Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). "Predicting business failure of construction contractors using long short-term memory recurrent neural network." Journal of Construction Engineering and Management, doi.org/10.1061/(ASCE)CO.1943-7862.0001709

Predicting Business Failure of Construction Contractors using Long Short-Term Memory Recurrent Neural Network

Youjin Jang[1], In-Bae Jeong[2], Yong K. Cho[3] and Yonghan Ahn[4]

[1]Postdoctoral Research Associate, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, Email: uzjang@gatech.edu

[2]Postdoctoral Research Associate, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, Email: inbae.jeong@gatech.edu

[3]Associate Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, E-mail: yong.cho@ce.gatech.edu (Corresponding author)

[4]Associate Professor, School of Architecture and Architectural Engineering, Hanyang University ERICA, Ansan-si 15588, Republic of Korea, Email: yhahn@hanyang.ac.kr

Abstract

Predicting business failure of construction contractors is critical for both contractors themselves and other stakeholders such as project owners, surety underwriters, investors and government entities. In an attempt to identify a new model with better prediction of business failure of the construction contractors, this study utilized long short-term memory (LSTM) recurrent neural network (RNN). The financial ratios of the construction contractors in the United States are collected, and SMOTE+Tomek links is employed to obtain a balanced dataset. The proposed LSTM RNN model is evaluated by comparing its accuracy and F1-score with feedforward neural network (FNN) and support vector machine (SVM) models for the optimized parameters selected from a grid search with five-fold cross-validation. The results successfully demonstrate that the prediction performance of the proposed LSTM RNN model outperforms FNN and SVM models for both test and original dataset. Therefore, the proposed LSTM RNN model is a promising alternative to assist managers, investors, auditors, and government entities in predicting business failure of construction contractors which can be also adapted to other industry cases.

Keywords: Business failure; Construction contractors; Prediction model; Long Short-Term Memory (LSTM); Recurrent Neural Network (RNN).

Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). "Predicting business failure of construction contractors using long short-term memory recurrent neural network." Journal of Construction Engineering and Management, **doi.org/10.1061/(ASCE)CO.1943-7862.0001709**

Introduction

Predicting business failure of construction contractors has been a critical and challenging issue for both contractors themselves and stakeholders (Yeh et al. 2010, Tserng et al. 2011, Horta and Camanho 2013, Cheng and Hoang 2015). Since the construction business is intertwined in a complicated system involving numerous subcontractors, business failure of a single construction contractor could lead a chain reaction of another business failure (Tserng et al. 2011, Heo and Yang 2014). Predicting business failure of construction contractors is essential for project owners to avoid contracting with the contractors with the high probability of business failure (Tserng et al. 2011). It is also crucial for surety underwriters, investors and financial lending institutions because increased business failure can be a significant burden to them (Horta and Camanho 2013).

US construction market value is $1.2 trillion in 2016 (US Census Bureau 2016), and global construction output is $ 8.5 trillion in 2015, and $10 trillion by 2020 (Market Report Store 2016). As such, the construction industry has been one of the largest contributors to the U.S. economy as well as global economy. Nevertheless, compared to other sectors of the economy, construction contractors are vulnerable to business failure due to the characteristics of the construction industry, such as the uniqueness of projects, the long time period of the project completion, the complexity of the construction process, the integration of different types of companies, and the uncertainty and risk of the construction activity (Kangari et al. 1992, Cheng et al. 2014, Tserng et al. 2012). Recently, construction contractors are increasingly likely to fail in their business because of the significant structural changes such as globalization, technological evaluation, increased competition and regulation (Horta and Camanho 2013). These situations highlight the importance of developing a robust model to enable the business failure prediction of construction contractors.

There is abundant literature on the development of business failure prediction models for construction contractors. The models of predicting business failure usually employed financial data as the input variable and were based on statistical techniques, such as multiple regression (Kangari et al. 1992, Russell and Zhai 1996), multiple discriminant analysis (Mason and Harris 1979, Abidali and Harris 1995), logistic regression (Aelelye et al. 2013, Tserng et al. 2014). In recent years, business failure prediction methods have been gradually shifting from traditional statistical models to soft computing, which is part of a separate field of science defining as computational intelligence. Prediction models found in the literature include artificial neural networks (Al-Sobiei et al. 2005), support vector machines (Lam et al. 2009, Horta and Camanho 2013) and numerous hybrid models (Chen 2012, Cheng et al. 2014, Cheng and Hoang 2015, Tserng et al. 2015). However, most of the existing studies have not considered the historical patterns of the input variables in the modeling. In other words, they predict the business failure with only information at a certain point in time. For example, if predicting the business failure of the year 2019, the existing studies only used financial information of the year 2018. The importance of predictive variables varies over time since the financial information used as predictive variables of a company changes over time (Grice 2001). Therefore, it is necessary to consider the changes over time when developing a business failure prediction model. The performance of a prediction model can be improved if predictive variables include information for several periods (Niemaan et al. 2008).

To address this issue, this study employs long short-term memory (LSTM) recurrent neural network (RNN) to determine the temporal patterns in the data associated with the business failure of construction contractors. RNN is one of deep learning algorithms, which can effectively learn sequential patterns from data containing temporal or sequential information. In particular, LSTM, a kind of RNN, can solve the vanishing gradient and exploding gradient problem of standard RNN. In this regard, this study proposes the business failure prediction model for construction contractors using LSTM RNN. This study aims to improve the business failure prediction performance of the construction contractors. Considering data accessibility, this study employs U.S. construction contractors listed on the New York Stock Exchange (NYSE), NASDAQ, and the American Exchange (AMEX), and their financial data is used to predict business failure of construction contractors. To evaluate the prediction performance of the proposed LSTM RNN model, feedforward neural network (FNN) and support vector machine (SVM) models, which are commonly used in the business failure prediction literature, are served as benchmarks. The proposed model is expected to be able to help decision-makers in dealing with of construction contractor financial status. It can also be used to assist investors, auditors, and government entities to predict business failure of construction contractors in the United States.

The remainder of this paper is organized as follows. The next section reviews the literature on the prediction models of business failure and LSTM RNN. Then, dataset preparation, prediction model development, and performance evaluation are described. Lastly, this study provides test results followed by discussions and conclusions with suggestions for a future study.

Literature Review

Prediction models of Business Failure

Business failure refers to cease operations following its inability to make a profit or to bring in enough revenue to cover its expenses (Adeleye et al. 2013). Terms such as bankruptcy, default, distress, and insolvency are usually used interchangeably in the literature. The increasing scale and technical complexity of construction projects often lead to the collaboration of many stakeholders to accomplish the projects (Lam et al. 2009). If a construction contractor’s business fails, other stakeholders will also be affected. In this respect, business failure prediction has attracted considerable attention in the construction industry.

Most traditional business failure prediction models are based on statistical techniques. These models usually use financial data as predictive variables. Kangari et al. (1992) and Russell and Zhai (1996) presented a model for predicting financial performance using multiple regression analysis. Mason and Harris (1979) and Abidali and Harris (1995) attempted to predict business failure of construction contractors with the multiple discriminant analysis (MDA). Adelelye et al. (2013) and Tserng et al. (2014) constructed a business failure prediction model based on logistic regression (LR).

The advances in artificial intelligent (AI) techniques have demonstrated that data mining techniques outperform traditional statistical techniques. Al-Sobiei et al. (2005) utilized the artificial neural networks (ANN) to estimate the risk of a construction contractor’s default. Lam et al. (2009) presented the support vector machine (SVM) method for contractor prequalification transactions and compared with ANN. Horta and Camanho (2013) employed the SVM to develop decision support systems for assessing contractor’s financial status and showed the results which outperformed the logistic regression.

Recent research trends have also employed hybrid methodologies successfully. Chen (2012) integrated the concepts of self-organizing feature map optimization, fuzzy, and hyper-rectangular composite neural networks (SFNN) to provide a new method for forecasting the financial distress of construction contractors. Cheng et al. (2014) presented the evolutionary least squares support vector machine inference model for predicting contractor default status (ELSIM-PCDS) by hybridizing the synthetic minority over-sampling technique (SMOTE), least squares support vector machine (LS-SVM), and differential evolution (DE) algorithms. Cheng and Hoang (2015) proposed a hybrid fuzzy instance-based classifier for construction contractor default prediction by incorporating the fuzzy k-nearest neighbor classifier (FKNC), the synthetic minority over-sampling technique (SMOTE), and the firefly algorithm (FA). Tserng et al. (2015) suggested a model for predicting the default of construction contractors based on LS-SVM and grey system theory.

Although many researchers have attempted to develop prediction models using various techniques such as statistical, AI and hybrid methodologies as an effort to improve prediction performance, they have not considered the historical patterns of the predictive variables in the modeling. Most existing studies used the financial data which are time-series as the predictive variables. The importance of predictive variables generally changes over time (Grice 2001). However, previous studies in the construction industry predicted business failure with information at a certain point in time. The prediction models can be improved performance if predictive variables contain financial information over several periods (Niemaan et al. 2008). In this respect, this study predicts business failure of construction contractors using the LSTM RNN which can take into account changes in financial information by providing the output of the previous time step as input for the next step.

Long Short-Term Memory Recurrent Neural Network (LSTM RNN)

Recurrent Neural Network (RNN), which is a neural network with recurrent connections, has been developed in order to learn sequential and temporal patterns from time-series or sequences of data. As illustrated in Fig. 1, recurrent connections, called feedback connections, can reflect temporal information during training. Here, x_t is the input layer, h_t is the hidden layer and y_t is the output layer. The calculated h_t value is used to calculate y_t value as shown in Equation (1) and to cacluate the hidden state h_{t+1} at the next time $t+1$ concurrently. Hence, the temporal pattern can be learned. The h_t at time t is calculated by receiving input information at time $t(x_t)$ and using the previous hidden state vector (h_{t-1}) as following Equation (2).

$$y_t = f_1(Vh_t + b_1) \quad (1)$$

$$h_t = f_2(Ux_t + Wh_{t-1} + b_2) \quad (2)$$

Where U, V and W are weight matrices and b_1, b_2 are bias vectors. f_1, f_2 are activation functions such as the sigmoid function or hyperbolic tangent function.

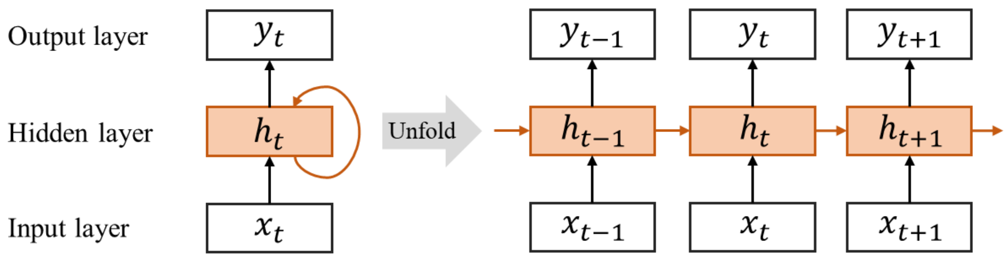

Fig. 1. The architecture of RNN

As RNNs are very deep neural networks in the time direction, vanishing or exploding gradients can occur (Bengio et al. 1994). It also can store short-term memory but are vulnerable to long-term dependency in terms of time. Hochreiter and Schmidhuber (1997) proposed LSTM to hold short-term memory for a longer period. As shown in Fig. 2(a), the LSTM RNN is comprised of LSTM memory blocks instead of the hidden neurons in the RNN. These LSTM memory blocks consist of memory cells and gates and play an important role in training long-range dependency while controlling information storage and flow. The LSTM memory block has one memory cell ($\boldsymbol{c_t}$) and four gates which are input ($\boldsymbol{i_t}$), forget ($\boldsymbol{f_t}$), output gates ($\boldsymbol{o_t}$), and input modulation gate ($\boldsymbol{g_t}$) as illustrated in Fig. 2 (b). The gates are used as a mechanism to determine the information that is able to be received by the cell. The memory cell in each gate consists of a sigmoid neural net layer and a pointwise multiplication operation. The sigmoid layer outputs numbers between zero and one and describes how much of each component can be forwarded to the cell. For time step $\boldsymbol{t}$, the cell state can be updated by using the following Equations (3) – (8).

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma\left(U_f x_t + W_f h_{t-1} + b_f\right) \tag{4}$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \tag{5}$$

$$g_t = tanh\left(U_g x_t + W_g h_{t-1} + b_g\right) \tag{6}$$

$$c_t = f_i \otimes c_{t-1} + i_i \otimes g_t \tag{7}$$

$$h_t = o_i \otimes \tan h\,(c_t) \tag{8}$$

Where, σ is activate function sigmoid defined as $\sigma(x) = (1 + e^{-x})^{-1}$, i_t, f_t, o_t, g_t, c_t are vectors for the outputs of the 'input', 'forget', 'output' and 'input modulation' gates, cell and the LSTM layer at time t, respectively. h_i, b_i, b_f, b_o, and b_g are offset vector, U_i, U_f, U_o, U_g, W_i, W_f, W_o and W_g are the coefficient matrix.

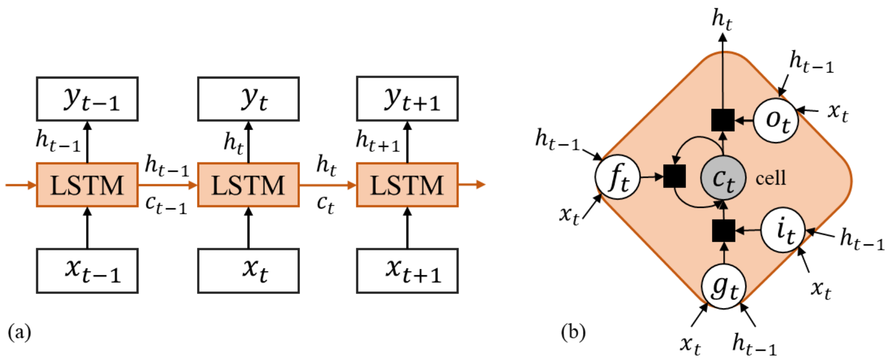

Fig. 2. The architecture of LSTM RNN: (a) LSTM architecture with one hidden LSTM layer; and (b) LSTM memory block

Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). “Predicting business failure of construction contractors using long short-term memory recurrent neural network.” Journal of Construction Engineering and Management, doi.org/10.1061/(ASCE)CO.1943-7862.0001709

Methodology

Dataset Preparation

The financial ratio serves as the fundamental basis for evaluating the company’s financial capabilities and provides useful information for predicting business failure of construction contractors (Tserng et al. 2012). The value of the financial ratio of construction contractors is significantly different from those of other industries due to the characteristics of the construction industry. For example, construction contractors’ values may include total firm assets because they deal with large projects. As a result, the capital structures of construction contractors are entirely different from other types of industries. Also, the probability of business failure of construction contractors may differ from other industries even though the financial ratio are the same because different industries face different levels of competition. The degree of financial risk of a company may vary depending on the industries in which they are based even if they have the same value of the fianaical ratio.

Therefore, this study considers that the characteristics of the construction industry are reflected in the value of the financial ratio of the construction contractors, and twelve financial ratios are selected as input variables : (1) return on asset; (2) return on equity; (3) return on sales; (4) current ratio; (5) current assets to net assets; (6) working capital to total asset; (7) total liabilities to net worth; (8) retained earnings to sales; (9) debt ratio; (10) working capital turnover; (11) equity turnover; and (12) total asset turnover. These twelve ratios are commonly used in the previous studies regarding contractors’ business failure models, with the aspects of profitability, liquidity, leverage and activity (Tserng et al. 2011, Chen 2012, Tsai et al. 2012, Tserng et al. 2012, Bal and Wu 2013, Horta and Camanho 2013, Cheng et al. 2014, Heo and Yang 2014, Tserng et al. 2014, Tserng et al. 2015, Cheng and Hoang 2015). The detailed equations used for computing financial ratios and the descriptive statistics of financial ratios can be seen in Table 1 and Table 2.

Table 1. Definition of input variables

Accounting variable	Calculation	Previous studies which employed a variable*
Profitability		
(1) Return on asset	Net income/total assets	1,2,3,4,5,6,7,8,9,10,11
(2) Return on equity	Net income/shareholder’s equity	1,2,3,4,5,7,9,10,11
(3) Return on sales	Net income/net sale	1,3,4,5,6,7,10,11
Liquidity		
(4) Current ratio	Current assets/current liabilities	1,3,4,5,7,9,10,11
(5) Current assets to net assets	Current assets/(total assets-current liabilities)	1,3,4,5,10,11

(6) Working capital to total asset	(current asset-current liabilities)/total assets	1,2,3,4,5,7,8,10,11
Leverage		
(7) Total liabilities to net worth	Total liabilities/shareholder's equity	1,2,4,5,10,11
(8) Retained earnings to sales	Retained earnings/net sales	1,2,4,5,10,11
(9) Debt ratio	Total liabilities/total assets	1,2,3,4,5,9,10,11
Activity		
(10) Working capital turnover	Net sales/(current assets-current liabilities)	1,3,4,5,10,11
(11) Equity turnover	Net sales/shareholder's equity	1,2,3,4,5,10,11
(12) Total asset turnover	Net sales/total assets	1,2,3,4,5,8,9,10,11

*Note. 1: Cheng & Hoang (2015), 2: Tserng et al. (2011), 3: Tserng et al. (2012), 4: Tserng et al. (2014), 5. Tserng et al. (2015), 6: Chen (2012), 7: Horta & Camanho (2013), 8: Heo & Yang (2014), 9: Bal & Wu (2013), 10: Cheng et al. (2014), 11: Tsai et al. (2012)

Table 2. Descriptive statistics of input variables

Variable	Min.	Max.	Mean	S.D
(1) Return on Asset	-1457.00	7.07	-0.95	36.74
(2) Return on Equity	-189.62	21.12	-0.18	5.41
(3) Return on Sales	-2722.00	451.00	-7.68	126.27
(4) Current ratio	0.00	136.37	3.71	6.52
(5) Current assets to net assets	-100.76	1909.09	4.54	51.67
(6) Working capital to total asset	-145.00	142.94	0.27	6.08
(7) Total liabilities to net worth	-137.14	387.29	2.33	14.25
(8) Retained earnings to sales	-10619.00	11608.00	-14.01	577.15
(9) Debt ratio	0.01	146.00	0.73	3.72
(10) Working capital turnover	-838.89	7433.50	12.53	196.33
(11) Equity turnover	-134.63	965.07	4.72	26.78
(12) Total asset turnover	-0.50	10.97	1.38	0.91

The data sample of construction contractors is obtained from selecting North America public listed firms in the Standard & Poor's COMPUSTAT database (Wharton Research Data Service). This study chooses the firm with Standard Industrial Classification (SIC) codes between 1500 and 1799 which include the following three construction categories: (1) building construction (SIC codes 1500 to 1599); (2) heavy construction (SIC codes 1600 to 1699); and (3) special trade construction (SIC codes 1700 to 1799). Data is collected from each construction contractor for the period from 1980 to 2016. To avoid sample selection biases, this study uses firm-year data for analysis. The construction contractors who do not have a continuous financial statement for at least five years are removed from the sample. Similar to the studies of Tserng et al. (2015), this study defines a construction contractor with the delisting codes between 550 and 585 assigned by the Center for Research in Securities Prices (CRSP) as a business failure contractors. These delisting codes represent bankruptcy, liquidity, or poor performance. As the goal of this study is to predict business failure within one year, the financial data in the year before being delisted is considered as data sample of business failure otherwise data sample of healthy.

The dataset used in this study consists of 1,337 healthy observations and 42 business failure observations when considering all the firm-year data. The number of healthy observations dramatically exceeds the number of business failure observations. To cope with this imbalance classification problem, this study uses SMOTE + Tomek links, which is a hybrid method that combines the SMOTE (Synthetic Minority Over-sampling Technique) and Tomek links to obtain a balanced dataset. SMOTE is an over-sampling method where the minority class is oversampled to generate new synthetic examples (Chawla et al. 2002). The basis idea is to create new synthetic examples that are located between each of the minority class examples and one of its *k*-nearest neighbors combining their feature vectors. Although over-sampling minority class examples can balance class distributions, class clusters are not well defined since some majority class examples might be invading the minority class spaces. The opposite can also be true since interpolating the minority class examples can expand the minority class clusters, introducing artificial minority class examples too deeply in the majority class space. Inducing a classifier under such a situation can lead to overfitting. In order to create better-defined class clusters, Tomek links is applied to the over-sampled training set as a data cleaning method (Tomek 1976). Thus, instead of removing only the majority class examples that form Tomek links, the examples from both classes are removed. After being processed by SMOTE+Tomek links, the ratio of healthy and business failure observations became 13370 : 13370 and there were 26,740 observations used for analysis. An illustration of the original dataset and the dataset with SMOTE+Tomik links in three-dimensional space is provided in Fig. 3. The X, Y and Z axes in Fig. 3 represnet the most principal components extracted from Principal Components Analysis (PCA).

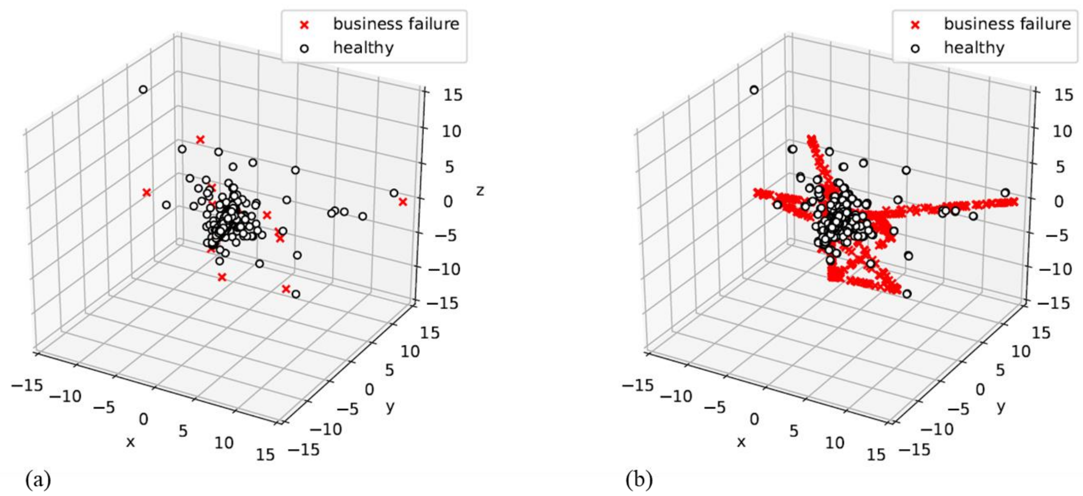

Fig. 3. Illustration of the dataset in three-dimensional space: (a) Original dataset; and (b) Resampled dataset using SMOTE+Tomek links

Model Development

This study develops a business failure prediction model using LSTM RNN as depicted in Fig 4. The prediction model uses dataset of three years as input, which are x_{t-2}, x_{t-1}, and x_t where t is the year of predicting business failure. We chose 12 input variables regarding the profitability (PR), liquidity (LI), leverage (LE), and activity (AC). As the numerical ranges of the input data, standardization, which is widely used for normalization, is employed. Standardization makes the values of each data have zero-mean and unit-variance, expressed in Equation (9).

$$X_{STAND} = \frac{X_{raw}-X_{mean}}{\sigma} \tag{9}$$

Where $X_{raw}$is the original data value, X_{mean} is the mean value of the variables, and σ is its standard deviation.

A given input value passes through the LSTM layer at time t, which is the most recent time, through calculation of the LSTM layer according to Equations (3) – (8). The value h_t passing through the LSTM layer at the last point in time t is the predicted class of the business failure in the next year. This prediction is output as y_t thought the softmax layer. Softmax function is used to map the non-normalized output to a probability distribution over predicted output classes. The output classes, which are healthy firm and business failure firm, are numbered 0 and 1, respectively. The value of y_t is calculated as following Equation (10).

$$z_t = W_z h_t + b_z \tag{10}$$

$$y_t = softmax(z_t) = \frac{1}{\sum_{k=0}^{1}\exp(z_{t,i})}\begin{bmatrix}\exp(z_{t,0})\\ \exp(z_{t,1})\end{bmatrix} \tag{11}$$

Where W_z is the weight matrix, b_z is a bias term and z_t is a three-dimensional vector from which the softmax function in Equation (1) is obtained. In that equation, $z_{t,i}$ is the *i*-th unit value of z_t, called a logit. The final $y_t$value is the probability of each class.

This study uses cross-entropy loss function, expressed in Equation (12).

$$\mathcal{L}(p_t, y_t) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{3} p_{t,i,n}\log(y_{t,i,n}) \tag{12}$$

Where *N* is the total number of training data elements, p_t is the target value, $p_{t,i,n}$ is the *i*-th value of the p_t of the *n*-th sample and $y_{t,i,n}$ is the *i*-th value of the y_t of the *n*-th sample.

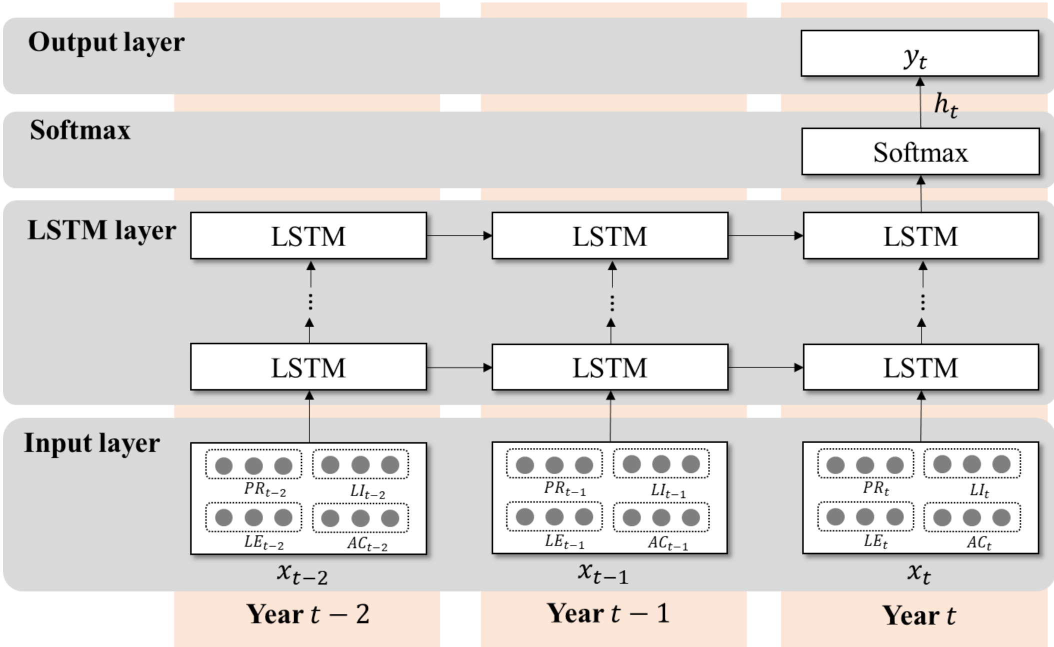

Fig. 4. The architecture of business failure prediction model using LSTM RNN

Performance Evaluation

To evaluate the performance of the prediction model using LSTM RNN, Feedforward Neural Network (FNN) and SVM were used as a comparison reference. The FNN, one of the most common structures of the artificial neural network (ANN), is generally composed of three layers of input, hidden and output. Each layer of the FNN has a certain number of nodes, and each node in a layer is connected to other nodes in the next layer with a specific weight and bias. The SVM, suggested by Vapnik (1995), is based on the structural risk minimization (SRM) rather than the empirical risk minimization (ERM) of the ANN. SVM first transforms the original input data into the high-dimensional feature space using nonlinear mapping (or an alternative using a kernel function) and then searches for a linear separating hyperplane in the new space. The optimal hyperplane found in the new space corresponds to a nonlinear separating hypersurface in the original space. This study uses the linear, radial basis function (RBF) and polynomial as the kernel functions of SVM model.

To measure the performance of the classification techniques, two different metrics, the accuracy, and F1-score, are used. They can be measured by a confusion matrix containing True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN), as shown in Table 3. The accuracy is defined as the number of correctly classified data elements among the total number of test data elements. Therefore, the average prediction accuracy is obtained by following Equation 13. The accuracy reaches its best at 1 and worst at 0.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+TN}} \tag{13}$$

The F1-score is the harmonic mean of the precision and recall. Precision measures how many of the points predicted as rare are in fact rare, whereas recall measures how many of the rare points are predicted to be rare. Both precision and recall are desirable, but they typically trade off against each other. The F1-score is measured as following Equation 14. The F1-score reaches its best at 1 and worst at 0.

$$\text{F1} - \text{score} = \frac{\text{2TP}}{\text{2TP+FP+FN}} \tag{14}$$

Table 3. Confusion Matrix

Actual	Predicted	
	Negative = class0 (healthy)	Positive = class1 (business failure)
Negative = class0	True Negative (TN)	False Positive (FP)
Positive = class1	False Negative (FN)	True Positive (TP)

Results

Experimental Setting

To avoid the randomness of selected testing cases, this study used a five-fold cross-validation approach when examining the prediction performance. As aforementioned, the original dataset was resampled by SMOTE+Tomek links. The resampled dataset was divided into the training and validation dataset (70%) and the test dataset (30%) as illustrated in Fig. 5. The training and validation dataset was divided randomly into five mutually exclusive folds. Each fold was used once to validate the performance, and the remaining four folds was used for training, which obtained five independent performance values. This procedure was repeated five times by changing the remaining folds, and five prediction performances were generated. The performance of the prediction model can be appraised by the average predictive results of the five folds. In this study, the five-fold cross-validation was used to evaluate the accuracy and F1-score of not only the proposed LSTM model but also the FNN and SVM models.

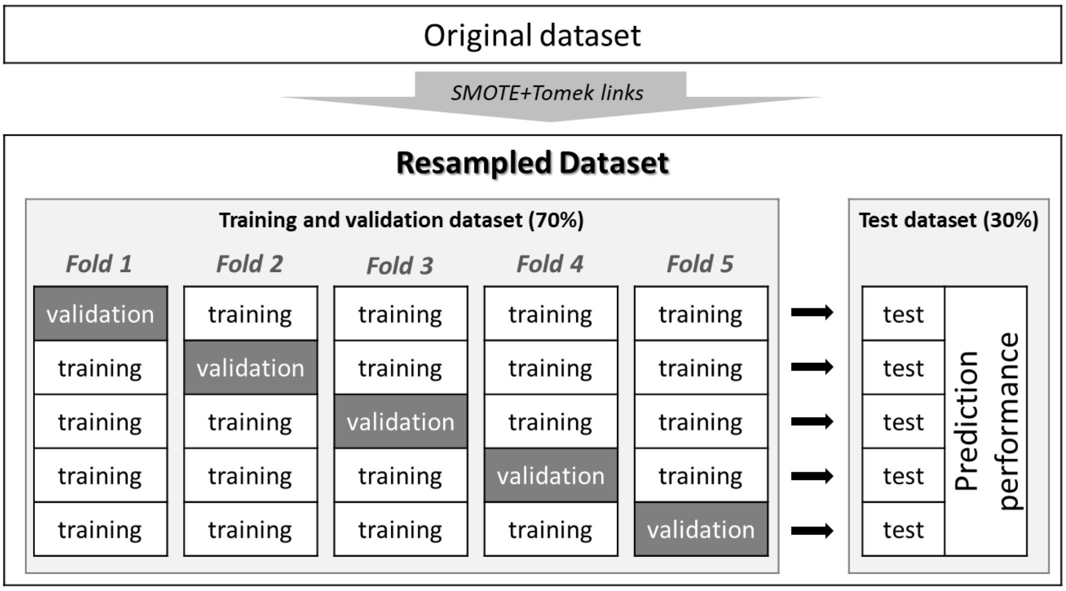

Fig. 5. Five-fold cross-validation

When comparing the performance of the proposed LSTM RNN model with FNN and SVM models, the parameter values of each classifier were optimized using a grid search. The basic concept of a grid search involves applying various parameter values and choosing the one with the best five-fold cross-validation performance. It is noted that when using the ANN, it is needed to set the number of hidden layers (HL) and the number of neurons in the hidden layer (NN) (Samarasinghe 2006). In order to obtain the optimal performance of the LSTM RNN model, this study set the grid space to HL $\in$ [1, 2, 3, 4, 5] and NN $\in$ [40, 80, 120, 160, 200]. The FNN model also set the grid space to HL $\in$ [1, 2, 3, 4, 5] and NN $\in$ [40, 80, 120, 160, 200]. When using the SVM, there are tuning parameters needed to be determined. For the SVM model with RBF kernel, this study set the grid space to 30 different penalty parameter (C) values and 30 different kernel parameter (γ) values, that makes C and γ evenly distributed in logarithmic space of the range [10^{-2}, 10^{-5}]. Grid search of SVM with linear kernel also used the same C value. The same C values and 30 different the kernel parameter (d) values were used in grid search of SVM with polynomial kernel.

In this study, the dropout probability was set to 0.5, and the learning model was trained for 100,000 epochs. The learning rate began from 0.001 and exponentially decayed to 0.00001 at the end of the iterations. The models were trained with a workstation equipped with four GPUs (CPU: Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40GHz, RAM: 64GB and VGA: NVIDIA TITAN Xp × 4ea).

Experimental Results

The results of a grid search with five-fold cross-validation are depicted in Fig. 6. Here, the brighter the cell color, the higher the F1-Score. The best F1-score for LSTM RNN model was found to be 0.986, with the corresponding values of the number of hidden layers as 2 and the number of neurons in the hidden layer as 160, as depicted in Fig. 6(a). It was found that the prediction performance of the LSTM RNN model is not primarily affected by the number of hidden layers, but rather by the number of neurons in the hidden layer. This is because the LSTM RNN model can memorize temporary or sequential information. The highest F1-score of the validation dataset of FNN model was 0.869, with the corresponding values of the number of hidden layers and the number of neurons in the hidden layer as 3 and 200, respectively (Fig. 6(b)). The optimum C and γ of the SVM model with RBF kernel were 4.52 and 2.59 respectively at 0.955, which is the highest F1-score in the validation dataset (Fig. 6(c)). The optimum C of SVM with linear kernel was 0.03 at 0.640, the highest F1-score. The best F1- score for the SVM model with polynomial kernel was found to be 0.850, with the corresponding value of C and d as 7.880 and 4, respectively (Fig. 6(d)). Meanwhile, it could be confirmed that overfitting of the LSTM RNN model does not occur through both accuracy curve and F1-score curve for the validation and test dataset, as can be seen in Fig. 7.

Using the optimized parameters for the validation dataset mentioned above, this study evaluated the prediction performance of the LSTM, FNN and SVM models for the test dataset. Table 4 shows the confusion matrix and Table 5 illustrates the accuracy and F1-score calculated based on Table 4. In Table 5, the higher values of accuracy and F1-score, the better performance. LSTM RNN model achieves the best accuracy of 0.982 for the testing dataset, while the testing results of the FNN model, SVM models with RBF, linear and polynomial kernels are 0.855, 0.956, 0.630 and 0.844 respectively. The best model for determining F1-score was found to be LSTM RNN (0.982), while the SVM model with linear kernel exhibited the worst predictive capabilities (0.582). These results indicate a significant improvement of the proposed LSTM RNN model over SVM model with linear kernel and a moderate improvement over SVM model with RBF kernel.

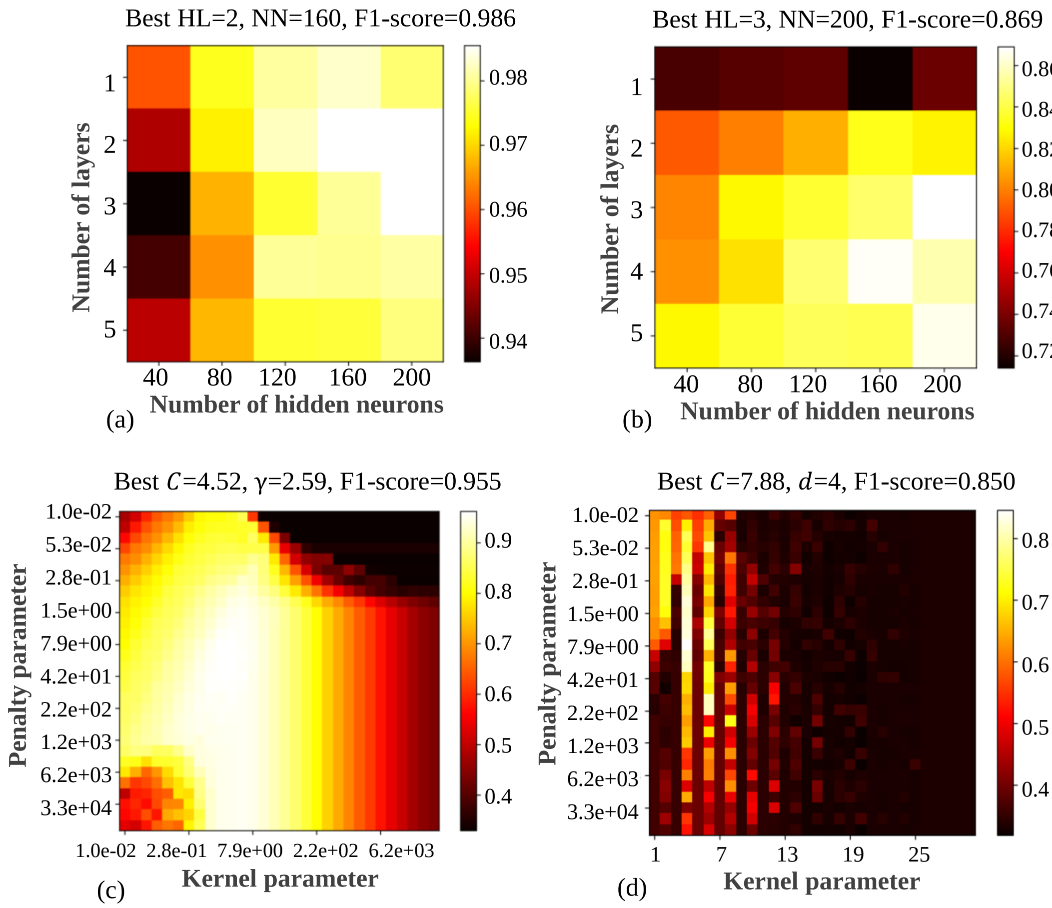

Fig. 6. Parameter selection by a grid search with five-fold cross-validation: (a) LSTM RNN model; (b) FNN model; (c) SVM model with RBF kernel; and (d) SVM model with polynomial kernel

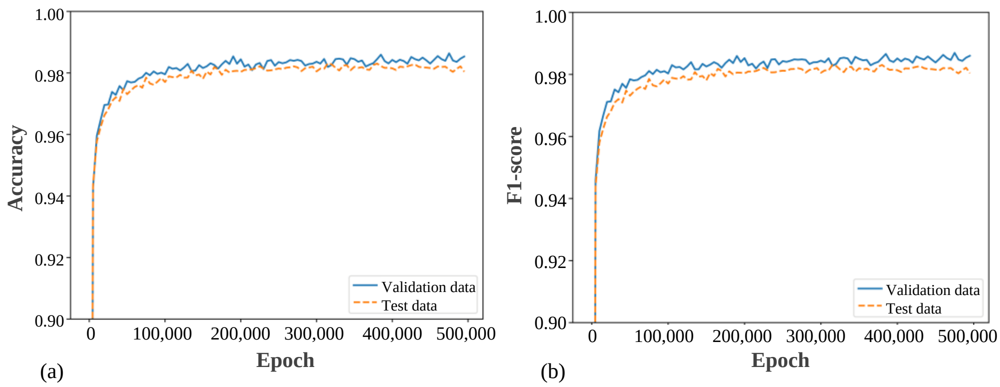

Fig. 7. Prediction performance of LSTM RNN model for each epoch: (a) Accuracy curve; and (b) F1-score curve

Since the resampled dataset used to train, validate, and test the prediction model is a synthesized dataset, the prediction models need to be evaluated against the original dataset. The prediction models may be overfitted to the resampled dataset and may not properly predict on the original dataset. We therefore additionally evaluated the prediction models with the original dataset (Table 4 and 5). The accuracy and F1-score of LSTM RNN model (0.986 and 0.812, respectively) had higher values compared to the FNN and SVM models. The accuracy of the FNN model and SVM model with RBF for the original dataset (0.862, and 0.953, respectively) was not significantly different from that of the test dataset. However, the F1-score of the FNN model and SVM models with RBF, linear and polynomial kernels for the original dataset showed the largely lower values (0.270, 0.557, 0.263 and 0.062, respectively) than that of the test dataset. The LSTM RNN model notably outperformed FNN and SVM models for original dataset.

To sum up, the experiment results showed that the accuracy and F1-score of the proposed LSTM RNN were higher than those of FNN and SVM models for both test and original dataset. The proposed LSTM RNN model has relatively better prediction performance than FNN and SVM models. These results indicate that the LSTM RNN model is a reliable technique to predict the business failure of construction contractors.

Table 4. Comparison of the prediction performance of LSTM RNN, FNN, and SVM models

Actual	Predicted (test dataset)		Predicted (original dataset)	
	Class0	Class1	Class0	Class1
		LSTM RNN model		
Class0	3,938	73	1,317	20
Class1	72	3,939	1	41
		FNN model		
Class0	3,337	674	1,149	188
Class1	488	3,523	9	33
		SVM model (RBF kernel)		
Class0	3,740	271	1,209	128
Class1	80	3931	1	41
		SVM model (linear kernel)		
Class0	2,988	1023	1,240	97
Class1	1,944	2,067	21	21
		SVM model (polynomial kernel)		
Class0	2,895	1,116	66	1,271
Class1	135	3,876	0	42

Table 5. Comparison of the prediction performance for LSTM RNN, FNN, and SVM models

Models	Test dataset		Original dataset	
	Accuracy	F1-score	Accuracy	F1-score
LSTM RNN	0.982	0.982	0.986	0.812

FNN	0.855	0.858	0.862	0.270
SVM				
RBF kernel	0.956	0.957	0.953	0.557
Linear kernel	0.630	0.582	0.914	0.263
Polynomial kernel	0.844	0.861	0.078	0.062

Discussion

The business failure of construction contractors may cause not only substantial losses to the economy but also another business failure to various stakeholders (Tserng et al. 2011, Horta and Camanho 2013, He0 and Yang 2014). Therefore, predicting the potential business failure of construction contractors has always been an important issue for both researchers and practitioners. To develop a robust model to enable the business failure prediction of construction contractors, this study applied LSTM RNN which has a time step of three. LSTM RNN feeds the output of the previous time step to the input of the next step. LSTM RNN has achieved great success in various applications on sequential data because of the gate and memory mechanism. However, it is usually considered 'black box' model whose internal structure and learned parameters are not interpretable. It is because LSTM RNN blindly blends the information of all variables into the hidden states and memory cells for subsequent prediction. Accordingly, it is intractable to distinguish the contribution of individual variables by looking into hidden states. Despite these limitations, LSTM RNN is powerful and effective for processing sequential data. To the best of our knowledge, this study is the first attempt to apply the LSTM RNN to the research on business failure prediction. On the basis of the results of this study, the use of LSTM RNN model has been demonstrated as a feasible approach to the development of models for business failure prediction.

Although this study proves that the proposed LSTM RNN model is a reliable technique to help predict the business failure of construction contractors, there are several issues to be improved for a practical application. First, the proposed LSTM RNN model in this study was designed to predict the business failure of construction contractors within one year. Understanding the possibility of business failure within one year is meaningful from a short-term perspective. However, many construction projects require a long period of construction time exceeding one year. Therefore, it would be necessary to develop the LSTM RNN model for business failure prediction covering a relatively long construction project period. Second, the proposed LSTM RNN model used only financial ratios as input variables. However, many factors affect the performance of the company's business failure. Macroeconomic factors are important to predict business failure since construction contractors are highly susceptible to macroeconomic effect (Arditi et al. 2000, Sang et al. 2013). The level of business riskiness may also vary due to the managerial factors. Thus, other variables such as the managerial and economic related factors need to be quantified and included in the model. Lastly, the proposed LSTM RNN model used twelve input variables with two hidden layers and 160 neurons for each hidden layer. It is important to choose a group of variables with more prediction information since reducing the number of redundant features reduces the running time of learning algorithm (Tasi 2009). It is also critical to select the optimal parameter of a neural network for the selected group of variables. However, theses require a lot of trial-and-error runs. Therefore, it would be necessary to develop a new approach to optimize the group of variables selection and parameter setting for an LSTM RNN model for business failure prediction.

Conclusions

This study proposed the business failure prediction model of construction contractors using LSTM RNN. Twelve financial ratios of the listed U.S. construction contractors from 1980 to 2016 were used to develop a model which can predict business failures within one year. To overcome the imbalanced classification problem, SMOTE+Tomek links was employed. A feedforward neural network (FNN) and a support vector machine (SVM) were used as a benchmark to evaluate the performance of the proposed LSTM RNN model. This study compared the accuracy and F1-score of three different prediction models for the optimized parameters selected from a grid search with five-fold cross-validation.

The empirical results showed that the proposed LSTM RNN model achieved the best prediction performance in both accuracy and F1-score compared to the FNN and SVM models. Moreover, the proposed LSTM RNN model outperformed the FNN and SVM models not only for the test dataset but also the original dataset. Therefore, the proposed LSTM RNN model is proven to be a promising alternative to assist managers, investors, auditors, and government entities in predicting business failure of construction contractors, and the approach may also be used as a reference for the other industry.

Nevertheless, there is still room for improvement for predicting performance and applicability of the LSTM RNN model for business failure prediction. The LSTM RNN model needs to cover a relatively long project period in the construction industry including managerial and economic variables. It also remains an interesting topic for further study to derive optimization producers to select proper group of variables and the corresponding parameter values.

Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). "Predicting business failure of construction contractors using long short-term memory recurrent neural network." Journal of Construction Engineering and Management, **doi.org/10.1061/(ASCE)CO.1943-7862.0001709**

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Science, ICT & Future Planning (No. 2015R1A5A1037548).

References

Arditi, D., Koksal A. and Kale, S. (2000). "Business failures in the construction industry." *Engineering Construction and Architecture Management*, 7(2), 120–132.

Abidali, A.F. and Harris, F.C. (1995). "A methodology for predicting company failure in the construction industry." *Construction Management and Economics*, 13(3), 189–96.

Adeleye, T., Huang, M., Huang, Z. and Sun, L. (2013). "Predicting loss for large construction companies." *Journal of construction engineering and management*, 139(9), 1224–1236.

Al-Sobiei, O.S., Arditi, D. and Polat, G. (2005). "Predicting the risk of contractor default in Saudi Arabia utilizing artificial neural network (ANN) and genetic algorithm (GA) techniques." *Construction Management and Economics*, 23(4), 423–430.

Bal, J., Cheung, Y. and Wu, H.C. (2013). "Entropy for business failure prediction: An improved prediction model for the construction industry." *Hindawi Publishing Corporation Advances in Decision Science*s, 2013, 1–14.

Bengio, Y., Simard, P. and Frasconi, P. (1994). "Learning long-term dependencies with gradient descent is difficult." *IEEE Transactions on Neural Network*, 5, 157–166.

Chawla, N.V., Bowyer, K.W., Hall L.O. and Kegelmeyer, W.P. (2002). "Synthetic Minority Over-sampling Technique." *Journal of Artificial Intelligence Research*, 16, 321–357.

Chen, J.H. (2012). "Developing SFNN models to predict financial distress of construction companies." *Expert Systems with Applications*, 39(1), 823–827.

Cheng, M.Y. and Hoang, N.D. (2015). "Evaluating contractor financial status using a hybrid fuzzy instance-based classifier: case study in the construction industry." *IEEE Transactions on Engineering Management*, 62(2), 184–192.

Cheng, M.Y., Hoang, N.D., Limanto, L. and Wu, Y.W. (2014). "A noble hybrid intelligent approach for contractor default status prediction." *Knowledge-based Systems*, 71, 314–321.

Grice, J. S. and M. T. Dugan (2001). "The Limitations of Bankruptcy Prediction Models: Some Cautions for the Researcher." *Review of Quantitative Finance and Accounting*, 17(2), 151–166.

Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). “Predicting business failure of construction contractors using long short-term memory recurrent neural network.” Journal of Construction Engineering and Management, **doi.org/10.1061/(ASCE)CO.1943-7862.0001709**

Edum-Fotwe, F., Price, A. and Thorpe, A. (1996). “A review of financial ratio tools for predicting contractor insolvency.” *Construction Management and Economics*, 14(3), 189–98.

Heo, J. and Yang, J.Y. (2014). “AdaBoost based bankruptcy forecasting of Korean construction companies.” *Applied Soft Computing*, 24, 494–499.

Hochreiter, S. and Schmidhuber, J (1997). “Long short-term memory.” *Neural Computation*, 9(8), 1735–1780.

Horta, I.M. and Camano, A.S. (2013). “Company failure prediction in the construction industry.” Expert Systems with Applications, 40(16), 6253–6257.

Hwee, N. G. and Tiong, R. L. K. (2002). “Model on cash flow forecasting and risk analysis for contracting firms.” *International Journal of Project Management*, 20(5), 351–63.

Kangari, R., Farid, F. and Elgharib, H. (1992). “Financial performance analysis for construction industry.” *Journal of Construction Engineering and Management, 118(2), 349–361.*

Keerthi, S. and Lin, C.J. (2003). “Asymptotic behaviors of support vector machines with Gaussian kernel.” *Neural Computation*, 15(7), 1667–1689.

Khosrowshahi, F. and Kaka, A.P. (2007). “A decision support model for construction.” *Computer-Aided Civil and Infrastructure Engineering*, 22(7), 527–39.

Lam, K.C., Palaneeswaran, E. and Yu, C. (2009). “A support vector machine model for contractor prequalification.” *Automation in Construction*, 18(3), 321–329.

Mason, R. J. and Harris, F. C. (1979). “Predicting company failure in the construction industry.” in *Proceedings Institution of Civil Engineers*, London, 301–307.

Niemann, M., Schmidt, J.H. and Neukirchen, M. (2008). “Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity.” *Journal of Banking and Finance*, 32(3), 434–446.

Russell, J. S. and Zhai, H. (1996). “Predicting contractor failure using stochastic dynamics of economic and financial variables.” *Journal of Construction Engineering and Management*, 122(2), 183–191.

Samarasinghe, S. (2006). *Neural Networks for Applied Sciences and Engineering*, Taylor and Francis, 2006.

Sang, J., Ham, N.H., Kim, J.H., and Kim, J.J. (2014). “Impacts of macroeconomic fluctuations on insolvency: case of Korean construction companies.” *Journal of Management in Engineering*, 20(5), 10.1061/(ASCE)ME.1943-5479.0000235

Tomek, I. (1976). “Two modifications of CNN.” *IEEE Transactions on Systems, Man, and Cybernetics*, 6 (11), 769–772

Jang, Y., Jeong, I., Cho, Y., and Ahn, Y. (2019). "Predicting business failure of construction contractors using long short-term memory recurrent neural network." Journal of Construction Engineering and Management, **doi.org/10.1061/(ASCE)CO.1943-7862.0001709**

Tsai, C.F. (2009). "Feature selection in bankruptcy prediction." Knowledge-based Systems, 22 (2009), 120-127

Tsai, L. K., Tserng, H., Liao, H. H., Chen, P. C. and Wang, W. (2012). "Integration of accounting-based and option-based models to predict construction contractor default." *Journal of Marine Science and Technology,* 20(5), 479–484.

Tserng, H.P., Lin, G.F., Tsai, L.K. and Chen, P.C. (2011). "An enforced support vector machine model for construction contractor default prediction." *Automation in Construction*, 20(2011), 1242-1249.

Tserng, H.P. Liao, H.H., Jaselskis, E.J., Tsai, L.K. and Chen, P.C. (2012), "Predicting construction contractor default with barrier option model option model." *Journal of construction engineering and management*, 138(5), 621-630.

Tserng, H.P., Chen, P.C., Huang, W.H., Lei, M.C. and Tran, Q.H. (2014), "Predicting of default probability for construction firms using the logit model." *Journal of Civil Engineering and Management*, 20(2), 247-255.

Tserng, H.P., Ngo, T.L., Chen, P.C. and Tran, L.Q. (2015). "A grey system theory-based default prediction model for construction firms." *Computer-Aided Civil and Infrastructure Engineering*, 30(2), 120-134.

Vapnik, V.N. *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.

Yeh, C. C., Chi, D. J. and Hsu, M. F. (2010). "A hybrid approach of DEA, rough set and support vector machines for business failure prediction." *Expert Systems with Applications*, 37(2), 1535–1541.