



PROJECT DISPUTE PREDICTION BY HYBRID MACHINE LEARNING TECHNIQUES

Jui-Sheng CHOU^a, Chih-Fong TSAI^b, Yu-Hsin LU^c

^a*Department of Construction Engineering, National Taiwan University of Science and Technology, 43, Sec. 4, Keelung Rd, Taipei, 106, Taiwan (R.O.C.)*

^b*Department of Information Management, National Central University, No. 300, Zhongda Rd, Zhongli City, Taoyuan County, 32001, Taiwan*

^c*Department of Accounting, Feng Chia University, 100, Wenhwa Rd. Seatwen, Taichung 40724, Taiwan*

Received 27 Jul. 2011; accepted 20 Jan. 2012

Abstract. This study compares several well-known machine learning techniques for public-private partnership (PPP) project dispute problems. Single and hybrid classification techniques are applied to construct models for PPP project dispute prediction. The single classification techniques utilized are multilayer perceptron (MLP) neural networks, decision trees (DTs), support vector machines, the naïve Bayes classifier, and *k*-nearest neighbor. Two types of hybrid learning models are developed. One combines clustering and classification techniques and the other combines multiple classification techniques. Experimental results indicate that hybrid models outperform single models in prediction accuracy, Type I and II errors, and the receiver operating characteristic curve. Additionally, the hybrid model combining multiple classification techniques perform better than that combining clustering and classification techniques. Particularly, the MLP + MLP and DT + DT models perform best and second best, achieving prediction accuracies of 97.08% and 95.77%, respectively. This study demonstrates the efficiency and effectiveness of hybrid machine learning techniques for early prediction of dispute occurrence using conceptual project information as model input. The models provide a proactive warning and decision-support information needed to select the appropriate resolution strategy before a dispute occurs.

Keywords: machine learning; clustering and classification; hybrid intelligence; public-private partnership; project management; dispute prediction.

Reference to this paper should be made as follows: Chou, J.-S.; Tsai, C.-F.; Lu, Y.-H. 2013. Project dispute prediction by hybrid machine learning techniques, *Journal of Civil Engineering and Management* 19(4): 505–517. <http://dx.doi.org/10.3846/13923730.2013.768544>

Introduction

During the last decade, many PPP projects were not as successful as expected due to project disputes occurring during the build, operate, and transfer (BOT) phase. According to the Taiwan Public Construction Commission (TPCC), the dispute rate was 23.6% during 2002–2009 (PCC 2011). These disputes were resolved by mediation and non-mediation procedures. Non-mediation procedures include arbitration, litigation, negotiation, and administrative appeals. In Taiwan, up to 84% of PPP projects disputes are settled by mediation or negotiation within 1–9 months (PCC 2011). Notably, arbitration or litigation costs to all parties are considerably more in time and money than those associated with mediation or negotiation.

Most research has focused on predicting litigation outcomes (Arditi, Tokdemir 1999a, b; Arditi, Pulket 2005, 2010; Arditi *et al.* 1998; Chau 2007;

Pulket, Arditi 2009a, b) rather than providing a proactive dispute warning. Additionally, most studies examined the relationship between the project owner and general contractor; however, PPP projects involve many stakeholders, including the government, participating private investors, and financial institutions. This study intends to provide early dispute warnings by predicting when disputes will occur based on preliminary project information.

For effective control of PPP projects and to design proactive dispute management strategies, early knowledge of PPP project dispute propensity is essential to provide the governmental PPP taskforce with the information needed to implement a win-win resolution strategy and even prevent disputes. Further, depending on possible dispute outcomes, precautionary measures can be implemented proactively during

project execution. Additional preparation in preventive actions can prove beneficial once a dispute occurs by reducing future effort, time, and cost to multiple parties during dispute settlement processes.

To achieve this goal, this study compares different prediction models using a series of machine learning techniques for predicting PPP dispute likelihood and thereby eliminates future adverse impacts of disputes on project delivery, operation, and transfer. Particularly, this study uses single and hybrid machine learning techniques. The single machine learning models are based on neural networks, decision trees (DTs), support vector machines (SVMs), the naïve Bayes classifier, and k -nearest neighbor (k -NN). Two hybrid learning models are developed, one combining clustering and classification techniques and the other combining multiple classification techniques.

The rest of this paper is organized as follows. Section 1 thoroughly reviews artificial intelligence (AI) literature and its accuracy in predicting conventional construction disputes and litigation outcomes. Section 2 then introduces the single and hybrid machine-learning schemes. Next, Section 3 discusses the experimental setup and results from comparing the single and hybrid machine learning techniques for dispute outcome prediction. Conclusions are finally drawn in the final section, along with recommendations for future research.

1. Literature review

Management personnel typically benefit when the taskforce has a decision-support tool for estimating dispute propensity and for early planning of how disputes should be resolved before project initiation (Marzouk *et al.* 2011). Several studies have attempted to minimize construction litigation by predicting the outcomes of court decisions. In Arditi *et al.* (1998), a network was trained using data from Illinois appellate courts, and 67% prediction accuracy was obtained. Arditi *et al.* (1998) argued that if the parties in a dispute know with some certainty how a case will be resolved in court, the number of disputes can be reduced markedly.

In another series of studies, AI techniques achieved superior prediction accuracy with the same dataset – 83.33% in a case-based reasoning study (Arditi, Tokdemir 1999b), 89.95% with boosted DTs (Arditi, Pulket 2005), and 91.15% by integrated prediction modeling (Arditi, Pulket 2010). These studies used AI to enhance prediction of outcomes in conventional construction procurement litigation.

However, Chau (2007) determined that, other than in the above case studies, AI techniques are rarely applied in the legal field. Thus, Chau (2007) applied AI techniques based on particle swarm optimization to predict construction litigation outcomes, a field in which new data mining techniques are rarely applied. The network achieved an 80% prediction accuracy

rate, much higher than mere chance. Nevertheless, Chau (2007) suggested that additional case factors, such as cultural, psychological, social, environmental, and political factors, be used in future studies to improve accuracy and reflect real world.

For construction disputes triggered by change orders, Chen (2008) applied a k -NN pattern classification scheme to identify potential lawsuits based on a nationwide study of US court records. Chen (2008) demonstrated that the k -NN approach achieved a classification accuracy of 84.38%. Chen and Hsu (2007) further applied a hybrid artificial neural networks case-based reasoning (ANN-CBR) model with dispute change order dataset to obtain early warning information of construction claims. The classifier attained a prediction rate of 84.61% (Chen, Hsu 2007).

Despite the numerous studies of CBR and its variations for identifying similar dispute cases for use as references in dispute settlements, Cheng *et al.* (2009) refined and improved the conventional CBR approach by combining fuzzy set theory with a novel similarity measurement that combines Euclidean distance and cosine angle distance. Their model successfully extracted the knowledge and experience of experts from 153 historical construction dispute cases collected manually from multiple sources.

Generally, all previous studies focused on either specific change order disputes or on conventional contracting projects using a single accuracy performance measure. Characteristics and environments of construction projects under the PPP strategy, however, differ markedly from the general contractor and owner relationships and require machine learning techniques with rigorous model performance measures to assist governmental agencies in predicting disputes with excellent accuracy.

Since disputes always involve numerous complex and interconnected factors and are difficult to rationalize, machine learning techniques is now among the most effective methods for identifying hidden relationships between available or accessible attributes and dispute-handling methods (Arditi, Pulket 2005, 2010; Arditi, Tokdemir 1999a; El-Adaway, Kandil 2010; Kassab *et al.* 2010; Pulket, Arditi 2009b). Approaches based on machine learning are related to computer system designs that attempt to resolve problems intelligently by emulating human brain processes (Lee *et al.* 2008) and are typically used to solve prediction or classification problems.

Researchers in various scientific and engineering fields have recently combined different learning techniques to increase their efficacy. Numerous studies have demonstrated that hybrid schemes are promising applications in various industries (Arditi, Pulket 2010; Chen 2007; Chou *et al.* 2010, 2011; Kim, Shin 2007; Lee 2009; Li *et al.* 2005; Min *et al.* 2006; Nandi *et al.* 2004; Wu 2010; Wu *et al.* 2009). However, selecting

the most appropriate combinations is difficult and time consuming, such that further attempts are not worthwhile unless significant improvements in accuracy are achieved. This study constructs PPP project dispute-prediction models using single and hybrid machine learning techniques.

2. Machine learning techniques

2.1. Classification techniques

2.1.1. Artificial neural networks

ANN consists of information-processing units that resemble neurons in the human brain, except that a neural network consists of artificial neurons (Haykin 1999). Generally, a neural network is a group of neural and weighted nodes, each representing a brain neuron; connections among these nodes are analogous to synapses between brain neurons (Malinowski, Ziembicki 2006).

Multilayer perceptron (MLP) neural networks are standard neural network models. In an MLP network, the input layer contains a set of sensory input nodes, one or more hidden layers contain computation nodes, and an output layer contains computation nodes.

In a multilayer architecture, input vector x passes through the hidden layer of neurons in the network to the output layer. The weight connecting input element i to hidden neuron j is W_{ji} , and the weight connecting hidden neuron j to output neuron k is V_{kj} . The net input of a neuron is derived by calculating the weighted sum of its inputs, and its output is determined by applying a sigmoid function. Therefore, for the j^{th} hidden neuron:

$$net_j^h = \sum_{i=1}^N W_{ji}x_i \text{ and } y_i = f(net_j^h), \quad (1)$$

and for the k^{th} output neuron:

$$net_k^o = \sum_{j=1}^{J+1} V_{kj}y_j \text{ and } o_k = f(net_k^o). \quad (2)$$

The sigmoid function $f(net)$ is the logistic function:

$$f(net) = \frac{1}{1 + e^{-\lambda net}}, \quad (3)$$

where λ controls the function gradient.

For a given input vector, the network produces an output o_k . Each response is then compared to the known desired response of each neuron d_k . Weights in the network are modified continuously to correct or reduce errors until total error from all training examples stays below a pre-defined threshold.

For the output layer weights V and hidden layer weights W , update rules are given by Eqs (4) and (5), respectively:

$$V_{kj}(t+1) = v_{kj}(t) + c\lambda(d_k - o_k)o_k(1 - o_k)y_j(t); \quad (4)$$

$$W_{ji}(t+1) = w_{ji}(t) + c\lambda^2 y_j(1 - y_j)x_i(t) \times \left(\sum_{k=1}^K (d_k - o_k)o_k(1 - o_k)v_{kj} \right). \quad (5)$$

2.1.2. Decision trees

DTs have a top-down tree structure, which splits data to create leaves. In this study, the C4.5 classifier, a recent version of the ID3 algorithm (Quinlan 1993), is used to construct a DT for classification. A DT is constructed in which each internal node denotes a test of an attribute and each branch represents a test outcome. Leaf nodes represent classes or class distributions. The top-most node in a tree is the root node with the highest information gain. After the root node, the remaining attribute with the highest information gain is then chosen as the test for the next node. This process continues until all attributes are compared or no remaining attributes exist on which samples may be further partitioned (Huang, Hsueh 2010; Tsai, Chen 2010).

Assume one case is selected randomly from a set S of cases and belongs to class C_j . The probability that an arbitrary sample belongs to class C_j is estimated by:

$$P_i = \frac{\text{freq}(C_j, S)}{|S|}, \quad (6)$$

where $|S|$ is the number of samples in set S and, thus, the information it conveys is $-\log_2 p_i$ bits.

Suppose a probability distribution $P = \{p_1, p_2, \dots, p_n\}$ is given. The information conveyed by this distribution, also called entropy of P , is then:

$$\text{Info}(P) = \sum_{i=1}^n -p_i \log_2 p_i. \quad (7)$$

If a set T of samples is partitioned based on the value of a non-categorical attribute X into sets T_1, T_2, \dots, T_m , then the information needed to identify the class of an element of T becomes the weighted average of information needed to identify the class of an element of T_i , that is, the weighted average of $\text{Info}(T_i)$:

$$\text{Info}(X, T) = \sum_{i=1}^m \frac{|T_i|}{T} \times \text{Info}(T_i). \quad (8)$$

Information gain, $\text{Gain}(X, T)$, is then derived as:

$$\text{Gain}(X, T) = \text{Info}(T) - \text{Info}(X, T). \quad (9)$$

This equation represents the difference between information needed to identify an element of T and

information needed to identify an element of T after the value of attribute X has been determined. Thus, it is the gain in information due to attribute X .

2.1.3. Support vector machines

SVMs, which were introduced by Vapnik (1998), perform binary classification, that is, they separate a set of training vectors for two different classes $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, where $x_i \in R^d$ denotes vectors in a d -dimensional feature space and $y_i \in \{-1, +1\}$ is a class label. The SVM model is generated by mapping input vectors onto a new higher dimensional feature space denoted as $\Phi : R^d \rightarrow H^f$, where $d < f$. In classification problems, SVM identifies a separate hyperplane that maximizes the margin between two classes. Maximizing the margin is a quadratic programming problem, which can be solved from its dual problem by introducing Lagrangian multipliers (Han, Kamber 2001; Tan et al. 2006; Witten, Frank 2005). An optimal separating hyperplane in the new feature space is then constructed by a kernel function $K(x_i, x_j)$, which is the product of input vectors x_i and x_j and where $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$.

2.1.4. Naïve Bayes classifier

The naïve Bayes classifier requires all assumptions be explicitly built into models that are then utilized to derive ‘optimal’ decision/classification rules. This classifier can be used to represent the dependence between random variables (features) and to generate a concise and tractable specification of a joint probability distribution for a domain (Witten, Frank 2005). The classifier is constructed using training data to estimate the probability of each class, given feature vectors of a new instance. For an example represented by feature vector X , the Bayes theorem provides a method for computing the probability that X belongs to class C_i , which is denoted as $p(C_i|X)$:

$$P(C_i|X) = \prod_{j=1}^N P(x_j|C_i). \quad (10)$$

That is, the naïve Bayes classifier determines the conditional probability of each attribute $x_j (j=1, 2, \dots, N)$ of X given class label C_i . Therefore, the (image) classification problem can be stated as follows: given a set of observed features x_j from an image X , classify X into one class C_i .

2.1.5. k -Nearest neighbor

In pattern classification, the k -NN classifier is a conventional non-parametric classifier (Bishop 1995). To classify an unknown instance represented by some feature vectors as a point in a feature space, the k -NN classifier calculates distances between the point and points in a training dataset. It then assigns the point to the class among its k -NNs (where k is an integer).

The k -NN classifier differs from the *inductive* learning approach described previously; thus, it has also been called instance-based learning (Mitchell 1997) or a *lazy learner*. That is, without off-line training (i.e. model generation) the k -NN algorithm only needs to *search* all examples of a given training dataset to classify a new instance. Therefore, the primary computation of the k -NN algorithm is on-line scoring of training examples to find the k -NNs of a new instance. According to Jain et al. (2000), 1-NN can be conveniently used as a benchmark for all the other classifiers since it achieves reasonable classification performance in most applications.

2.2. Hybrid classification techniques

In literature, hybridization improves the performance of single classifiers. Hybrid systems can address relatively more complex tasks because they combine different techniques (Lenard et al. 1998). Generally, hybrid models are based on combining two or more machine learning techniques (e.g. clustering and classification techniques).

According to Tsai and Chen (2010), two methods can be applied to construct hybrid models for classification – the sequential combination of clustering and classification techniques and the sequential combination of different classification techniques. These two methods are described as follows.

2.2.1. Clustering + Classification techniques

The method combining clustering and classification techniques uses one clustering algorithm as the first component of the hybrid system. This study uses the k -means clustering algorithm to combine classification techniques.

The k -means clustering algorithm, a simple and efficient clustering algorithm, iteratively updates the means of data items in a cluster; the stabilized value is then regarded as representative of that cluster. The basic algorithm has the following steps (Hartigan, Wong 1979):

- Randomly select k data items as cluster centers;
- Assign each data item to the group that has the closest centroid;
- When all data items have been assigned, recalculate the positions of k centroids;
- If no further change exists, end the clustering task; otherwise, return to step 2 NOTE: if you need to use this phrase, you have to change the bullets into step 1, step 2, etc.

Therefore, clustering can be used as a pre-processing stage to identify pattern classes for subsequent supervised classification. Restated, the clustering result can

be used for pre-classification of unlabelled collections and to identify major populations in a given dataset.

Alternatively, clustering can be used to filter out unrepresentative data. That is, the data that cannot be clustered accurately can be considered noisy data. Consequently, representative data, which are not filtered out by the clustering technique, are used during the classification stage.

Next, the classification stage is the same as that for training or constructing a classifier. The clustering result becomes the training dataset to train a classifier. After the classifier is trained, it can classify new (unknown) instances.

Given a training dataset D , which contains m training examples, the aim of clustering is to “pre-process” D for data reduction. That is, the correctly clustered data D' by the cluster are collected, where D' contains n examples ($n < m$ and $D' \in D$). Then, D' is used to train the classifier. Hence, given a test dataset, the classifier provides better classification results than single classifiers trained with the original dataset D .

2.2.2. Classification + Classification techniques

Another hybrid approach combines multiple classification techniques sequentially; that is, multiple classifiers are cascaded. As with the combination of clustering and classification techniques, the first classifier can be used to reduce the amount of data.

The way of cascading two classification techniques is as follows: given a training dataset D , which contains m training examples, it is used to train and test the first classifier. Notably, 100% classification accuracy is impossible. Therefore, the correctly classified data D' by the first classifier are collected, where D' contains o examples, where $o < m$ and $D' \in D$. Then, D' is utilized to train the second classifier. Again, the hybrid classifier could provide better classification results than single classifiers trained with the original dataset D over a given test dataset.

3. Modeling experiments

3.1. Experimental setup and design

3.1.1. The dataset

To demonstrate the accuracy and efficiency of the dispute classification schemes, this study used PPP project data collected by the TPCC, the authority overseeing infrastructure construction in Taiwan, to construct classification models to predict dispute likelihood. The study database contains 584 PPP projects overseen by the TPCC during 2002–2009. Of 584 surveys issued, 569 were returned completed, for a response rate of 97.4%. The questionnaire included items to collect social demographic data of respondents, background information, project characteristics, and project dispute resolutions.

Several projects had more than one dispute – one project had nine disputes – at various project stages. Thus, the overall dataset comprised data for $N = 645$ cases (i.e. $N_2 = 493$ cases without disputes and $N_1 = 152$ dispute cases). Through expert feedback, project attributes and their derivatives that were clearly relevant to the prediction output of interest were identified by survey items. However, quantitative techniques were still needed to construct and validate hidden relationships between selected project predictors and the response (output) variable.

Table 1 summarizes the statistical profile of categorical labels and numerical ranges for study samples. For PPP-oriented procurement, 59.5% of projects were overseen by the central government. Over the last eight years, most public construction projects have been for cultural and education facilities (25.3%), sanitation and medical facilities (20.8%), transportation facilities (18.1%), and major tourist site facilities (10.5%). In accordance with economic planning and development policy, 48.5% of projects were located in northern Taiwan. Based on the standard industry definition, most private sector investment was in industrial (38.6%) and service departments (50.7%). In most cases (91.0%), the government provided land and planned the facility to attract investors.

The three major PPP strategies for delivering public services are BOT (23.7%); operate and transfer (OT) (52.7%); and rehabilitate, operate, and transfer (ROT) (23.6%). Specifically, the World Bank Group (WBG 2011) defines the BOT scheme as a strategy in which a private sponsor builds a new facility, operates the facility, and then transfers the facility to the government at the end of the contract period. The government typically provides revenue guarantees through long-term take-or-pay contracts. When a private sponsor renovates an existing facility, and then operates and maintains the facility at its own risk for the contract period, the PPP strategy is ROT, according to WBG (2011) classifications. Projects involving only management and lease contracts are classified as OT projects.

Further, flagship infrastructure projects refer to those that are important and generally large. Average project value was approximately New Taiwan Dollar (NTD) 841 million (i.e. 1 USD is approximately equal to 30 NTD). Based on collected data, the overall procurement amount via PPP was roughly NTD 543 billion. Mean capital investment by the government and private sector per project was NTD 63.5 million and NTD 777.8 million, respectively. Notably, the average private capital investment ratio was as high as 91.4%. The mean duration of licensed facility operations was about 12 years (maximum, 60 years).

To assess the dependencies between categorized data, contingency table analyses were compared between particular predictors and the response variable via chi-square testing to infer relationships (Table 2).

Table 1. Project attributes and their descriptive statistics

Attribute	Data range, categorical label or statistical description
Input variables	
Type of government agency in charge	Central authority (59.5%); Municipality (11.5%); Local government (29%)
Type of public construction and facility	1: Transportation facilities (18.1%); 2: Common conduit (0%); 3: Environmental pollution prevention facilities (2.3%); 4: Sewerage (1.1%); 5: Water supply facilities (0.5%); 6: Water conservancy facilities (2.5%); 7: Sanitation and medical facilities (20.8%); 8: Social welfare facilities (3.9%); 9: Labor welfare facilities (1.2%); 10: Cultural and education facilities (25.3%); 11: Major tour-site facilities (10.5%); 12: Power facilities (0%); 13: Public gas and fuel supply facilities (0%); 14: Sports facilities (3.3%); 15: Parks facilities (2.5%); 16: Major industrial facilities (0.5%); 17: Major commercial facilities (1.9%); 18: Major hi-tech facilities (0.2%); 19: New urban development (0%); 20: Agricultural facilities (5.6%)
Project location	North (48.5%); Center (21.2%); South (24.5%); East (5.3%); Isolated island (0.5%)
Executive authority	Central authority (36.0%); Municipality (36.1%); Local government (27.9%)
Type of invested private sector	Standard industry classification-Primary (0.2%); Secondary (38.6%); Tertiary (50.7%); Quaternary (10.5%)
Planning and design unit	Government provides land and plans facility (91.0%); Government provides land and private investor designs facility (5.9%); Private provides land and designs facility (3.1%)
PPP contracting strategy	BOT (23.7%); OT (52.7%); ROT (23.6%)
Major public infrastructure/facility	Promoted as major public infrastructure/facility in PPP Act (80.1%); Not major infrastructure/facility (19.9%)
Project scale	Range: 0–60,000,000; Sum: 5.43E8; Mean: 841337.1776; Standard deviation: 3.52061E6 (Thousand NTD; USD:NTD is about 1:30 as of Apr. 2011)
Government capital investment	Range: 0–9,600,000; Sum: 40,975,392.41; Mean: 63527.7402; Standard deviation: 5.11192E5 (Thousand NTD)
Private capital investment amount	Range: 0–60,000,000; Sum: 5.02E8; Mean: 777809.4374; Standard deviation: 3.32433E6 (Thousand NTD)
Private capital investment ratio (PCIR)	Range: 0–100; Mean: 91.4729; Standard deviation: 25.42269 (%)
Licensed operations duration	Range: 0–60; Mean: 11.9778; Standard deviation: 13.39007 (Year)
Output variable	
Dispute propensity	No dispute occurred (76.4%); Dispute occurred (23.6%)

All tests obtained statistically significant results at the 5% alpha level except variables (i.e. planning and design; PCIR) that were rejected by the null hypothesis, that is, no relationship was observed between the row variable (input variables) and column variable (output variable). For instance, among the dispute cases ($N_1 = 152$), the central government had a higher probability of encountering disputes (67.1% probability) than municipal (15.1%) and local governments (17.8%).

Particularly, in Nos. 1, 6, 7, 10, 11, 20 in type of public construction and facility of Table 1, disputes

occurred in 76.4% of projects. Data show that 85.5% of disputes occurred in northern and southern Taiwan. Interestingly, 92.1% of disputes occurred when the government provided land and planned the facility, while only 2% occurred when private investors provided land and designed the facility. Among the three PPP strategies, the probability of disputes was higher with BOT (49.3%) than with OT (32.2%) and ROT (18.4%). Notably, once a project was legally promoted as a major infrastructure project, the likelihood of a PPP dispute was 38.8%, lower than that for non-major infrastructure projects (61.2%).

Table 2. Contingency table and chi-square test results for dispute cases

Project attributes	<i>p</i> -value	Dispute occurred (%)
<i>Agency</i>	0.002	
Central authority		67.1
Municipality		15.1
Local government		17.8
<i>Type of public construction</i>	0.000	
Transportation facilities		10.5
Water conservancy facilities		9.9
Sanitation and medical facilities		17.1
Cultural and education facilities		13.2
Major tour-site facilities		14.5
Agricultural facilities		11.2
<i>Planning and design</i>	0.657	
Government provides land and plans facility		92.1
Government provides land and private investor designs facility		5.9
Private investor provides land and designs facility		2.0
<i>PPP strategy</i>	0.000	
BOT		49.3
OT		32.2
ROT		18.4
<i>Major public infrastructure</i>	0.000	
No		61.2
Yes		38.8
<i>Project scale (Thousand NTD)</i>	0.000	
< 5,000		15.8
5000–50,000		15.8
> 50,000		68.4
<i>PCIR (%)</i>	0.057	
< 25		3.3
25–50		0.0
50–75		3.9
> 75		92.8
<i>LOD (Year)</i>	0.000	
< 5		19.7
5–10		23.0
10–15		5.9
15–20		13.8
> 20		37.5

Moreover, once project value exceeded NTD 50 million, dispute propensity was 4.33 times higher than that for projects valued at NTD 5–50 million and less than NTD 5 million. However, when private sector investment exceeded 75%, dispute likelihood increased to 92.8%. Notably, dispute patterns were not significantly related to licensed operating period. Table 2 summarizes statistical results of cross-analysis.

3.1.2. Single baseline model construction

The single baseline models using classification techniques are based on C4.5 DTs, the naïve Bayes classifier, SVMs, neural network classifier, and *k*-NN classifier.

Parameter settings for constructing the five baseline prediction models are described as follows:

- DTs. The C4.5 DT is established and the confidence factor for pruning the tree is set at 0.25. Parameters for the minimum number of instances per leaf and amount of data used to reduce pruning errors are 2 and 3, respectively;
- ANN. This study uses the MLP classifier. To avoid overtraining, this study constructs an MLP classifier by examining different parameter settings to obtain an average accuracy for further comparisons. Therefore, this study considers five different numbers of hidden nodes and learning epochs. The numbers of hidden nodes are 8, 12, 16, 24, and 32 and those of learning epochs are 50, 100, 200, 300, and 500;
- Naïve Bayesian classifier. In building the naïve Bayes classifier, this study uses supervised discretization to convert numerical attributes into nominal attributes, which can increase model accuracy. Additionally, the kernel estimator option is set as false because some attributes are nominal;
- SVM. The complexity parameter, *C*, and tolerance parameter are as 1.0 and 0.001, respectively. For the kernel function, the radial basis function with a gamma value of 1 is used;
- *k*-NN classifier. Different *k* values are assessed in this study, starting at 1 and increasing until the minimum error rate is reached.

When comparing the predictive performance of two or more methods, researchers often use *k*-fold cross-validation to minimize bias associated with random sampling of training and holdout data samples. As cross-validation requires random assignment of individual cases into distinct folds, a common practice is to stratify the folds. In stratified *k*-fold cross-validation, the proportions of predictor labels (responses) in folds should approximate those in the original dataset.

Empirical studies show that, compared to traditional *k*-fold cross-validation, stratified cross-validation reduces bias in comparison results (Han, Kamber 2001). Kohavi (1995) further demonstrated that 10-fold validation testing was optimal when computing time and variance. Thus, this study uses stratified 10-fold cross-validation to assess model performance. The entire dataset was divided into 10 mutually exclusive subsets (or folds), with class distributions approximating those of the original dataset (stratified). The subsets were extracted using the following five steps:

1. Randomize the dataset;
2. Extract one tenth of the original dataset from the randomized dataset (single fold);
3. Remove extracted data from the original dataset;

4. Repeat steps (1)–(3) eight times;
5. Assign the remaining portion of the dataset to the last fold (10th fold).

After applying this procedure to obtain 10 distinct folds, each fold was then used once for performance tests of the single flat and hybrid classification models, and the remaining nine folds were used for training model, which obtained 10 independent performance estimates. The cross-validation estimate of overall accuracy was calculated by averaging the k individual accuracy measures for cross-validation accuracy.

3.1.3. Hybrid model construction

For the hybrid models combining clustering and classification techniques, the k -means clustering algorithm is applied first as the clustering stage. Notably, the k value was set to 3, 4, 5, and 6. As dispute and no-dispute groups exist, there are two clusters out of k corresponding to these two groups, which provide higher accuracy rates than the other clusters. Then, they are selected as the clustering result.

For the example of k -means ($k = 4$), four clusters are produced and represented by C_1 , C_2 , C_3 , and C_4 based on a training dataset. According to the ground truth answer in the training dataset, one can identify two of the four clusters, which can be well ‘classified’ into the dispute and no-dispute groups. The other two clusters whose data are not well classified or difficult to classify by k -means clustering are filtered out.

Once the best k -mean is found, its clustered data (i.e. the clustering result) are used to train the five single classifiers. Notably, one specific clustering model for the 10 training datasets (by 10-fold cross validation) will yield 10 different clustering results. That is, data in the two representative clusters, which can best recognize the dispute and no-dispute groups using the 10 training datasets, are not duplicated. Therefore, the final clustering result of each k -means model is based on the union method for selecting dispute and no-dispute data. The clustering result is then used as the new training dataset to train the five baseline models.

Conversely, for the cascaded hybrid classifiers, the best baseline classification model is identified after performing 10-fold cross validation, that is, one of the C4.5 DTs, naïve Bayes classifier, SVMs classifier, k -NN classifier, and neural network classifier. The correctly predicted data from the training set by the best baseline model are used as new training data to train the five single baseline models.

3.1.4. Evaluation methods

To assess the performance of these single and hybrid prediction models, prediction accuracy and Type I and II errors, that is, false-positive and false-negative

errors, are examined. Table 3 shows a confusion matrix for calculating accuracy and error rates, which are commonly used measures for binary classification (Ferri et al. 2009; Horng 2010; Kim 2010; Sokolova, Lapalme 2009).

Prediction accuracy, which is defined as the percentage of records predicted correctly by a model relative to the total number of records among classification models, is a primary evaluation criterion. The classification *accuracy* is derived by:

$$\text{Accuracy} = \left(\frac{a + d}{a + b + c + d} \right). \quad (11)$$

Conversely, the Type I error is the error of not rejecting a null hypothesis when an alternative hypothesis is the true state. In this study, Type I error means that the event occurred when the model classified the event group into the non-event group. The Type II error is defined as the error in rejecting a null hypothesis when it is the true state, meaning the event occurred when the model classified the non-event group into the event group.

Moreover, the Receiver Operating Characteristic (ROC) curves reflect the ability of a classifier to avoid false classification. The ROC curve captures a single point, the area under the curve (AUC), in the analysis of model performance. As the distance between the curve and reference line increases, test accuracy increases. The AUC, sometimes referred to as balanced accuracy (Sokolova, Lapalme 2009), is derived easily by Eq. (12):

$$\text{AUC} = \frac{1}{2} \left[\left(\frac{a}{a + b} \right) + \left(\frac{d}{c + d} \right) \right]. \quad (12)$$

3.2. Experimental results

Table 4 lists the prediction performance of the five single classifiers, including their prediction accuracy, Type I and II errors, and the ROC curve. Experimental results indicate that the DT classifier performs best, providing the highest prediction accuracy at 83.72% and the lowest Type II error rate at 5.07%. The MLP classifier performs second best in prediction accuracy at 82.33%. Notably, the significant difference level is higher than 95% or 99% by t -test for all the performance measures of the individual models. Therefore, of the hybrid models combining multiple

Table 3. Confusion matrix

		Predicted	
		Positive	Negative
Actual	Positive	a (tp)	b (fn)
	Negative	c (fp)	d (tn)

Table 4. Prediction accuracy of single classifiers

Model	Accuracy	Type I error	Type II error	ROC Curve	Ranking by accuracy
MLP	82.33	44.08	9.53	0.781	2
DT	83.72	52.63	5.07	0.712	1
Naïve Bayes	78.91	63.82	7.91	0.720	4
SVMs	79.53	69.74	5.27	0.625	5
<i>k</i> -NN	80.93	29.17	13.59	0.768	3
<i>t</i> -value	91.62**	7.20**	5.27*	26.24**	

*Represents the level of significance is higher than 95% by *t*-test.

**Represents the level of significance is higher than 99% by *t*-test.

Table 5. Prediction performance of combined clustering and classification techniques

Model	Accuracy	Type I error	Type II error	ROC curve	Ranking by accuracy
<i>k</i> -means + MLP	84.66	59.78	5.67	0.749	2
<i>k</i> -means + DT	85.05	64.13	4.26	0.692	1
<i>k</i> -means + Naïve Bayes	82.72	68.48	6.14	0.720	5
<i>k</i> -means + SVM	82.33	94.56	0.95	0.522	4
<i>k</i> -means + <i>k</i> -NN	84.66	41.30	9.69	0.764	2
<i>t</i> -value	149.06**	7.65**	3.77*	15.80**	

*Represents the level of significance is higher than 95% by *t*-test.

**Represents the level of significance is higher than 99% by *t*-test.

Table 6. Prediction performance of the MLP and classification techniques combined

Model	Accuracy	Type I error	Type II error	ROC curve	Ranking by accuracy
MLP + MLP	97.08	8.82	2.08	0.987	1
MLP + DT	97.08	16.18	1.04	0.923	2
MLP + Naïve Bayes	91.61	35.29	4.58	0.918	5
MLP + SVM	96.53	13.24	2.08	0.923	4
MLP + <i>k</i> -NN	96.90	10.29	2.08	0.946	3
<i>t</i> -value	90.22**	3.49*	4.04*	73.08**	

*Represents the level of significance is higher than 95% by *t*-test.

**Represents the level of significance is higher than 99% by *t*-test.

Table 7. Prediction performances of the DT and classification techniques combined

Model	Accuracy	Type I error	Type II error	ROC curve	Ranking by accuracy
DT + MLP	93.12	32.53	2.48	0.826	3
DT + DT	95.77	16.87	2.07	0.957	1
DT + Naïve Bayes	88.36	51.81	4.75	0.853	4
DT + SVM	87.83	61.45	3.72	0.674	5
DT + <i>k</i> -NN	94.89	18.72	2.89	0.907	2
<i>t</i> -value	55.75**	4.09*	6.66**	17.58**	

*Represents the level of significance is higher than 95% by *t*-test.

**Represents the level of significance is higher than 99% by *t*-test.

classification techniques, the DT and MLP classifiers are chosen as the first classifiers for comparison.

Table 5 shows the prediction performance of the hybrid models combining clustering and classification techniques, which present the significance level of performance difference is higher than 95% or 99% by *t*-test. Notably, the *k*-means by the four clusters (i.e. *k* = 4) are combined with the five classifiers, since this combination performs best.

Analytical results demonstrate that the prediction models by hybrid learning techniques perform better than any single classification technique in terms of prediction accuracy and the Type II error. Particularly, *k*-means + the DT classifier performs best. However, the prediction accuracies of *k*-means + the MLP and *k*-means + *k*-NN classifiers are very close to that of *k*-means + the DT classifier. That is, performance differences are less than 1%.

Table 8. ANOVA analysis of average accuracy of three methods (*p* value)

Method	Cluster + Classifiers	Classifier + Classifiers	Single classifiers
Cluster + Classifiers	1.000	0.000*	0.112
Classifier + Classifiers		1.000	0.000*
Single classifiers			1.000
<i>F</i> -value 82.689*			

*Represents the level of significance is higher than 99% by *t*-test or *F*-test.

Table 9. ANOVA analysis of Type I error of three methods (*p* value)

Method	Cluster + Classifiers	Classifier + Classifiers	Single classifiers
Cluster + Classifiers	1.000	0.001**	0.412
Classifier + Classifiers		1.000	0.014*
Single classifiers			1.000
<i>F</i> -value 12.824**			

*Represents the level of significance is higher than 95% by *t*-test.
 **Represents the level of significance is higher than 99% by *t*-test or *F*-test.

Table 10. ANOVA analysis of Type II error of three methods (*p* value)

Method	Cluster + Classifiers	Classifier + Classifiers	single classifiers
Cluster + Classifiers	1.000	0.290	0.299
Classifier + Classifiers		1.000	0.021*
single classifiers			1.000
<i>F</i> -value 5.427*			

*Represents the level of significance is higher than 95% by *t*-test or *F*-test.

For hybrid models combining multiple classification techniques, Tables 6 and 7 show the prediction performance of MLP and DT combined and the five classification techniques, respectively. All the techniques indicate the significant level of performance difference is higher than 95% or 99% by *t*-test.

When using the MLP classifier as the first classifier, the MLP + MLP classifier performs best in terms of prediction accuracy, Type I and II errors, and the ROC curve, followed by the MLP + DT classifier. On the other hand, when the DT classifier was used as the first classifier, the DT + DT classifier achieved the highest prediction accuracy, lowest Type I and II error rates, and best ROC curve. Again, these hybrid models combining multiple classification techniques outperform single classifiers.

To determine which method is superior, the best single and hybrid models are compared by demonstrating difference statistically via analysis of variance (ANOVA). Tables 8–10 present the ANOVA of average accuracy, type I error, and type II error. The *p*-value indicates the single, cluster + classifier, and classifier + classifier models are statistically different at 1% or 5% alpha level except the *p*-value between cluster + classifiers and classifier + classifier. Notably, the three models show a statistical difference of performance measures (*F*-value) at either 1% or 5% alpha level.

Moreover, Figures 1–4 compare the best single and hybrid learning models in terms of prediction

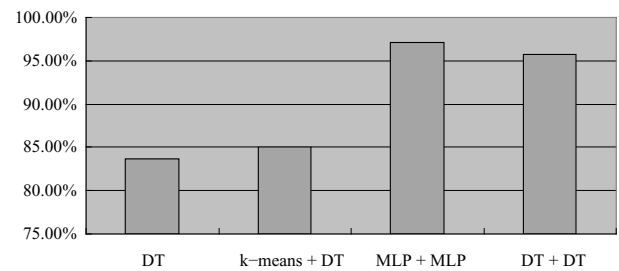


Fig. 1. Prediction accuracy of the best single and hybrid models

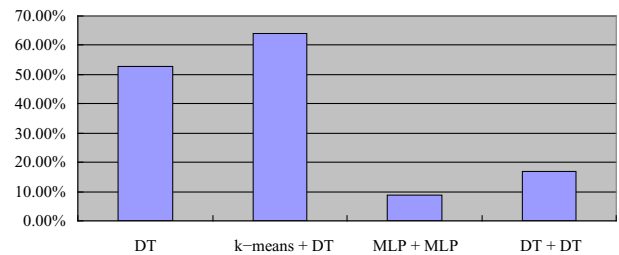


Fig. 2. Type I error of the best single and hybrid models

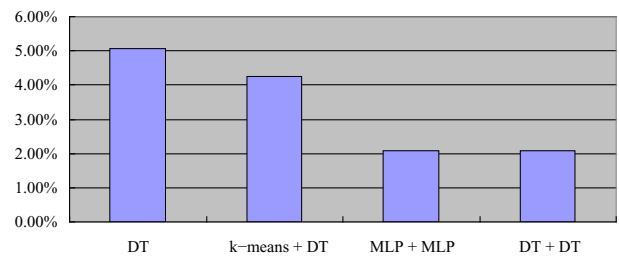


Fig. 3. Type II error of the best single and hybrid models

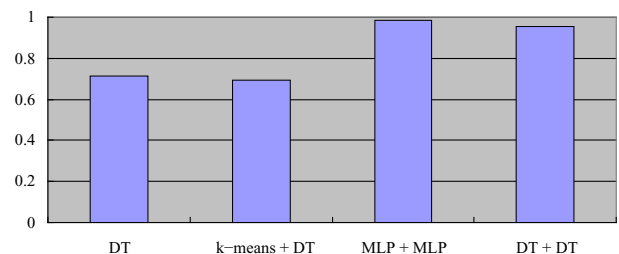


Fig. 4. ROC curve of the best single and hybrid models

accuracy, Type I and Type II errors, and the ROC curve, respectively. According to these comparison results, the MLP + MLP classifier is the best prediction model, achieving the highest prediction accuracy rate, lowest Type I and II error rates, and highest ROC curve, followed by the DT + DT model, indicating that hybrid learning models perform better than single learning models, and that multiple classification techniques combined outperform clustering and classification techniques combined.

Conclusions

Based on the spirit of partnership, Taiwan's governments function as promoters by building and operating public infrastructure or buildings with minimal out-of-pocket expense but full administrative support. For government agencies, the advantages of identifying dispute propensity early include reducing the time and effort needed to prepare a rule set to prevent disputes by improving the understanding of governments, private investors, and financial institutions of each side in a potential dispute.

This study compares 20 different classifiers using single and hybrid machine learning techniques. The best single model is the DT, achieving a prediction accuracy of 83.72%, followed by the MLP at 82.33%. For hybrid models, the combination of the k -means clustering algorithm and DT outperforms the combination of k -means and the other single classification techniques, including SVMs, the naïve Bayes classifier, and k -NN by achieving a prediction accuracy of 85.05%. Notably, all hybrid models (clustering + classification) perform better than single models.

Moreover, the hybrid models combining multiple classification techniques perform even better than that combining k -means and a DT. Specifically, the combination with multiple MLP classifiers and multiple DT classifiers outperforms other hybrid models, achieving prediction accuracy of 97.08% and 95.77%, respectively. Additionally, combining MLP classifiers is the best hybrid model based on having the highest prediction accuracy, lowest Type I and II error rates, and best ROC curve.

This study comprehensively compared the effectiveness of various machine learning techniques. Future work can focus on integration of proactive strategy deployment and preliminary countermeasures in early warning systems for PPP project disputes. Another fertile research direction is the development of second model for use once dispute likelihood is identified. For dispute cases, such a model is needed to predict which dispute category and which resolution methods are likely to be used during which phases of a project's lifecycle by mapping hidden classification or association rules.

References

- Arditi, D.; Pulket, T. 2005. Predicting the outcome of construction litigation using boosted decision trees, *Journal of Computing in Civil Engineering ASCE* 19(4): 387–393.
[http://dx.doi.org/10.1061/\(ASCE\)0887-3801\(2005\)19:4\(387\)](http://dx.doi.org/10.1061/(ASCE)0887-3801(2005)19:4(387))
- Arditi, D.; Pulket, T. 2010. Predicting the outcome of construction litigation using an integrated artificial intelligence model, *Journal of Computing in Civil Engineering ASCE* 24(1): 73–80.
[http://dx.doi.org/10.1061/\(ASCE\)0887-3801\(2010\)24:1\(73\)](http://dx.doi.org/10.1061/(ASCE)0887-3801(2010)24:1(73))
- Arditi, D.; Tokdemir, O. B. 1999a. Comparison of case-based reasoning and artificial neural networks, *Journal of Computing in Civil Engineering ASCE* 13(3): 162–169.
[http://dx.doi.org/10.1061/\(ASCE\)0887-3801\(1999\)13:3\(162\)](http://dx.doi.org/10.1061/(ASCE)0887-3801(1999)13:3(162))
- Arditi, D.; Tokdemir, O. B. 1999b. Using case-based reasoning to predict the outcome of construction litigation, *Computer-Aided Civil and Infrastructure Engineering* 14(6): 385–393.
<http://dx.doi.org/10.1111/0885-9507.00157>
- Arditi, D.; Oksay, F. E.; Tokdemir, O. B. 1998. Predicting the outcome of construction litigation using neural networks, *Computer-Aided Civil and Infrastructure Engineering* 13(2): 75–81.
<http://dx.doi.org/10.1111/0885-9507.00087>
- Bishop, C. M. 1995. *Neural networks for pattern recognition*. Oxford: Oxford University Press. 504 p.
- Chau, K. W. 2007. Application of a pso-based neural network in analysis of outcomes of construction claims, *Automation in Construction* 16(5): 642–646.
<http://dx.doi.org/10.1016/j.autcon.2006.11.008>
- Chen, J.-H. 2008. KNN based knowledge-sharing model for severe change order disputes in construction, *Automation in Construction* 17(6): 773–779.
<http://dx.doi.org/10.1016/j.autcon.2008.02.005>
- Chen, J.-H.; Hsu, S. C. 2007. Hybrid ANN-CBR model for disputed change orders in construction projects, *Automation in Construction* 17(1): 56–64.
<http://dx.doi.org/10.1016/j.autcon.2007.03.003>
- Chen, K.-Y. 2007. Forecasting systems reliability based on support vector regression with genetic algorithms, *Reliability Engineering & System Safety* 92(4): 423–432.
<http://dx.doi.org/10.1016/j.ress.2005.12.014>
- Cheng, M.-Y.; Tsai, H.-C.; Chiu, Y.-H. 2009. Fuzzy case-based reasoning for coping with construction disputes, *Expert Systems with Applications* 36(2): 4106–4113.
<http://dx.doi.org/10.1016/j.eswa.2008.03.025>
- Chou, J.-S.; Chiu, C.-K.; Farfoura, M.; Al-Taharwa, I. 2011. Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data mining techniques, *Journal of Computing in Civil Engineering ASCE* 25(3): 242–253.
[http://dx.doi.org/10.1061/\(ASCE\)CP.1943-5487.0000088](http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000088)
- Chou, J.-S.; Tai, Y.; Chang, L.-J. 2010. Predicting the development cost of tft-lcd manufacturing equipment

- with artificial intelligence models, *International Journal of Production Economics* 128(1): 339–350.
<http://dx.doi.org/10.1016/j.ijpe.2010.07.031>
- El-Adaway, I. H.; Kandil, A. A. 2010. Multiagent system for construction dispute resolution (MAS-COR), *Journal of Construction Engineering and Management* ASCE 136(3): 303–315.
[http://dx.doi.org/10.1061/\(ASCE\)CO.1943-7862.0000144](http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000144)
- Ferri, C.; Hernández-Orallo, J.; Modroiu, R. 2009. An experimental comparison of performance measures for classification, *Pattern Recognition Letters* 30(1): 27–38. <http://dx.doi.org/10.1016/j.patrec.2008.08.010>
- Han, J.; Kamber, M. 2001. *Data mining: concepts and techniques*. San Francisco: Morgan Kaufmann Publishers. 744 p.
- Hartigan, J. A.; Wong, M. A. 1979. Algorithm AS 136: a k-means clustering algorithm, *Applied Statistics* 28(1): 100–108. <http://dx.doi.org/10.2307/2346830>
- Haykin, S. 1999. *Neural networks: a comprehensive foundation*. 2nd ed. New Jersey: Prentice Hall. 842 p.
- Hornig, M.-H. 2010. Performance evaluation of multiple classification of the ultrasonic supraspinatus images by using ml, rbfn and svm classifiers, *Expert Systems with Applications* 37(6): 4146–4155.
<http://dx.doi.org/10.1016/j.eswa.2009.11.008>
- Huang, C. F.; Hsueh, S. L. 2010. Customer behavior and decision making in the refurbishment industry—a data mining approach, *Journal of Civil Engineering and Management* 16(1): 75–84.
<http://dx.doi.org/10.3846/jcem.2010.07>
- Jain, A. K.; Duin, R. P. W.; Mao, J. 2000. Statistical pattern recognition: a review, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1): 4–37.
<http://dx.doi.org/10.1109/34.824819>
- Kassab, M.; Hegazy, T.; Hipel, K. 2010. Computerized dss for construction conflict resolution under uncertainty, *Journal of Construction Engineering and Management* ASCE 136(12): 1249–1257.
[http://dx.doi.org/10.1061/\(ASCE\)CO.1943-7862.0000239](http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000239)
- Kim, H.-J.; Shin, K.-S. 2007. A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets, *Applied Soft Computing* 7(2): 569–576.
<http://dx.doi.org/10.1016/j.asoc.2006.03.004>
- Kim, Y. S. 2010. Performance evaluation for classification methods: a comparative simulation study, *Expert Systems with Applications* 37(3): 2292–2306.
<http://dx.doi.org/10.1016/j.eswa.2009.07.043>
- Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection, in *The International Joint Conference on Artificial Intelligence*, Montreal, Quebec, Canada: Morgan Kaufmann, 1137–1143.
- Lee, J.-R.; Hsueh, S.-L.; Tseng, H.-P. 2008. Utilizing data mining to discover knowledge in construction enterprise performance records, *Journal of Civil Engineering and Management* 14(2): 79–84.
<http://dx.doi.org/10.3846/1392-3730.2008.14.2>
- Lee, M.-C. 2009. Using support vector machine with a hybrid feature selection method to the stock trend prediction, *Expert Systems with Applications* 36(8): 10896–10904.
<http://dx.doi.org/10.1016/j.eswa.2009.02.038>
- Lenard, M. J.; Madey, G. R.; Alam, P. 1998. The design and validation of a hybrid information system for the auditor's going concern decision, *Journal of Management Information Systems* 14(4): 219–237.
- Li, L.; Jiang, W.; Li, X.; Moser, K. L.; Guo, Z.; Du, L.; Wang, Q.; Topol, E. J.; Wang, Q.; Rao, S. 2005. A robust hybrid between genetic algorithm and support vector machine for extracting an optimal feature gene subset, *Genomics* 85(1): 16–23.
<http://dx.doi.org/10.1016/j.ygeno.2004.09.007>
- Malinowski, P.; Ziembicki, P. 2006. Analysis of district heating network monitoring by neural networks classification, *Journal of Civil Engineering and Management* 12(1): 21–28.
- Marzouk, M.; El-Mesteckawi, L.; El-Said, M. 2011. Dispute resolution aided tool for construction projects in egypt, *Journal of Civil Engineering and Management* 17(1): 63–71.
<http://dx.doi.org/10.3846/13923730.2011.554165>
- Min, S.-H.; Lee, J.; Han, I. 2006. Hybrid genetic algorithms and support vector machines for bankruptcy prediction, *Expert Systems with Applications* 31(3): 652–660.
<http://dx.doi.org/10.1016/j.eswa.2005.09.070>
- Mitchell, T. 1997. *Machine learning*. New York: McGraw Hill. 432 p.
- Nandi, S.; Badhe, Y.; Lonari, J.; Sridevi, U.; Rao, B. S.; Tambe, S. S.; Kulkarni, B. D. 2004. Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: study of benzene isopropylation on hbeta catalyst, *Chemical Engineering Journal* 97(2–3): 115–129.
[http://dx.doi.org/10.1016/S1385-8947\(03\)00150-5](http://dx.doi.org/10.1016/S1385-8947(03)00150-5)
- PCC. 2011. *Engineering evaluation forum of ppp strategy (in Chinese)* [online]. Public Construction Commission, Executive Yuan, [cited 5 May 2011]. Available from Internet: <http://ppp.pcc.gov.tw/PPP/frontplat/search/showViews.do?indexID=0&PK=1002>.
- Pulket, T.; Arditi, D. 2009a. Construction litigation prediction system using ant colony optimization, *Construction Management and Economics* 27(3): 241–251.
<http://dx.doi.org/10.1080/01446190802714781>
- Pulket, T.; Arditi, D. 2009b. Universal prediction model for construction litigation, *Journal of Computing in Civil Engineering* ASCE 23(3): 178–187.
[http://dx.doi.org/10.1061/\(ASCE\)0887-3801\(2009\)23:3\(178\)](http://dx.doi.org/10.1061/(ASCE)0887-3801(2009)23:3(178))
- Quinlan, J. R. 1993. *C4.5: programs for machine learning*. San Francisco: Morgan Kaufmann. 302 p.
- Sokolova, M.; Lapalme, G. 2009. A systematic analysis of performance measures for classification tasks, *Information Processing and Management* 45(4): 427–437. <http://dx.doi.org/10.1016/j.ipm.2009.03.002>
- Tan, P.-N.; Steinbach, M.; Kumar, V. 2006. *Introduction to data mining*. London: Pearson Education, Inc. 769 p.
- Tsai, C.-F.; Chen, M.-L. 2010. Credit rating by hybrid machine learning techniques, *Applied Soft Computing* 10(2): 374–380.
<http://dx.doi.org/10.1016/j.asoc.2009.08.003>
- Vapnik, V. N. 1998. *Statistical learning theory*. New York: John Wiley and Sons. 736 p.

- WBG. 2011. [online]. The World Bank Group [cited 5 April 2011]. Available from Internet:
http://ppi.worldbank.org/resources/ppi_glossary.aspx.
- Witten, I. H.; Frank, E. 2005. *Data mining: practical machine learning tools and techniques*. 2nd ed. San Francisco: Morgan Kaufmann. 664 p.
- Wu, C.-H.; Tzeng, G.-H.; Lin, R.-H. 2009. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression, *Expert Systems with Applications* 36(3): 4725–4735.
<http://dx.doi.org/10.1016/j.eswa.2008.06.046>
- Wu, Q. 2010. The hybrid forecasting model based on chaotic mapping, genetic algorithm and support vector machine, *Expert Systems with Applications* 37(2): 1776–1783.
<http://dx.doi.org/10.1016/j.eswa.2009.07.054>

Jui-Sheng CHOU. He received his Bachelor's and Master's degrees from National Taiwan University, and PhD in Construction Engineering and Project Management from The University of Texas at Austin. Chou is a professor in the Department of Construction Engineering at National Taiwan University of Science and Technology. He has over a decade of practical experience in engineering management and consulting services for the private and public sectors. He is a member of several international and domestic professional organizations. His teaching and research interests primarily involve Project Management (PM) related to knowledge discovery in databases (KDD), data mining, decision, risk & reliability, and cost management.

Chih-Fong TSAI. He received a PhD at School of Computing and Technology from the University of Sunderland, UK in 2005. He is now an associate professor at the Department of Information Management, National Central University, Taiwan. He has published more than 50 technical publications in journals, book chapters, and international conference proceedings. He received the Highly Commended Award (Emerald Literati Network 2008 Awards for Excellence) from Online Information Review, and the award for top 10 cited articles in 2008 from Expert Systems with Applications. His current research focuses on multimedia information retrieval and data mining.

Yu-Hsin LU. She received her PhD in Accounting and Information Technology from National Chung Cheng University, Taiwan. She is an assistant professor at the Department of Accounting, Feng Chia University, Taiwan. Her research interests focus on data mining applications and financial information systems.