SVM) was originally developed by Vapnik et al. in 1995 [2]. The SVM has been widely used for classification because of its high learning capabilities. An SVM performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. The main idea of SVM is to find the largest margin between two categories. The least squares version of support vector machines (LSSVM) classifiers is close to conventional SVM formulation. Alternatively, it solves linear problems, not quadratic programming problems [3]. This algorithm applies a least squares cost function to obtain a linear set of equations in the dual space by modifying the conventional SVM as shown in Eqs. (1) and (2):

$$L = \frac{1}{2}\| \vec{w} \|^2 - \frac{1}{2} C \sum_{i=1}^{N} e_i^2 \quad (1)$$

Subject to the equality constraint

$$y_i(\overrightarrow{w_i} \bullet \overrightarrow{x_i} + b) = 1 - e_i \quad (2)$$

The LSSVM method is attractive because it has a low computational cost compared to the conventional SVM and is as accurate as the conventional SVM. The LSSVM with RBF kernel already proved its performance by solving a two-spiral classification problem, which is known to be hard for multilayer perceptron [4]. The LSSVM also uses all samples to find a good approximation model. Therefore, LSSVM is widely used to solve real-world problems.

B. *Swarm and Evolutionary Optimization Algorithm*

The firefly algorithm (FA) developed by Yang is based on the flashing patterns and behavior of tropical fireflies [5]. Equation (3) gives the movement of the j^{th} firefly when attracted to another more attractive (brighter) k^{th} firefly at x_j and x_k, respectively.

$$x_j = x_j + \beta_0 e^{-\gamma r_{jk}^2}(x_k - x_j) + \alpha_0(rand - 0.5) \quad (3)$$

The FA parameters are fixed and do not change during iterations. However, an important component in swarm intelligence and modern meta-heuristics is the use of randomization to enable an algorithm to jump out of any local optimum during a global search. Fine-tuning the randomness and balance of local search and global search are essential for controlling the performance of any meta-heuristic algorithm. Thus, FA must be incorporated with other components to enhance FA performance. In this study, chaotic Gauss/Mouse map is used to fine tune β_0 parameter, chaotic Logistic map is used to diversify the FA initialization, adaptive inertia weight is used to maintain α_0 in a reasonable range, and Lévy flight is used to increase optimization capability of FA by mimicking the movements of insects.

CASE STUDY AND DISCUSSION

The historical data set was recapped by Goh and Goh [6]. The 226 cases in the soil liquefaction database include 133 liquefied cases (class 1) and 93 non-liquefied cases (class 0). The data represents the field performance of 52 sites taken from six different earthquakes. The six input variables considered were the cone tip resistance (q_c), the sleeve friction ratio (R_f), the effective stress at the depth of interest (σ'_v), the total stress at the same depth (σ_v), the maximum horizontal ground surface acceleration (a_{max}), and the earthquake moment magnitude (M_w).

Table 1 shows that the proposed model can predict soil liquefaction existence with 94.31% accuracy in average. Using feature scaling increases accuracy to 95.18%. Notably, the TACO-miner algorithm [7] is highly effective for predicting soil liquefaction existence. It predicts soil liquefaction existence with 100% accuracy. Unfortunately, k-fold cross-validation algorithm was not performed to minimize prediction bias in their studies. The accuracy presented in literature may be a one-time luck. Although the proposed algorithm is not as accurate as previous algorithms, its results are relatively reliable based on the 10-fold cross validation.

TABLE 1. COMPARISON RESULTS.

Literature	Technique	Cross fold validation	Accuracy (%)
Goh and Goh, 2007	SVM	-	98.00%
Baykasoglu, 2009	NBTree	-	86.67%
	Decision table	-	93.33%
	PART	-	84.00%
	C4.5	-	90.67%
	MEPAR-Miner	-	97.73%
	TACO-miner	-	100.00%
This study	SFA-LSSVM (original value)	10	94.31%
	SFA-LSSVM (Feature scaling)	10	95.18%

The performance of the proposed SFA-LSSVM system was validated with the actual case to confirm the practicality of a hybrid swarm intelligence system. The SFA-LSSVM has consistent and adequate prediction accuracy compared to previous prediction methods and can be considered as an effective and accurate decision-support system.

REFERENCES

[1] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi, and O. K. Hussain, "Support vector regression with chaos-based firefly algorithm for stock market price forecasting," Applied Soft Computing, vol. 13, pp. 947-958, 2// 2013.

[2] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.

[3] J. A. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, Least squares support vector machines vol. 4: World Scientific, 2002.

[4] J. A. K. Suykens, "Nonlinear modelling and support vector machines," in Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE, 2001, pp. 287-294 vol.1.

[5] X.-S. Yang, "Firefly algorithms for multimodal optimization," in Stochastic algorithms: foundations and applications, ed: Springer, 2009, pp. 169-178.

[6] A. T. C. Goh and S. H. Goh, "Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data," Computers and Geotechnics, vol. 34, pp. 410-421, 9// 2007.

[7] A. Baykasoğlu, A. Çevik, L. Özbakır, and S. Kulluk, "Generating prediction rules for liquefaction through data mining," Expert Systems with Applications, vol. 36, pp. 12491-12499, 12// 2009.