

## Article

# Artificial Intelligence in Corporate Sustainability: Using LSTM and GRU for Going Concern Prediction

Der-Jang Chi <sup>1</sup> and Chien-Chou Chu <sup>2,\*</sup><sup>1</sup> Department of Accounting, Chinese Culture University, Taipei City 11114, Taiwan; qdz@ulive.pccu.edu.tw<sup>2</sup> Graduate Institute of International Business Administration, Chinese Culture University, Taipei City 11114, Taiwan

\* Correspondence: A6001001@ulive.pccu.edu.tw

**Abstract:** “Going concern” is a professional term in the domain of accounting and auditing. The issuance of appropriate audit opinions by certified public accountants (CPAs) and auditors is critical to companies as a going concern, as misjudgment and/or failure to identify the probability of bankruptcy can cause heavy losses to stakeholders and affect corporate sustainability. In the era of artificial intelligence (AI), deep learning algorithms are widely used by practitioners, and academic research is also gradually embarking on projects in various domains. However, the use of deep learning algorithms in the prediction of going concern remains limited. In contrast to those in the literature, this study uses long short-term memory (LSTM) and gated recurrent unit (GRU) for learning and training, in order to construct effective and highly accurate going-concern prediction models. The sample pool consists of the Taiwan Stock Exchange Corporation (TWSE) and the Taipei Exchange (TPEX) listed companies in 2004–2019, including 86 companies with going concern doubt and 172 companies without going concern doubt. In other words, 258 companies in total are sampled. There are 20 research variables, comprising 16 financial variables and 4 non-financial variables. The results are based on performance indicators such as accuracy, precision, recall/sensitivity, specificity, F1-scores, and Type I and Type II error rates, and both the LSTM and GRU models perform well. As far as accuracy is concerned, the LSTM model reports 96.15% accuracy while GRU shows 94.23% accuracy.

**Keywords:** going concern prediction; artificial intelligence (AI); corporate sustainability; deep learning algorithm; long short-term memory (LSTM); gated recurrent unit (GRU)



**Citation:** Chi, D.-J.; Chu, C.-C. Artificial Intelligence in Corporate Sustainability: Using LSTM and GRU for Going Concern Prediction. *Sustainability* **2021**, *13*, 11631. <https://doi.org/10.3390/su132111631>

Academic Editor: Fabio Carlucci

Received: 13 September 2021

Accepted: 18 October 2021

Published: 21 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

After AlphaGo defeated numerous top human Go players in 2014, the public and the media have become highly interested in and attentive to AI given the continuous upgrade of robots and the success of driverless car tests on highways. The technological breakthrough of AI over the last few years came from the gradual maturity and readiness of both software and hardware such as networking, big data, cloud computing, algorithms, and semiconductor chips. The requirement to process and analyze the large volume of data generated from each applied field, which is critical technology and competence for corporate operations, further pushes the development of AI. The fundamental applications of AI include deep learning, voice to texts, Natural Language Processing (NLP), Optical Character Recognition (OCR), and voice recognition, and smart technologies constructed with algorithms are everywhere these days. With further development, AI systems can directly interpret business activities, obtain and analyze financial information, manage risks, and issue warnings. In the future, AI will surely be combined with fundamental technology and commercial intelligence to create business value.

The cognitive insight provided by deep learning is different from that provided by traditional analysis and typically comes with more dense data, in bigger volumes, and with greater details. The learning and training with a certain dataset usually enhance the

models, and the speed is also amazing, often in just seconds. Deep learning can make predictions with new data or classify items. This capability improves over time. Deep learning is able to perform outstanding functions such as voice and image recognition. It seeks to imitate how human brains function in order to identify different patterns. Deep learning also provides new data, makes better analysis, and builds models [1]. The use of deep learning algorithms to process data characteristics, change and adjust parameters, and provide optimal prediction models with given datasets and prediction targets will help decision-makers to reach wiser, faster, and more accurate decisions.

“Going concern” is a professional term in the domain of accounting and auditing. It is about the assessment by certified public accountants (CPAs) and auditors based on financial statements and relevant risks to determine whether a company has a going concern doubt [2]. It has a profound impact on whether a company can continue as a going concern. If certified public accountants (CPAs) and auditors do not issue audit reports or audit opinions about going concern doubt before companies experience financial distress, then this will greatly hurt the stakeholders of the companies [2–4]. There is an extensive literature on the importance of external auditors in the financial market [5–7]. Going concern assumes a company will continue to operate in its current scale and status in the foreseeable future at least 12 months after the balance sheet date. Researchers indicated that one of the main causes of the global financial crisis in 2008–2009 starting in the U.S. and subsequently engulfing the world was the failure of CPAs and auditors to fulfill their audit responsibility or even their issuance of false audit reports [8–10]. In other words, it was due to erroneous judgments and opinions from CPAs regarding going concern and sustainability of companies [10].

Auditing regulators worldwide demand that CPAs and auditors discuss key audit matters (KAMs) and provide more information on audit reports [10,11]. KAMs are determined and carried out by CPAs and auditors according to industry characteristics, company specifics, and respective risks. After the financial crisis in 2008–2009, there was an increasing call for reforms in audit reports around the world. A new format of audit reports was created via a series of discussions by the European Union and IAASB (International Auditing and Assurance Standards Board). The biggest change for the new audit reports was the requirement for CPAs to disclose KAMs; i.e., the risks encountered during the audit process, material judgment, or significant events during the audit period, in a language investors can understand.

Taiwan also started to adopt the new format of audit reports in the fourth quarter of 2016. In addition to KAMs [12] and before the communication of KAMs in audit reports, SAS No. 58 of Taiwan (2018) requires that CPAs and auditors implement tests of control, substantive analytical procedures, and tests of details of balances during the audit process where necessary [13]. Large international accounting firms such as Deloitte, KPMG, PwC, and EY have robust audit systems and regulations to assist, support, and supervise CPAs and auditors so as to reduce the risk of audit failures. For example, computer-aided audit techniques and walk-throughs are conducted during the audit process to enhance the quality and accuracy of audit reports and audit opinions.

CPAs may collude with senior management and issue audit reports and audit opinions that differ from the facts, which can cause great damage to corporate stakeholders. After the Enron fraud in the U.S. in 2001, the U.S. Congress passed the Sarbanes-Oxley Act (SOX) in 2002. This legislation stipulates that internal control in relation to financial reports enhances the transparency of company operations and financial information and establishes relevant regulations to ensure the accuracy and reliability of financial information. There is now a heavier burden on the independence and legal liability of CPAs. CPAs should act with professionalism and independence in the offering of audit services and issue appropriate opinions regarding whether there is going concern doubt. However, it concerns the issuance of appropriate audit opinions on customers with going concern under the dual consideration of economic dependence and reputation protection and at the risk of being replaced [2,14].

According to SAS No. 57 of Taiwan [15], when an event casts material doubt or there is significant uncertainty regarding the audited entity as a going concern, then CPAs should issue reports by following relevant audit standards. CPAs should refer to the audit evidence obtained and reach conclusions about the appropriateness of the accounting basis adopted by management. They should also include the event that may cast material doubt or any circumstance that may cause great uncertainty regarding the audited unit's going concern capability. Risk assessments are additionally required by the IFRSs.

Based on the aforesaid financial statements and risks, CPAs issue audit reports and audit opinions, such as (1) unqualified opinion; (2) qualified opinion; (3) disclaimer opinion; and (4) adverse opinion [16]. As per SAS No. 57 of Taiwan "Audit Reports on Financial Statements" [15] and No. 61 "going concern" [16], CPAs should issue qualified opinions or adverse opinions depending on the materiality of the impact after assessment on the reasonableness of going concern, if the audited unit's financial statements do not make appropriate disclosure. In addition to going concern doubt, the reasons for CPAs to issue qualified opinions or adverse opinions include limitation of the audit scope and differing opinions with the audited management's choice of accounting policies or disclosure in financial statements according to SAS No. 57 of Taiwan [15]. SAS No. 61 of Taiwan [16] stipulates that the basis of financial statement preparation is often based on going concern assumptions. If the doubt is cleared after the assessment of the reasonableness of going concern assumptions, then CPAs may issue unqualified audit opinions in audit reports. If CPAs believe that the responding measures taken by the audited party are reasonable, but it is necessary to disclose relevant contents in financial reports, then qualified opinions or adverse opinions should be issued in audit reports. If the going concern doubt remains, but the audited financial statements have made appropriate disclosure, then CPAs should issue qualified opinions or adverse opinions in audit reports in accordance with the degrees of the impact. If CPAs are certain that the basis of going concern assumptions with which the audited financial statements are prepared does not reflect the reality and the impact is extremely significant, then adverse opinions should be issued in audit reports.

The auditing by CPAs of financial statements serves as one type of external supervisory mechanism. In contrast with unqualified opinions, qualified opinions mean CPAs have doubts about certain contents of the financial reports or think there are uncertainties. In Taiwan, the issuance of audit reports and audit opinions issued by CPAs affects the trading of shares of TWSE/TPEX listed companies as follows: (1) unqualified opinions—no effects; (2) qualified opinions—change of trading methods; (3) disclaimer opinions—suspension of trading; (4) adverse opinions—suspension of trading. Additionally unique to Taiwan is the double sign system; i.e., signing for reviews or audits by two CPAs for accounts of the same TWSE/TPEX listed company in order to enhance the accuracy of audit reports and audit opinions and to show a better reflection of financial status. The purpose is to protect corporate stakeholders, financial report users, and potential investors. In SAS No. 62 of Taiwan "Communication with those Charged with Governance" [17], CPAs and auditors must communicate with the company's governance unit (such as the Audit Committee) so that the governance unit's members can better understand the audit process (such as KAMs) of CPAs and auditors.

While the government has set forth stringent regulations on the audit process by CPAs and auditors, CPAs assume legal liabilities for the issuance of audit reports and audit opinions. There is naturally always a slew of events where stakeholders, capital markets, and national economies suffer losses due to inaccurate audit reports and audit opinions issued by CPAs. CPAs themselves are also subject to penalties. Therefore, it is essential and imperative to construct an effective going concern prediction model to assist the audit work by CPAs and auditors and enable the issuance of audit reports and audit opinions to be better reflective of the reality.

To welcome the advent of the era of artificial intelligence (AI), the purpose of this study is to use efficient deep learning algorithms to construct going concern prediction models. In contrast to the existing literature, this study uses long short-term memory (LSTM)

and gated recurrent unit (GRU), two efficient deep learning algorithms, for learning and training in order to construct effective and highly accurate going concern prediction models. The research variables include both financial variables and non-financial variables. Data spanning over 16 years are sourced, in order to assist CPAs and auditors in the issuance of more accurate audit reports.

The structure of this study is described in order as follows: Section 1. Introduction, Section 2. Related Works, Section 3. Materials and Methods, Section 4. Results, Section 5. Discussion, and Section 6. Conclusions.

## 2. Related Works

The primary reason for audit failures is the error in the reasonableness of going concern assumption made by auditors, which is relevant to the professional judgment of auditors [2,18]. CPAs and auditors may be under the pressure of time and rewards, which may affect the judgment and decision by CPAs regarding going concern opinions [19].

When auditing financial statements, CPAs assess whether there is great uncertainty with the company's going concern. If any material uncertainty is confirmed, then CPAs will take into account the liquidity disclosed in the annual report for continued operations. Compared to companies without going concern doubt, companies with going concern doubt have worse financial structures, poorer liquidity, and lower efficiency and profitability [20].

Most of the past studies use traditional statistical methods such as factor analysis, regression analysis, discrimination analysis, and cluster analysis for going concern decisions. However, there are significant limitations and deficiencies in the research process and judgment and hence a likelihood of errors [2,10,18]. Some recent studies use data mining and machine learning techniques to boost the accuracy of going concern decisions. These techniques include artificial neural network (ANN), decision tree (DT), support vector machine (SVM), and Bayesian network (BN) [2,18,21–30]. As AI is gradually finding its way into research by practitioners and academia, deep learning algorithms and techniques are being used for going concern prediction [10]. Jan [10] samples 352 TWSE/TPEX listed companies in Taiwan in 2002–2019 and deploys deep neural network (DNN), recurrent neural network (RNN), and classification and regression tree (CART) to construct going concern prediction models. The most accurate is the CART-RNN model with a test dataset accuracy of 95.28%, Type I error rate and Type II error rate are 2.83% and 1.89% respectively. Two other relevant papers were written by Jan [31,32]: one discusses fraud in financial statements, two deep learning algorithms, recurrent neural network (RNN), and long short-term memory (LSTM). The research results show that the accuracy of the LSTM model is as high as 94.88%, Type I error rate and Type II error rate are both 2.56%. The other paper uses the chi-squared automatic interaction detector (CHAID), deep neural network (DNN), and convolutional neural network (CNN) to predict financial distress. According to the results, with the important variables selected by CHAID and modeling by CNN, the CHAID-CNN model has the highest financial distress prediction accuracy rate of 94.23%, Type I error rate and Type II error rate are 0.96% and 4.81% respectively. Based on the research results of Jan's three papers application of machine learning and deep learning algorithms to finance, accounting, and auditing topics, it can be assured that the prediction accuracy of using machine learning and deep learning algorithms is relatively high.

A study by Hamal and Senvar [33] discusses fraud in financial accounting. Six machine learning algorithms and the logistic regression are used, and their results show that the Random Forest classifier always performs the best or the second best among the seven classifiers in terms of all the performance metrics. The overall accuracy of the Random Forest without feature selection-oversampling model is the highest at 93.74%. Goo et al. [18] apply the least absolute shrinkage and selection operator (LASSO) and three machine learning algorithms to construct prediction going concern models, such as neural network (NN), classification and regression tree (CART), and support vector machine (SVM). According to their results, the prediction accuracy of the LASSO-NN model is

88.96%, the prediction accuracy of the LASSO-CART model is 88.75%, and the prediction accuracy of the LASSO-SVM model is 89.79%. Yeh et al. [29] use a hybrid random forest (RF) and rough set theory (RST) approach to predict going concern; their results show that the average accuracy is 96.10%. Chen and Lee [28] also use a hybrid decision tree CART, decision tree CHAID, artificial neural network (ANN), and stepwise regression (SR) approach to predict going concern; their results show that the CART-ANN model has the highest prediction accuracy (96.77%) for identifying going concern doubts and also has the highest overall accuracy (94.66%).

It is worth mentioning that many studies [2,10,18,24–30] clearly state that both machine learning and deep learning algorithms are more rigorous and accurate than traditional statistical methods. In other words, compared with traditional statistical methods, machine learning and deep learning algorithms have higher accuracy and lower error rates.

In summary, the prediction accuracy of using machine learning and deep learning algorithms is quite high. Both can be as high as 90%, but deep learning algorithms seem to be more stable and fast. Moreover, the academic presentation tools of artificial intelligence (AI) are deep learning algorithms [10,31,32].

### 3. Materials and Methods

This study samples the Taiwan Stock Exchange Corporation (TWSE) and the Taipei Exchange (TPEX) listed companies in Taiwan in 2004–2019. The sample pool consists of 258 companies: 86 companies with going concern doubt and 172 companies without going concern doubt. Two powerful deep learning algorithms, long short-term memory (LSTM) and gated recurrent units (GRU), are used to construct going concern prediction models.

#### 3.1. Research Design

To achieve the research objectives, this study designs a three-step research process, as shown in Figure 1. The first step is data acquisition and preprocessing. Financial variables and non-financial variables data are sourced from Taiwan Economic Journal (TEJ) on the companies with going concern doubt and without going concern doubt. All the data are randomly distributed into the training dataset, the validation dataset, or the test dataset. The second step is modeling, by inputting the data from the training dataset and the validation dataset into the LSTM model and the GRU model for deep learning. The third step is evaluation, by testing the data in the test dataset to assess the model performance and present the effectiveness of going concern prediction models.

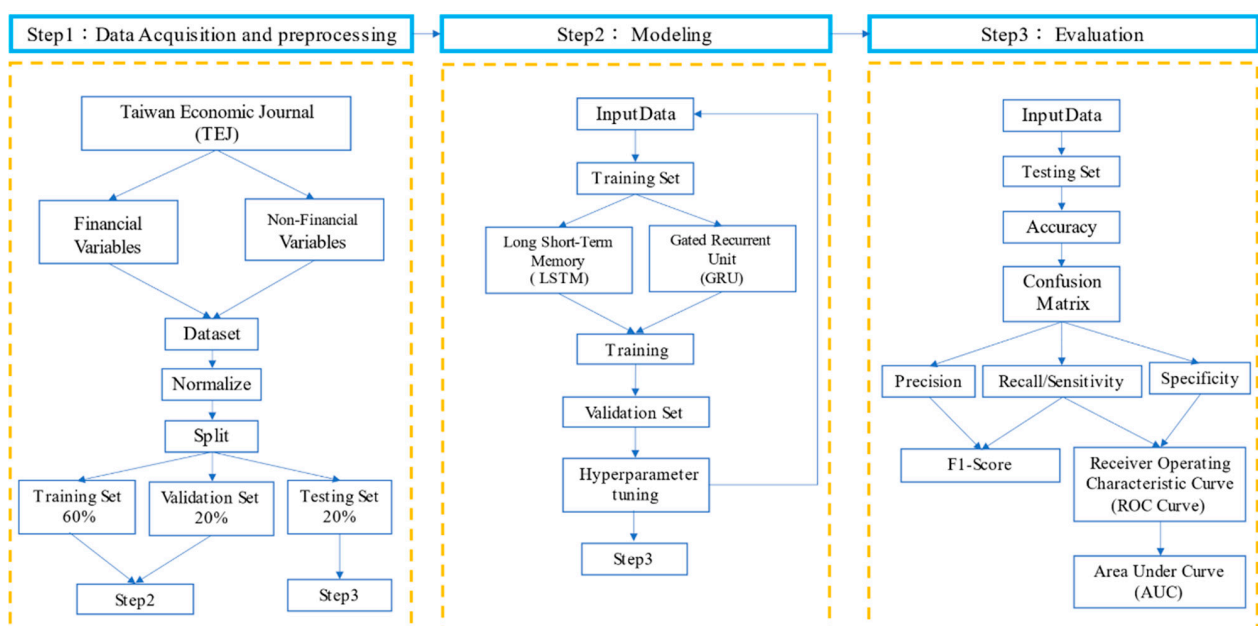


Figure 1. Research design.

### 3.2. Samples and Datasets

This study sources financial data and non-financial data on the 258 sampled companies from Taiwan Economic Journal (TEJ). All the data are randomly selected for the training dataset (60%) in learning and modeling, the validation dataset (20%) to assist the modeling, and the test dataset (20%) for testing of the model performance. The distribution of the sampled companies by industry is summarized in Table 1.

**Table 1.** Sample distribution.

Industry	Going Concern Doubt	No Going Concern Doubt	Total
Semiconductor	6	12	18
Optoelectronics	13	26	39
Electric machinery	3	6	9
Electronic components	11	22	33
Electric appliances and cables	2	4	6
Computers and peripherals	4	8	12
Communication and networking	1	2	3
Information services	2	4	6
Other electronics	5	10	15
Biotech and medicare	3	6	9
Steel	4	8	12
Building materials and construction	5	10	15
Textiles	5	10	15
Food	1	2	3
Cultural and creative industry	4	8	12
Trade and department stores	2	4	6
Shipping	2	4	6
Oil, electricity, and gas	1	2	3
Tourism	4	8	12
Others	8	16	24
Total	86	172	258

### 3.3. Variables

The dependent variable is categorized according to audit opinions expressing going concern doubt. It is a dummy variable, with 1 indicating going concern doubt and 0 if not. Independent variables (research variables) are 20 variables frequently used to measure going concern. They include 16 financial variables and 4 non-financial variables. The research variables are summarized in Table 2.

### 3.4. Methods

This study uses two efficient deep learning algorithms, long short-term memory (LSTM) and gated recurrent unit (GRU), for modeling.

#### 3.4.1. Long Short-Term Memory

Long short-term memory (LSTM) is a deep learning model derived from a recurrent neural network (RNN). It was proposed in 1997 by Hochreiter and Schmidhuber [34]. It is designed for the processing of serial data, but can effectively handle non-serial data. Given its unique structures, LSTM is also suitable for processing and predicting key events with long intervals and delays of time series. LSTM generalizes the problem domain well. This is important because some tasks can no longer be resolved with the existing recurrent neural networks. It is also a great advantage compared to RNNs. The functioning architecture is depicted in Figure 2. The model is designed to resolve the problem of discontinued learning due to the two major RNN flaws—i.e., inability to retain long-term memory and vanishing of gradient that prevents neural network weights at shallower levels from updating during

backpropagation. To address this, LSTM adds a gate control mechanism and memory cells on the basis of RNNs.

Table 2. Research variables.

Code	Variable	Variable Definition or Equation (The Year before the Year of Going Concern Doubt)
X01	Non-current liabilities/total assets	Non-current liabilities/total assets
X02	Working capital/total assets	Working capital/total assets
X03	Retained earnings/total assets	Retained earnings/total assets
X04	Fixed assets/total assets	Fixed assets/total assets
X05	Liability ratio	Total liabilities/total assets
X06	Liability to equity ratio	Total liabilities/shareholders' equity
X07	Current liabilities/revenues	Current liabilities/revenues
X08	Working capital/revenues	Working capital/revenues
X09	Cash flows from operating activities/revenues	Cash flows from operating activities/revenues
X10	Operating cash flow ratio	Cash flows from operating activities/current liabilities
X11	Cash flow adequacy ratio	Operating cash flows over the past five years/(capital expenditures + increase in inventory + cash dividends over the past five years)
X12	Inventory turnover	Cost of goods sold/average inventory
X13	Current ratio	Current assets/current liabilities
X14	Return on equity	Net income/average shareholders' equity
X15	Growth in asset	(Total assets at the end of the current year—total assets at the end of the previous year/total assets at the end of the previous year
X16	Growth in revenues	(Revenues during the current year—revenues during the previous year)/revenues during the previous year
X17	No. of independent directors	No. of independent directors
X18	Percentage of shareholdings by directors and supervisors	No. of shares held by directors and supervisors/No. of ordinary shares outstanding at the end of the period
X19	Percentage of shares pledged by directors and supervisors	No. of shares pledged by directors and supervisors/No. of shares held by directors and supervisors
X20	CEO/Chairperson duality	1 if Chairperson is also CEO and 0 if not

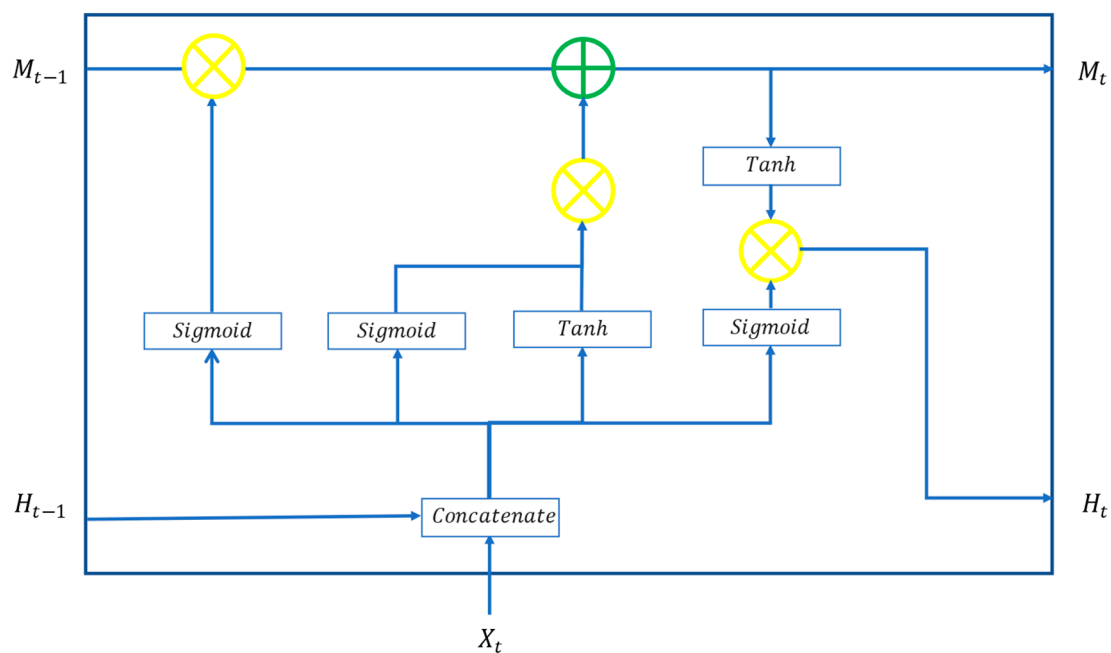


Figure 2. LSTM architecture.

The gate control mechanism consists of a forget gate ( $F_t$ ), an input gate ( $I_t$ ), and an output gate ( $O_t$ ). The forget gate ( $F_t$ ), as illustrated in Figure 3, serves to determine which information in memory cells should be forgotten with a Sigmoid activation function, based on the input results from the previous period ( $H_{t-1}$ ) and new information inputted during this period ( $X_t$ ). This is expressed with Equation (1) where  $W_F$  is the weight of the forget gate and  $b_F$  is the bias of the forgotten door.

$$F_t = \text{Sigmoid}[W_F(X_t + H_{t-1}) + b_F] \tag{1}$$

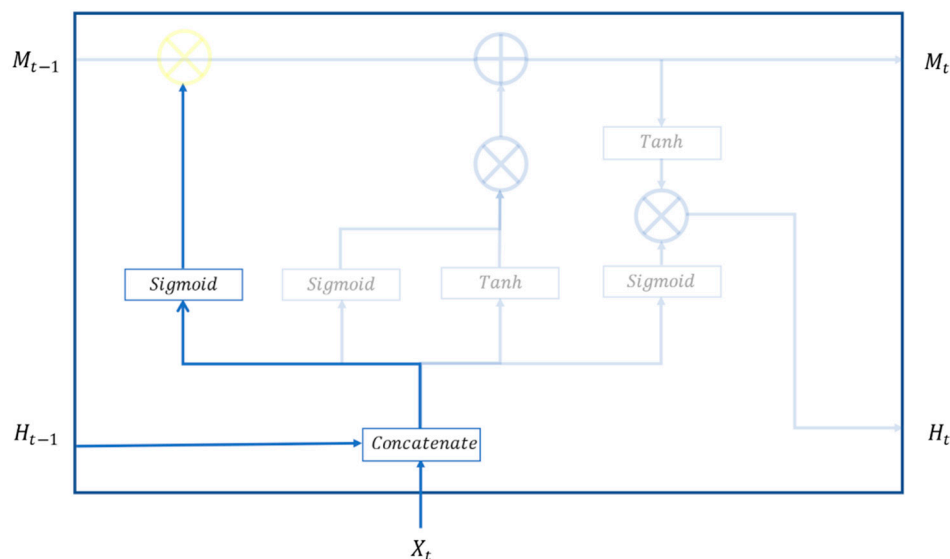


Figure 3. Forget gate.

The input gate determines which information is to be inputted into memory cells, as shown in Figure 4. The Sigmoid function is used to decide which outputs from the previous period and inputs during the current period are to be updated to the memory cells. The hyperbolic tangent function ( $\tanh$ ) generates the update parameters ( $\tilde{M}_t$ ) in the memory state, expressed with Equations (2) and (3).

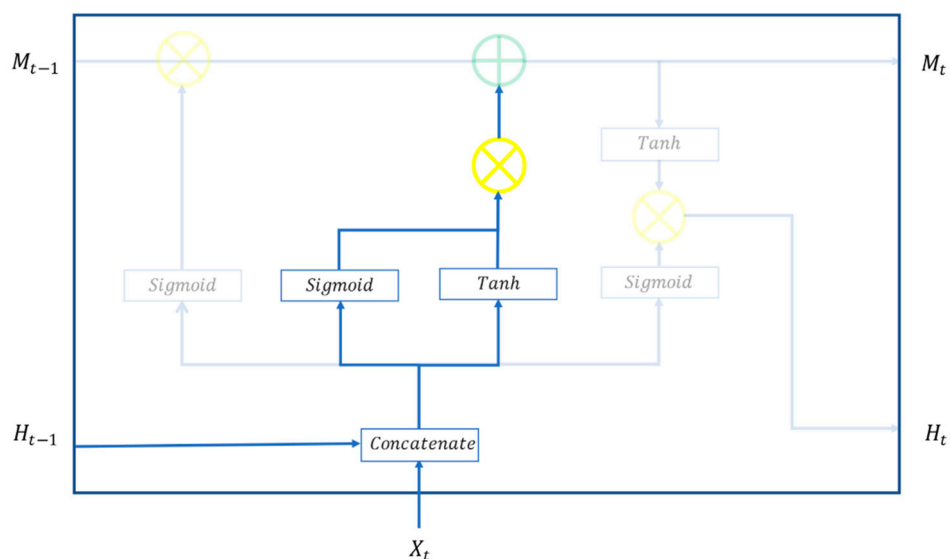


Figure 4. Input gate.



Here,  $W_I$  and  $W_M$  are the weight of the input gate and the weight of the update parameter, respectively;  $b_I$  and  $b_M$  are the weight of the input gate and the weight of the update parameter, respectively.

$$I_t = \text{Sigmoid}[W_I(X_t + H_{t-1}) + b_I] \tag{2}$$

$$\tilde{M}_t = \text{tanh}[W_M(X_t + H_{t-1}) + b_M] \tag{3}$$

Memory cells are the databank of long-term memory for the calculation of each input value, as shown in Figure 5. The forget gate determines which information from the previous memory cell ( $M_{t-1}$ ) is to be forgotten, and the input gate calculates the current memory cell ( $M_t$ ) by updating the information, as expressed in Equation (4). The output gate determines, with the Sigmoid function, the output from the previous period and the new information for the current period to form into Equation (5). The final memory cell calculates the current output (i.e., the hidden state,  $H_t$ ) by multiplying individual elements with the hyperbola and the output gate to form into Equation (6). The process is outlined in Figure 6.

$$M_t = M_{t-1} \times F_t + \tilde{M}_t \times I_t \tag{4}$$

$$O_t = \text{Sigmoid}[W_O(X_t + H_{t-1}) + b_O] \tag{5}$$

$$H_t = \text{tanh}(M_t) + O_t \tag{6}$$

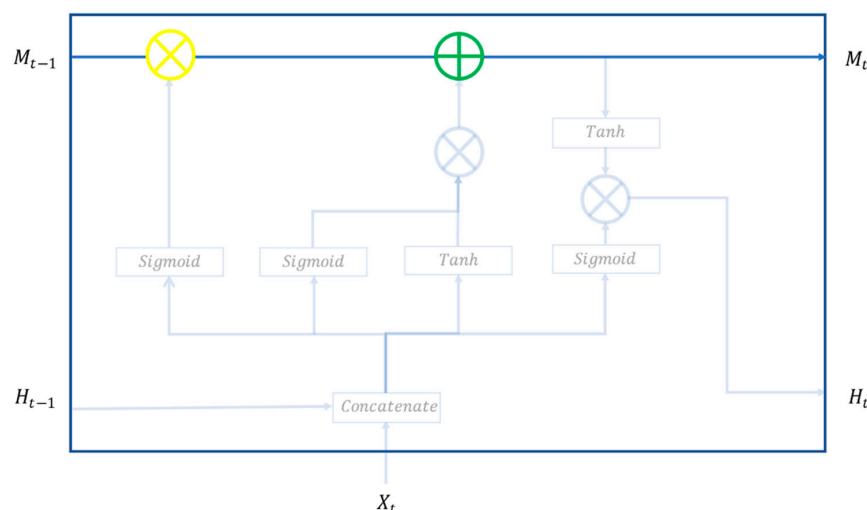


Figure 5. Memory cell.

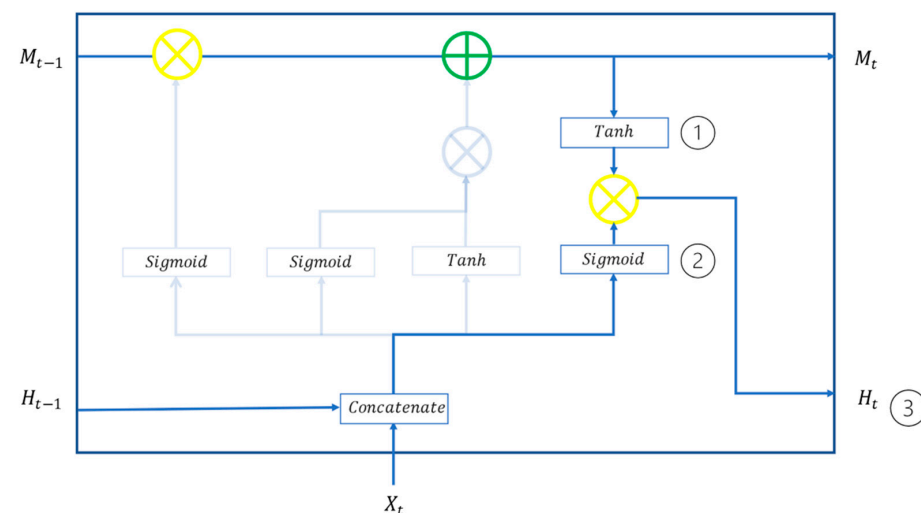


Figure 6. Output process.

### 3.4.2. Gated Recurrent Unit

Gated recurrent unit (GRU) was developed by Chung et al. [35]. Similar to LSTM, GRU also aims to address RNN's two major flaws—i.e., inability to retain long-term memory and vanishing of gradients. However, GRU does not have memory cells and only relies on the gate control mechanism to resolve these two RNN problems. It seeks to maintain equivalent effectiveness by significantly reducing parameters and computing. Therefore, quick computing is a great advantage of GRU, which has achieved successful results in serial data and time data. It is suitable for voice recognition, natural language processing, and machine translation. Just like LSTM, GRU performs well in long series problem domains. Its functioning process is illustrated in Figure 7. In contrast to LSTM, GRU only has two gates—i.e., the reset gate ( $R_t$ ) and the update gate ( $U_t$ ). The reset gate, as shown in Figure 6, determines with the Sigmoid function which information from the previous time steps ( $H_{t-1}$ ) is to be forgotten, which is similar to the function served by the forget gate in LSTM. The difference is that LSTM selects the to-be-forgotten information based on memory cells and GRU determines the to-be-forgotten information according to the past time steps (in the hidden state). The reset gate is expressed by Equation (7). The update gate, shown in Figure 8, serves as a function similar to the input gate of LSTM by controlling the ratio of the information for new input ( $X_t$ ) to the output value from the previous calculation. The input gate is expressed by Equation (8), and the calculation of the reset parameter ( $\tilde{H}_t$ ) in the current hidden state is described in Figure 9. The calculation is based on the hidden state ( $R_t \times H_{t-1}$ ) reset with the hyperbolic tangent function to form Equation (9). As shown in Equation (10) and Figure 10, the ratio of the reset parameter in the current hidden state to the previous hidden state is calculated with the reset gate as the outcome for the current iteration (Figure 11).

$$R_t = \text{Sigmoid}[W_R(X_t + H_{t-1}) + b_R] \quad (7)$$

$$U_t = \text{Sigmoid}[W_U(X_t + H_{t-1}) + b_U] \quad (8)$$

$$\tilde{H}_t = \text{tanh}[W_H(X_t + R_t \times H_{t-1}) + b_H] \quad (9)$$

$$H_t = (1 - U_t) \times H_{t-1} + U_t \times \tilde{H}_t \quad (10)$$

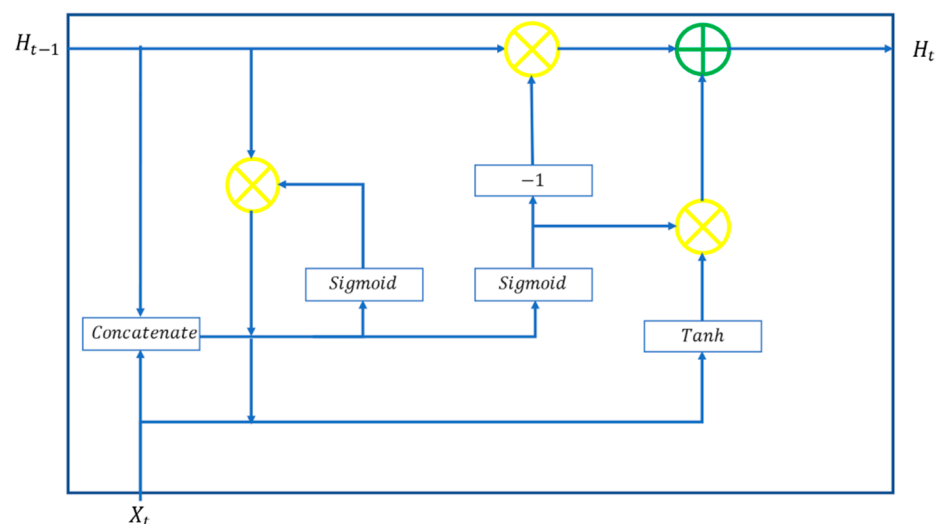


Figure 7. Gated recurrent unit network.

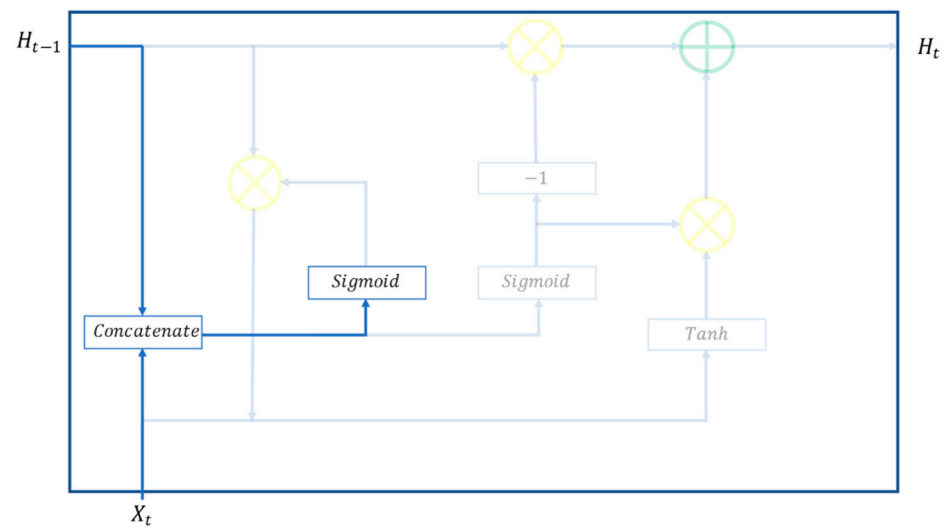


Figure 8. Reset gate.

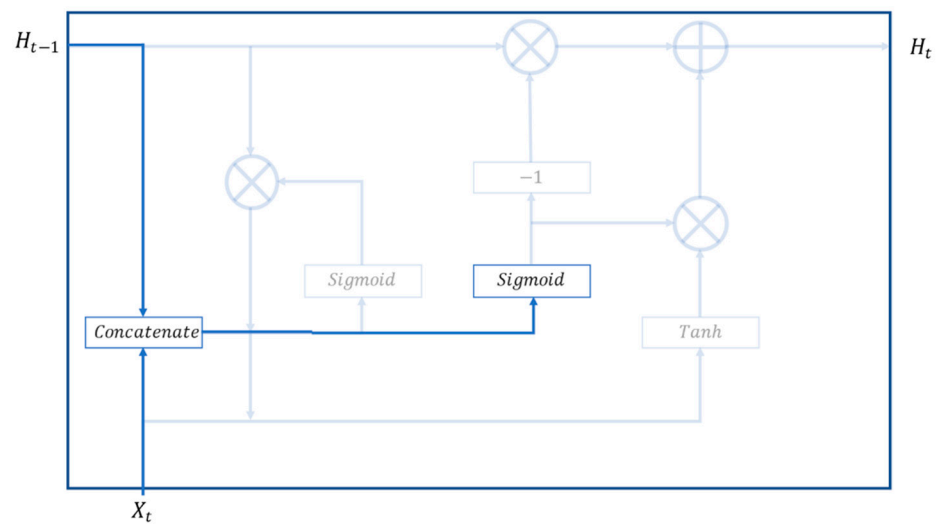


Figure 9. Update gate.

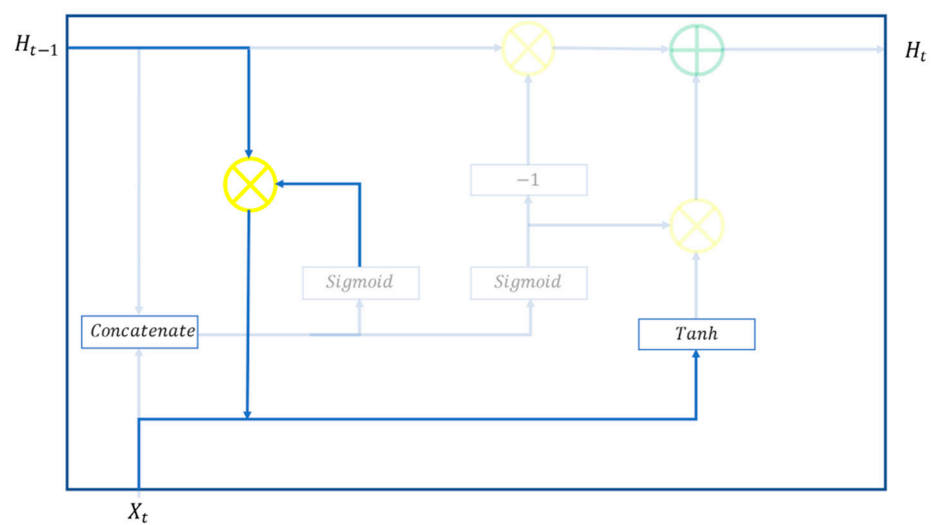


Figure 10. Calculation of update parameters in hidden state.

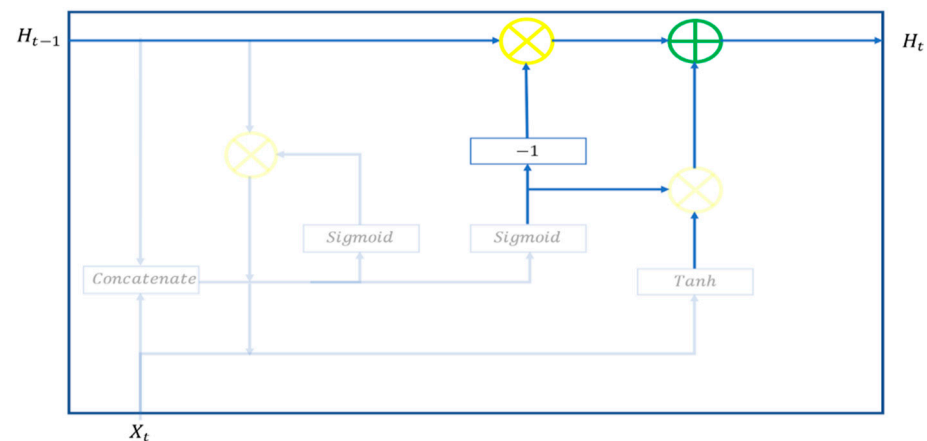


Figure 11. Calculation of output values.

#### 4. Results

Validation is essential to modeling. This paper adopts a few effective methods by referring to prior studies so as to ensure the validation of models [10,32].

First, raw data are normalized and standardized so that data values are between 0 and 1. The varying degrees of change in data due to different units or representations of numbers may affect the results of statistical analysis. Normalization and standardization seek to resolve this problem. Raw data are converted into dimensionless values to facilitate comparison and analysis. Normalization and standardization can optimize gradient descent and enhance accuracy for deep learning algorithms.

Second, the randomly selected data are not sent back to the sampling pool to avoid bias due to repeated data selection. In the RNN and GRU modeling processes, this paper randomly selects 60% from the sourced data as the training dataset, 20% as the validation dataset, and 20% as the test dataset. The training dataset is used for model training and fitting and fine-tuning parameters in the network. The validation dataset shows whether the model is overfitted by showing the change in the loss value of the training dataset and the validation dataset in each epoch of the training process. If so, the training can be stopped in time. The model structure and hyperparameters are then adjusted accordingly. Hyperparameters are validated and confirmed after many iterations, so that the model is in its best status. This greatly saves time and avoids model overfitting. The test dataset is used to assess the generalization ability of the finalized model. The assessment with the test dataset derives some performance indicators.

Third, the loss function serves to indicate the accuracy of model predictions. The smaller the loss function is, the higher is the model accuracy. This study uses the binary cross-entropy. When the loss function is minimized, the classification error rate is the lowest, and the model accuracy is the highest. The loss function is optimized with gradual convergence by updating parameters in multiple iterations to avoid either overfitting or underfitting.

Finally, this study uses multiple model performance indicators, rather than relying on a single indicator. The binary classification indicators include the confusion matrix, suitable for machine learning and deep learning model performance assessments, and Type I error rates and Type II error rates frequently seen in statistics. The confusion matrix consists of accuracy, precision, recall/sensitivity, specificity, and F1-score. The reductions of Type I and Type II error rates are critical to the control of audit failure risks and costs.

The LSTM and the GRU modeling processes and the results in this study are explained as follows.

##### 4.1. Modeling Process

The LSTM and GRU modeling processes divide the raw data into three datasets to provide model overfitting. The data are randomly allocated to the training dataset, the

validation dataset, and the test dataset. The training dataset is used to train the models and the validation dataset is for model validation and adjustment along the way, in order to select the best models. The training with the training dataset and the validation dataset derive a final model to be assessed with the test dataset. The detailed steps are as follows. All data are randomly selected for the training dataset (60% of the data), for deep learning, and for adjusting and fitting of the LSTM model and the GRU model. The parameters of the classifiers are adjusted to establish the best classification models. The model adjustment and fitting process may also involve feature selections and parameter estimates. Feature selection refers to the process of selecting the subsets of relevant features (i.e., attributes) for model construction, in order to avoid redundancy, simplify models, shorten training time, and reduce overfitting. This is followed by the random selection of 20% of the total data for the validation dataset.

The purpose of the validation dataset is to conduct validation and prediction by using the model derived from the training dataset. The prediction made with the validation dataset aims to identify, from the models trained with the training dataset, the one that yields the best results. Model accuracies are recorded in order to select the parameters corresponding to the models reporting the best outcomes, so that model parameters can be adjusted accordingly. In the adjustment of model hyperparameters, the validation dataset derives unbiased estimates by adjusting the models developed with the training dataset. If the error rate goes up with the validation dataset, then it is a signal for overfitting with the training dataset. At this juncture, the training should be stopped. Finally, the remaining 20% of data are used as the test dataset to assess the model's generalization capability. The optimal model derived with the training dataset and the validation dataset is tested with the test dataset for predictions, in order to measure the model's functionality and classification capability. Once the model parameters have been determined, the test dataset is used to assess the model performance.

#### 4.2. LSTM Model and Performance Assessment

As described above, the training dataset and the validation dataset are used for the LSTM deep learning, and repeated adjustments are made until the model becomes stable. The loss function in Figure 12 and the accuracy in Figure 13 gradually converge during the training process and stabilize after 200 epochs (which requires 1.4 s). This suggests that the model showed no overfitting nor underfitting. The training dataset reports an accuracy of 98.70% while the validation dataset is 94.23%.

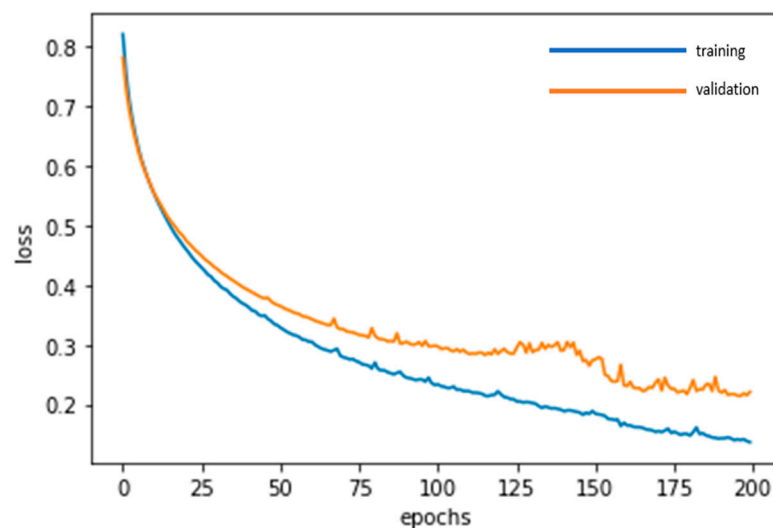
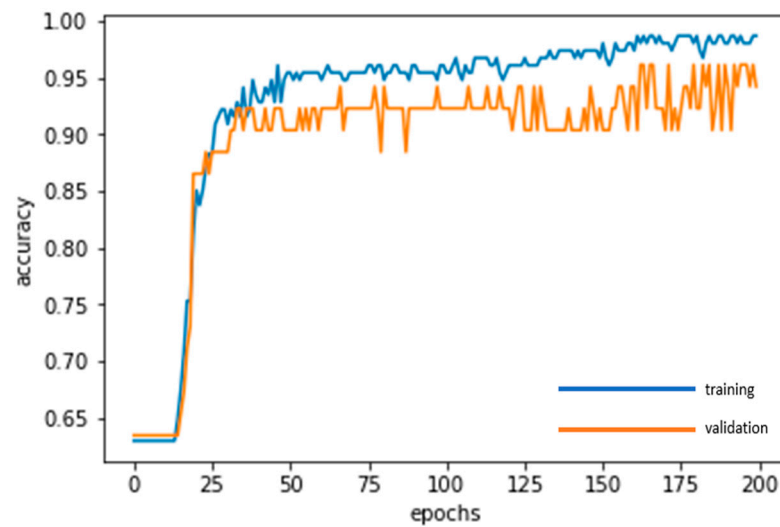


Figure 12. Loss curve of the LSTM model.

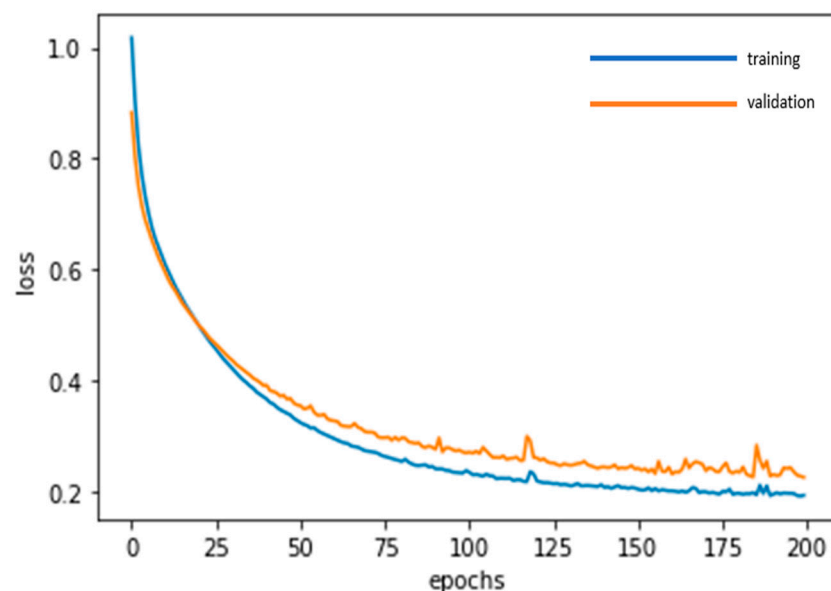


**Figure 13.** Accuracy curve of the LSTM model.

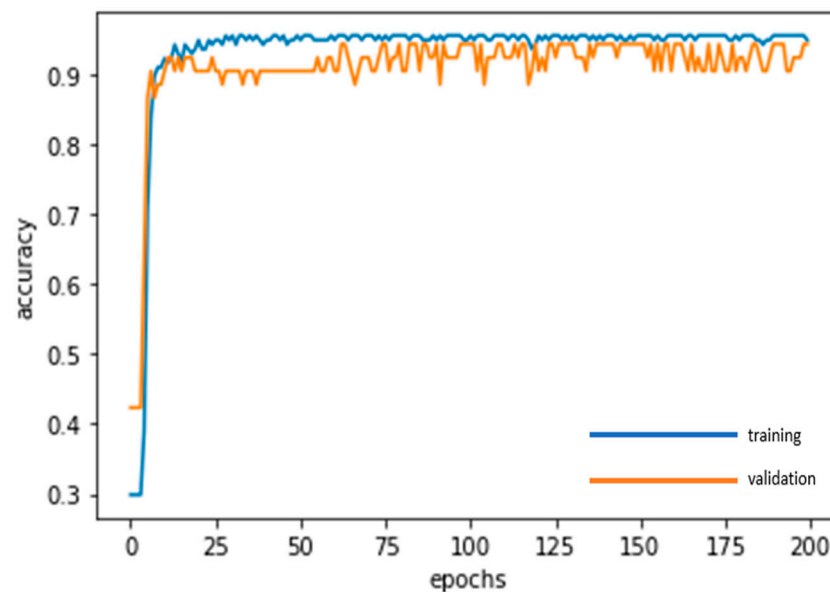
The model performance is assessed with the test dataset. The confusion matrix indicators for the LSTM model are accuracy = 96.15%, precision = 90.00%, recall/sensitivity = 90.00%, specificity = 97.62%, and F1-score = 90.00%. These results suggest the LSTM model performs well. Both the Type I error rate and Type II error rate of the test dataset are 1.92%, or an extremely low level.

#### 4.3. GRU Model and Performance Assessment

In a similar vein, the training dataset and the validation dataset are used for the GRU deep learning and repeated adjustments until the model becomes stable. The loss function in Figure 14 and the accuracy in Figure 15 gradually converge during the training process and stabilize after 200 epochs (which requires 1.4 s). The training dataset yields an accuracy of 94.81%, while the validation dataset is 94.23%.



**Figure 14.** Loss curve of the GRU model.



**Figure 15.** Accuracy curve of the GRU model.

The model performance is assessed with the test dataset. The confusion matrix indicators for the GRU model are accuracy = 94.23%, precision = 94.12%, recall/sensitivity = 88.89%, specificity = 97.06%, and F1-score = 91.43%. These results suggest the GRU model also performs well. The Type I error rate and Type II error rate of the test dataset are 1.92% and 3.85%, respectively, showing an extremely low level of errors.

## 5. Discussion

Both LSTM and GRU are efficient deep learning algorithms, developed to address the two major flaws of the RNN network: inability to retain long-term memory and vanishing gradients. The deep learning models constructed with LSTM or GRU can rapidly and effectively process large volumes of data. The modeling with LSTM or GRU algorithms requires the import of TensorFlow as `tf.keras`. The Adam optimizer and Sigmoid activation function are used for training of 200 epochs at a batch size of 3. This study refers to prior studies in the adoption of a few effective methods to ensure model validation [10,33]. These methods are (1) normalization and standardization of raw data, so that data values come in the range of 0 to 1; (2) randomly selected data are not sent back to the sampling pool to avoid data caused by repeated data selection; (3) the loss function is deployed to assist model prediction accuracy; and (4) multiple model performance indicators are used, instead of single indicators.

This study selects 20 variables, consisting of 16 financial variables and 4 non-financial variables, which are the most frequently used to measure going concern. The training dataset and the validation dataset are used for deep learning and repeated adjustment of the LSTM model and the GRU model, until these models stabilize and the best models are derived. The test dataset is used to assess the generalization capability of models.

Table 3 summarizes the comparison of model performances. The confusion matrix indicators are accuracy, precision, recall/sensitivity, specificity, and F1-score. Both the LSTM model and the GRU model perform equally well according to these indicators. According to the most frequently used model performance indicator, the LSTM model reports an accuracy of 96.15% and the GRU model at 94.23%. The LSTM model and the GRU model have low Type I error rates and Type II error rates, based on the results with the test dataset. This shows very low prediction error rates, which can effectively reduce the risks and costs associated with audit failures. In other words, the empirical results prove that deep learning algorithms can be used in auditing. The LSTM model and the GRU model constructed by this study for going concern prediction are reliable and effective.

**Table 3.** Performance Comparison: LSTM vs. GRU.

Model	Accuracy	Precision	Recall/Sensitivity	Specificity	F1-Score	Type I Error Rate	Type II Error Rate
LSTM	96.15%	90.00%	90.00%	97.62%	90.00%	1.92%	1.92%
GRU	94.23%	94.12%	88.89%	97.06%	91.43%	1.92%	3.85%
Outperformer	LSTM	GRU	LSTM	LSTM	GRU	Equivalent performance	LSTM

The research results and models constructed in this study are not inferior to the previous literature [2,10,18,21–30] using machine learning or deep learning algorithms to predict the going concern. The models established in this study have excellent performance (see Table 3), high accuracy, and low error rates.

## 6. Conclusions

In the era of big data, artificial intelligence (AI), and Industry 4.0, deep learning algorithms have been widely used in practice, such as image and voice recognition, text editing and auto-correction, beating humans in Go, smart chatbots, social media, car/ride hailing, autonomous driving, Medicare, and business intelligence. Academic research is also embarking on projects in different domains. However, the use of deep learning algorithms in going concern predictions remains limited.

“Going concern” is a professional term in the domain of accounting and auditing. It is about the assessment by CPAs and auditors regarding a company as a going concern according to financial statements and relevant risks [2]. It has a profound impact on whether a company can maintain a sustainable operation. The global financial crisis emerging in the U.S. and engulfing the world in 2008–2009 resulted from erroneous assessments and opinions by CPAs on going concern and the sustainability of companies [10]. After the Enron fraud in the U.S. in 2001, the U.S. Congress passed the Sarbanes-Oxley Act (SOX) in 2002 by increasing the burden of independence and legal liabilities on CPAs. CPAs should act with professionalism and independence in the rendering of audit services and issue appropriate audit opinions regarding whether a company is a going concern.

In contrast to the approaches found in the literature, this study uses two powerful deep learning algorithms, LSTM and GRU, for learning and training in order to construct effective and highly accurate models for going concern predictions and to assist CPAs and auditors in the issuance of more accurate audit reports. The input variables include both financial and non-financial variables. Data spanning over a period of 16 years (2004–2019) are sampled. Non-financial variables are also known as corporate governance variables that are important to the assessment of going concern and corporate development and sustainability. The empirical results suggest that both the LSTM model and the GRU model perform well according to indicators such as accuracy, precision, recall/sensitivity, specificity, F1-score, and Type I and Type II error rates. As the most frequently used model performance indicators, the LSTM model and GRU model yield good accuracy performance with the values of 96.15% and 94.23%, respectively. In other words, the LSTM model and the GRU model constructed by this study for going concern predictions are both reliable and effective. These models are successful and useable, able to contribute to the going concern predictions in practice and academic research and can extend the scope of existing literature.

The research findings provide a reference to CPAs, research analysts, appraisers, business consultants, credit rating agencies, company management, and supporting staff and academics in corporate sustainability, risk management, and auditing.

This study provides the following suggestions for CPAs to make going concern decisions and to supervise their clients. First, financial and non-financial indicators should be taken into consideration in the variables to be measured. Second, CPAs should have the courage to issue “qualified opinions”, “disclaimer opinions”, and “adverse opinions” for their clients who have doubts or incomplete financial information that cannot be improved



after communicating [10,36]. Third, CPAs should also do their best to supervise their clients to achieve complete internal control, internal auditing, and corporate governance [10].

This study also provides the following suggestions for future research of going concern. First, in addition to finance ratios, non-financial variables (which can also be called corporate governance variables) should also be included in the research variables. Second, the application and the reference of research variables should be adjusted according to the profile and the pattern of local companies, the environment and functioning of financial markets, audit standards and regulations, company laws, and capital market rules. Third, use machine learning and deep learning algorithms to engage in research on topics related to going concern, especially the more stable and fast deep learning algorithms. In addition to the LSTM and GRU used in this study, future researchers may consider adopting ANN, CNN, DNN, RNN, and other deep learning algorithms.

There are several research limitations in this study. First, this study is limited by the research sample pool of TWSE/TPEX listed companies in Taiwan in the construction of going concern prediction models. Second, the scale of the financial market in Taiwan is not very large; thus, the scale of listed companies is relatively small [36,37]. Third, most of the research variables used to construct the going concern prediction models in this study are from the past data, which may not be able to cope with sudden events that will affect the going concern of a company, such as the COVID-19 global crisis [31]. This is also inevitable in most academic research on similar topics.

**Author Contributions:** Conceptualization, D.-J.C. and C.-C.C.; methodology, D.-J.C. and C.-C.C.; software, D.-J.C. and C.-C.C.; validation, D.-J.C.; formal analysis, D.-J.C. and C.-C.C.; investigation, D.-J.C. and C.-C.C.; data curation, D.-J.C. and C.-C.C.; writing—original draft preparation, D.-J.C. and C.-C.C.; writing—review and editing, D.-J.C. and C.-C.C.; visualization, D.-J.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data can be provided upon request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Davenport, T.H.; Ronanki, R. Artificial intelligence for the real world. *Harv. Bus. Rev.* **2018**, 108–116.
2. Chen, S. An effective going concern prediction model for the sustainability of enterprises and capital market development. *Appl. Econ.* **2019**, *51*, 3376–3388. [\[CrossRef\]](#)
3. Bruynseels, L.; Knechel, W.R.; Willekens, M. Auditor differentiation, mitigating management actions, and audit-reporting accuracy for distressed firms. *Audit. A J. Pract. Theory* **2011**, *30*, 1–20. [\[CrossRef\]](#)
4. Carson, E.; Fargher, N.L.; Geiger, M.A.; Lennox, C.S.; Raghunandan, K.; Willekens, M. Audit reporting for going-concern uncertainty: A research synthesis. *Audit. A J. Pract. Theory* **2013**, *32*, 353–384. [\[CrossRef\]](#)
5. DeFond, M.; Zhang, J. A review of archival auditing research. *J. Account. Econ.* **2014**, *58*, 275–326. [\[CrossRef\]](#)
6. Hay, D.; Knechel, W.R.; Willekens, M. *The Routledge Companion to Auditing*; Routledge: New York, NY, USA, 2014.
7. Geiger, M.A.; Kumas, A. Anticipation and reaction to going-concern modified audit opinions by sophisticated investors. *Int. J. Audit.* **2018**, *22*, 522–535. [\[CrossRef\]](#)
8. Persakis, A.; Iatridis, G.E. Earnings quality under financial crisis: A global empirical investigation. *J. Multinat. Financ. Manag.* **2015**, *30*, 1–35. [\[CrossRef\]](#)
9. Sanoran, K. Auditors' going concern reporting accuracy during and after the global financial crisis. *J. Contemp. Account. Econ.* **2018**, *14*, 164–178. [\[CrossRef\]](#)
10. Jan, C.L. Using deep learning algorithms for CPAs' going concern prediction. *Information* **2021**, *12*, 73. [\[CrossRef\]](#)
11. Gutierrez, E.; Minutti-Meza, M.; Tatum, K.W.; Vulcheva, M. Consequences of adopting an expanded auditor's report in the United Kingdom. *Rev. Account. Stud.* **2018**, *23*, 1543–1587. [\[CrossRef\]](#)
12. Auditing Standard Committee. *Statements on Auditing Standards (SASs) No. 58*; Accounting Research and Development Foundation: Taipei, Taiwan, 2018.
13. Auditing Standard Committee. *Statements on Auditing Standards (SASs) No. 50*; Accounting Research and Development Foundation: Taipei, Taiwan, 2012.

14. Reynolds, J.K.; Francis, J.R. Does size matter the influence of large clients on office level auditor reporting decisions. *J. Account. Econ.* **2000**, *30*, 375–400. [[CrossRef](#)]
15. Auditing Standard Committee. *Statements on Auditing Standards (SASs) No. 57*; Accounting Research and Development Foundation: Taipei, Taiwan, 2018.
16. Auditing Standard Committee. *Statements on Auditing Standards (SASs) No. 61*; Accounting Research and Development Foundation: Taipei, Taiwan, 2018.
17. Auditing Standard Committee. *Statements on Auditing Standards (SASs) No. 62*; Accounting Research and Development Foundation: Taipei, Taiwan, 2018.
18. Goo, Y.J.; Chi, D.J.; Shen, Z.D. Improving the prediction of going concern of Taiwanese listed companies using a hybrid of LASSO with data mining techniques. *SpringerPlus* **2016**, *5*, 539. [[CrossRef](#)] [[PubMed](#)]
19. Gutierrez, E.; Krupa, J.; Minutti-Meza, M.; Vulcheva, M. Do going concern opinions provide incremental information to predict corporate defaults? *Rev. Account. Stud.* **2020**, *25*, 1344–1381. [[CrossRef](#)]
20. Zdolšek, D.; Jagrič, T.; Kolar, I. Auditor's going-concern opinion prediction: The case of Slovenia. *Econ. Res.-Ekonom. Istraz.* **2021**, *1*–16. [[CrossRef](#)]
21. Anandarajan, M.; Anandarajan, A. Comparison of machine learning techniques with a qualitative response model for auditors' going concern reporting. *Exp. Syst. Appl.* **1999**, *16*, 385–392. [[CrossRef](#)]
22. Etheridge, H.L.; Sriram, R.S.; Hsu, H.Y.K. A comparison of selected artificial neural networks that help auditors evaluate client financial viability. *Dec. Sci.* **2000**, *31*, 531–550. [[CrossRef](#)]
23. Brabazon, A.; Keenan, P.B. A hybrid genetic model for the prediction of corporate failure. *Comp. Manage. Sci.* **2004**, *1*, 293–310. [[CrossRef](#)]
24. Kirkos, E.; Spathis, C.; Nanopoulos, A.; Manolopoulos, Y. Identifying qualified auditors' opinions: A data mining approach. *J. Emerg. Tech. Account.* **2007**, *4*, 183–197. [[CrossRef](#)]
25. Martens, D.; Bruyneseels, L.; Baesens, B.; Willekens, M.; Vanthienen, J. Predicting going concern opinion with data mining. *Decis. Support Syst.* **2008**, *45*, 765–777. [[CrossRef](#)]
26. Mokhatab, R.F.; Manzari, S.M.; Bostanian, S. Financial health prediction models using artificial neural networks, genetic algorithm and multivariate discriminant analysis: Iranian evidence. *Exp. Syst. Appl.* **2011**, *38*, 10210–10217. [[CrossRef](#)]
27. Salehi, M.; Fard, F.Z. Data mining approach to prediction of going concern using classification and regression tree (CART). *Glob. J. Manage. Bus. Res. Account. Audit.* **2013**, *13*, 25–29.
28. Chen, S.; Lee, J. Going concern prediction using data mining. *ICIC-ELB* **2015**, *6*, 3311–3317.
29. Yeh, C.C.; Chi, D.J.; Lin, Y.R. Going-concern prediction using hybrid random forests and rough set approach. *Inf. Sci.* **2014**, *254*, 98–110. [[CrossRef](#)]
30. Chi, D.J.; Chu, C.C.; Chen, D. Applying support vector machine, C5.0, and CHAID to the detection of financial statements frauds. *Lect. Notes Artif. Intell.* **2019**, *11645*, 327–336.
31. Jan, C.L. Detection of financial statement fraud using deep learning for sustainable development of capital markets under information asymmetry. *Sustainability* **2021**, *13*, 9879. [[CrossRef](#)]
32. Jan, C.L. Financial information asymmetry: Using deep learning algorithms to predict financial distress. *Symmetry* **2021**, *13*, 443. [[CrossRef](#)]
33. Hamal, S.; Senvar, O. Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. *Int. J. Comput. Intell. Syst.* **2021**, *14*, 769–782. [[CrossRef](#)]
34. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural. Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
35. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. In NIPS 2014 Workshop on Deep Learning. *arXiv* **2014**, arXiv:1412.3555. Available online: <https://arxiv.org/abs/1412.3555> (accessed on 10 July 2021).
36. Chen, S. Detection of fraudulent financial statements using the hybrid data mining approach. *SpringerPlus* **2016**, *5*, 89. [[CrossRef](#)]
37. Jan, C.L. An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan. *Sustainability* **2018**, *10*, 513. [[CrossRef](#)]