An attention embedded DUAL-LSTM method for financial risk early warning of the three new board-listed companies

Xiaojing Chen [Corresp. 1]

[1] School of Accounting, Wuxi Taihu University, Wuxi 214064, Jiangsu, China., Wuxi, China

Corresponding Author: Xiaojing Chen
Email address: chengxj@wxu.edu.cn

Computer and financial fields are both involved in the interdisciplinary topic of financial risk early warning. We suggest an attention-embedded dual Long Short Term Memory (DUAL-LSTM) for the financial risk early warning to deal with the potential and constraints of rapid economic development to improve the precision of the financial risk prediction for the listed businesses on the New Third Board. First, feature fusion attentionally quantifies data characteristics, increasing the robustness and generalizability of data features. The model's predictive power is then increased by creating a dual LSTM model to meet the financial risk. The studies show that the attention-embedded dual LSTM model can achieve 96.9% of the F value scores and is superior to state-of-the-art model (SOTA) such as the Z-score model, Fisher discriminant method, logistic regression, and Back-Propagation network, achieves the advantage of time series in financial risk prediction. Additionally, for predicting financial risk, our algorithm performs flawlessly and effectively

An Attention Embedded Method for Financial Risk Early Warning of the Three New Board-Listed Companies

Xiaojing Chen*

School of Accounting, Wuxi Taihu University, Wuxi 214064, Jiangsu, China.

Corresponding Author:

Xiaojing Chen

68 Qianrong Road, Wuxi, Jiangsu

Email address: chengxj@wxu.edu.cn

Abstract

Computer and financial fields are both involved in the interdisciplinary topic of financial risk early warning. We suggest an attention-embedded dual Long Short Term Memory (DUAL-LSTM) for the financial risk early warning to deal with the potential and constraints of rapid economic development to improve the precision of the financial risk prediction for the listed businesses on the New Third Board. First, feature fusion attentionally quantifies data characteristics, increasing the robustness and generalizability of data features. The model's predictive power is then increased by creating a dual LSTM model to meet the financial risk. The studies show that the attention-embedded dual LSTM model can achieve 96.9% of the F value scores and is superior to state-of-the-art model (SOTA) such as the Z-score model, Fisher discriminant method, logistic regression, and Back-Propagation network, achieves the advantage of time series in financial risk prediction. Additionally, for predicting financial risk, our algorithm performs flawlessly and effectively

Keywords: Three New Board Listed Companies; Financial risk early warning; Attention; LSTM

1 Introduction

New three board-listed companies' financial risk prediction is a hot research point in the economic field. The Financial Risk Early Warning is to establish an early warning mechanism to predict the crisis for high-risk institutions with the intermittent change of economic and financial.

The prediction models of the early financial risk are often math models that calculate by traditional statistics. But these models always consist of linearization modules with strict

34   limitations, which draws the nonlinear relationship within the financial risk difficulty and always
35   obtains the lousy performance to be far away from the actual financial warning research. Later,
36   methods, such as SVM, NN and RF, have been exploited widely to forecast economics and
37   finances by the rapid development of pattern recognition and deep learning, which can deal with
38   complex nonlinear issues quickly and effectively by the nonlinear enhancing modules [1,14]. Among
39   them, SVM is the most popular method, which can easily achieve the optimal strategy in
40   hyperplane through the minimization of risk with perfect generalization. Huang Chao et al. [2] use
41   SVM to construct KPCA-SVM financial crisis, early warning model. Chun Weide and Xiao Yang
42   [3] find that SVM with polynomial kernel function has excellent prediction performance and learning
43   ability to predict financial risks. As we know, predicting financial troubles is a classification task
44   with a time sequence. However, SVM and NN are both traditional methods with invalid features
45   and can't learn sequence features effectively, such as time or steps.
46   In recent years, deep learning has occupied various fields in the community of scholars
47   through the continuous development of computer technology. The advantage of deep learning
48   lies in the gradual learning of multiple networks to extract complex and compelling features [4],
49   which can help models increase the prediction's accuracy. Dixon et al. [5] use DNN to predict the
50   futures price of the Chicago Board of Trade and demonstrate that DNN can simulate the financial
51   time sequence to obtain a better prediction result. Ouyang et al. [6] find that LSTM can effectively
52   capture the difference between the features with different economic time sequences of the Dow
53   Jones Industrial Index. Deep learning is exploited narrowly for financial risk prediction because
54   selecting parameters is hard, which effectively influences the prediction model. According to the
55   optimizing strategies, such as cross-validation and reinforcement learning, Marso and Merouani
56   [7] significantly improve the prediction performance of financial risks and achieve the 90.30% of F
57   value by optimizing the feedforward network with the CSA algorithm. The studies above directly
58   input the data and do not pay attention to the impact of abnormal data on financial risks. In
59   addition, the financial risk of enterprises listed on the New Third Board is affected by time, location,
60   etc. So, we must not only consider the prediction accuracy but pay attention to the time sequence
61   of the model.
62   To solve the mentioned problems and to enrich cycle application areas of the neural network
63   algorithm, we propose an attention-embedded dual LSTM model to predict the early financial risk
64   for the three board-listed companies, which can optimize the network model to adapt the task of
65   financial warning, improve the robustness of the model and increase the accuracy of the New
66   Third Board listed companies financial risk prediction. It is conducive to the empirical analysis of
67   the actual economic situation for the companies listed on the New Third Board.
68   Main contributions of our paper include:
69   To address the issue of the time sequence for financial risk prediction, we propose an
70   attention embedded dual LSTM model to determine the relationships among the attention features
71   from the financial situation of the New Third Board listed companies.
72   We propose a dual LSTM framework to classify the financial situation to risk or safety. The
73   first layer of LSTM is to attend the features and the second layer is to obtain the classification by
74   inputting attended financial features.

75

2 Related work

77   The task of financial risk early warning is to identify whether the financial risks will come in
78 the future. We often propose a discriminant model to help people to obtain an enterprise's future
79 financial situation with the previous financial or non-financial elements. Since then, the financial
80 early warning model has included parts: a) the appropriately selected analysis methods; b)
81 analysis for determining samples; c) the design of the application system, which uses some
82 methods to analyze the future financial situation and get the situation results [8,16,17].

83   Discriminant analysis [12,15], also known as linear discriminant analysis, is a typical
84 representative of statistical analysis. Discriminant analysis is to establish specific discriminant
85 criteria (discriminant functions) with the general condition of the classification results of the
86 predicted objects and the values of multiple characteristic variables that can affect the
87 classification results of the objects so that it can reduce the error probability of classifying the
88 newly observed objects that have not appeared in the past by the discriminant criteria.

89   The process of the discriminant analysis can be shown as formulas:

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|}{\tilde{s}_1^2 - \tilde{s}_1^2} \tag{1}$$

$$\tilde{s}_i^2 = \sum_{y \in w_i} (y - \tilde{\mu}_i)^2 \tag{2}$$

90

91   Logistic regression is also a multivariable model. The main difference between Logistic
92 regression and linear regression is that the explanatory variable Y of linear regression is
93 numerical, while that of Logistic regression is bivariate. The Logistic regression model classifies
94 the samples according to the value of the input variables and divides the samples into Logistic
95 binomial regression and multinomial Logistic regression according to the number of classified
96 categories. The difference lies in whether the input samples are divided into two classes or
97 multiple categories.

98   The process of the logistic regression can be shown as formulas:

$$\prod_{i=1}^{i=k} h(x_i) = \prod_{i=k+1}^{n} (1 - h(x_i)) \tag{3}$$

$$\prod_{i=1}^{i=k} h(x_i)^{y_i} (1 - h(x_i))^{1-y_i} \tag{4}$$

$$L(w) = \frac{1}{n} \sum_{i=1}^{n} -y ln(h(x_i)) - (1 - y_i) ln(1 - h(x_i)) \tag{5}$$

99

100   Artificial neural networks (ANNs) are used in this paper. ANNs can simulate the human brain
101 and the thinking system to process various information and features. There are many famous
102 models, such as the BP network, which are networks that process information for intelligent tasks.
103 The most popular is the BP network which consists of the feedforward network and loss function
104 by error backpropagation. Its key point is to use the gradient search method to find the
105 minimization of the difference between the ground truth and predicted output. It can process the
106 tasks such as classification, detection, recognition and so on.

107   The prediction of early financial risk is a classification problem whose results are generally
108 divided into two categories: with financial risk and without financial risk. It is a classification
109 problem and is the most suitable scenario for applying neural networks. Therefore, it is feasible
110 to predict early financial risk using neural networks for the listed companies on the New Third
111 Board.

112   The financial data of an enterprise is continuous in time. For general neural networks, the
113 data sets processed do not have the characteristics of time series, but the financial data of
114 enterprises are constant in time. The report items affect each other, and the financial information
115 of the previous year will also impact the following year. In particular, the ratio information, such
116 as the year-by-year growth rate and the growth rate compared with the beginning of the period,
117 are also critical financial data. Therefore, it is necessary to add time series based on a neural
118 network and the recurrent neural network [9] to process the sequence information better.

119   3 DUAL-LSTM embedded Attention model

120   This paper proposes a financial risk early warning method for enterprises listed on the New
121 Third Board based on the double-layer LSTM model. The technique takes LSTM as the primary
122 processing unit. It mainly includes three parts: financial data processing, financial data saliency
123 enhancement and financial risk judgment of enterprises listed on the New Third Board[10].

124 3.1 LSTM architecture

125   We apply the gradient decrease strategy in connection weights training to optimize the LSTM
126 network. Considering containing the rapidly disappearing input information in the reverse
127 transmission cycle sequence and the loss of information, we currently apply the LSTM network
128 [11], which is the most widely used network architecture. LSTM can store all the existing information
129 by building a memory function according to the structure of RNN. The memory function is an
130 essential principle for the LSTM model to solve the gradient vanishing problem.

131   Based on RNN, LSTM introduces a storage unit and gated storage unit to store historical
132 information, which can realize the long-term storage state of helpful information. There are three
133 types of gating in the basic unit of LSTM, including input gate, forget gate and output gate. By
134 working together, the information can be stored in the memory unit memory at a step. The
135 structure of neuronal cells in LSTM is presented in Figure 1.

136   LSTM learning comes from the RNN, the new learning process is the need to update the
137 weight and memory units, namely c_t, when the time step t. Set the LSTM hidden layer of input
138 and output vector to x_t and h_t, respectively. To control the current input data x_t stored in the

139   memory cell c_t of the amount of information, an input gate is implemented, which is how much
140   information can be saved to c_t. The formula is shown as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{6}$$

141   The forgetting gate is a critical component of LSTM. Its function is to control what information
142   is retained and what is forgotten and to give up unwanted information. It can mainly play the role
143   of dimensionality reduction and avoid the gradient backpropagation caused by the gradient
144   disappearance. At the same time, time change and the failure of convergence lead to the gradient
145   explosion problem, which can then converge loss function better. The forgetting gate controls the
146   self-connected unit and determines the discarded historical information to achieve the purpose of
147   dimension reduction. Its core is to select the information in the memory unit at the last moment,
148   which has an impact on the current memory unit:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{7}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{8}$$

149   We control the memory unit c_t on the current output value h_t to construct the output gate,
150   which part of the memory unit will output in time step t:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{9}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{10}$$

151   Where W is the linear matrix and b is bias.

152

153   3.2 Dual LSTM embedded Attention model

154

155   Aiming at the related problems of financial risk prediction for enterprises listed on the New
156   Third Board, this paper proposes an attention-embedded dual LSTM model. The model is divided
157   into two modules: Attention mechanism and double-layer LSTM architecture, as shown in the
158   figure.

159

160   3.2.1 Attention
161   There are two methods to classify financial state based on LSTM hidden state: One is to take
162   the final hidden state of LSTM as the input of the classifier, which will lose some information far
163   from the final hidden state; The other is to sum the hidden states of all time steps and then take
164   the average value as the input of the classifier, but it cannot distinguish the influence of the input
165   information of each time step on the composition classification. Therefore, we decide to use the
166   attention mechanism to learn the importance of the input information at each time step to capture
167   the critical information in many financial states and increase the weight.
168   Regard the input vector as LSTM $x_0, x_1,...,x_t$, and the hidden state corresponding to the input,
169   that is, vectors are $h_0, h_1,...,h_t$ Respectively. Then we calculate the similarity S_j between the current
170   hidden state h_j and the final hidden state $\tilde{h}$ by using an attention mechanism. The formula is as
171   follows:

$$S_j = \tanh(W_n h_j + b_n) \tag{11}$$

172   where W_n and b_n are the linearization matrix and bias. At each time step, the weight a_j, which
173   is to process input information, is shown in the following formula:

$$a_j = \frac{\exp(S_j \times \tilde{h})}{\sum_{m=0}^{t}(S_j \times \tilde{h})} \tag{12}$$

174   The hidden state vector v of LSTM after adding attention weight is shown in the following
175   formula:

$$v = \sum_{m=0}^{t}(a_j h_j) \tag{13}$$

176   3.2.2 Dual LSTM embedded Attention

177   Aiming at the financial feature representation $x_0, x_1,...,x_t$ of enterprises listed on the New Third
178   Board, this paper proposes a novel dual LSTM to generate attended vector and warning
179   classification results, respectively. The dual LSTM can be divided into the attention LSTM and the
180   classification LSTM, respectively, whose basic framework is shown in Figure 2.

181   In each time step, the attention LSTM includes the output of the particular classification LSTM
182   at the previous time, the financial feature vector and the generated financial status encoding
183   feature, respectively. The specific representation is shown in the following formula:

$$h_t^1 = \text{LSTM}_{att}\left(h_{t-1}^2, x_t\right) \tag{14}$$

184   where h_{t-1}^2 is the output of the Classification-LSTM at the previous time step. These inputs

185   provide the status information of the classified LSTM, the context information of the financial state
186   and the maximum context information of the generated financial state, respectively.

187   After obtaining the output h_t^1 of the attention LSTM at time t. At each time step, the attention

188   weight $a_{j,t}$ is generated for each financial feature x_t:

$$a_{j,t} = W^T \tanh\left(W_x x_t + W_h h_t^1\right) \tag{15}$$

189   where W^T, W_x, W_h are learnable matrixes. Then according to equation (8), the financial

190   features $\tilde{x}_t$ are input to the classification LSTM. After attention financial features of LSTM are
191   obtained, the classification results y are generated by combining the output of the classification
192   LSTM.

$$h_t^2 = \text{LSTM}_{cls}\left(\tilde{x}_t, h_t^1, h_{t-1}^2\right) \tag{16}$$

$$y = \text{softmax}\left(W_{cls} h_t^2 + b_{cls}\right) \tag{17}$$

193   where W_{cls} and b_{cls} are learnable linear transformation matrix and bias, respectively.

194   The Financial indicators adopted by the model training are shown in Table 1.

195

4 Experiment and analysis

4.1 Dataset

Considering the availability of data, our paper selects listed companies on the New Third Board to provide a research idea of financial data of a specific enterprise type. The 2014-2018 public financial data of 100 enterprises are selected as the explanatory variables of the model for the experiment. The 2018 financial data is taken as the actual value and the 2014-2017 data is regarded as the training sample.

4.2 Implement the details

We implement our experiments with the i7-12900k Cpu, 3080ti Gpu and the Pytorch deep learning framework.The parameter settings of the attention embedded dual LSTM model are shown in Table 2.

In this paper, we evaluate the performance of the early financial risk prediction by standard classification accuracy rate. Considering the unbalanced samples, we apply the evaluation index to represent the prediction performance. The classification accuracy is Recall (R), Precision (P) and F. The sample measure standard F is quoted to evaluate the classification of unbalanced samples. If F is large, it indicates the performance of the financial risk prediction is superior and vice versa. The formulas of the evaluation metrics are presented by:

$$R = \frac{TP}{TP + FN} \tag{18}$$

$$P = \frac{TP + TN}{TP + FP + TN + FN} \tag{19}$$

$$F = \frac{2 \times \frac{TP}{TP + FP} \times \frac{TN}{TN + FN}}{\frac{TN}{TN + FN} + \frac{TP}{TP + FP}} \tag{20}$$

4.3 Results and Discussion

4.3.1 Comparison results of different models

To evaluate our proposed model in this paper, several classical financial early warning models are selected for experimental comparisons, such as the Z-score model [11], Fisher discriminant method [12], Logistic regression [13] and BP network on the same dataset. We present the results as shown in Table 3.

Because this article focuses on the early warning of financial risk, the ST companies' identification into the recalling rate of ST companies is even more critical. As seen above, the dual LSTM embedded Attention model has better performance, and the dual LSTM model can achieve a value of more than 90% in terms of recall, precision and F measure, which is a specific identification model of ST and the l performance of the classifier is better. First, the F value of the

228 attention-embedded dual LSTM is increased by more than 18% compared with the original Z-
229 score because all the model structures designed in this paper are ahead. Secondly, our method
230 improves the F value by more than 5% to the Fisher's discriminant in all three evaluation indexes.
231 Then, the effect of these two methods, Logistic and BP neural network, are significantly improved,
232 which are still inferior to our proposed method. Finally, the proposed method achieves the best
233 experimental results by embedding the Attention module and creatively using the two-layer LSTM
234 as a financial warning.
235
236 4.3.2 Influence of Hidden layer size and Layers number
237   To discuss the number of LSTM and the size of hidden layers, we set the batch size to 4 and
238 epochs =50 during training to experiment. In the experimental process, we will complete the
239 determination of hyperparameters by comparing the recognition accuracy during training.
240   The dimension size of the hidden layer can be expressed as a bar graph for the model training
241 results (Acc refers to the correct rate of recognizing ST into ST). As shown in Figure 3, we can
242 find that the model's performance continues to rise with the increase of dimension until the hidden
243 layer dimension reaches a peak of 1024, where the model achieves the best training accuracy.
244 Therefore, we choose the dimension of the hidden layer to be 1024.
245   The number of layers can be represented as a histogram (Acc refers to the correct rate of
246 recognizing ST into ST). As shown in Figure 4, we can also find that the model's performance
247 continues to rise with the number of model layers increasing. When the length of the LSTM is 6,
248 our method achieves the best results and training accuracy. So, we set the number of the model
249 layer to 6.
250   The above results directly affect the training process's rigor and positively affect the financial
251 risk prediction of the Three New Board listed enterprises.
252 4.3.3 Comparison of Prediction Stability
253   To compare the predictive stability of our model with the Z-score model, Fisher discriminant
254 method, Logistic regression and BP neural network model Since the 1:1 of the rate can affect the
255 randomness. This paper changes the proportion between positive and negative samples by
256 adding more positive samples corresponding to the negative samples one by one to investigate
257 the robustness of the dual LSTM embedded Attention model. Specifically, the ratio between the
258 financial risk sample and the safe financial sample is reset, following the ratio of 1:2, 1:3, 1:4, 1:5
259 and 1:6, respectively. Then, we apply these five proportions to structure the five types of models,
260 which are to evaluate the prediction stability for the financial risks. Because the samples, which
261 are chosen with the five proportions, are not balanced samples, so we exploit a special
262 measurement F for the non-balanced samples to obtain the performance of the prediction. The
263 experiment results are presented in Figure 5.
264   We can find that the F value of our proposed model is higher than other models with any
265 proportions of positive and negative samples, which indicates that Z-score, Fisher discriminant,
266 Logistic regression and BP network are inferior to our method for both financial risk samples and
267 safe samples. Graphically, the F value of our proposed model is relatively more stable in trend
268 and less volatile, which indicates that the prediction of the financial risk samples and safe samples

269   are accurate with different proportions.

5 Conclusion

271   Financial risk early warning has been one of the hot spots research. Considering the

272   complexity of the financial risk prediction of the New Three Board listed companies, this paper

273   proposes an attention embedded dual LSTM model to predict financial early risk, which can be

274   successfully applied to New Three Board listed enterprise financial risk forecast. Firstly, the

275   financial features of listed enterprises on the New Three Board are fused and we quantify them to

276   input the attention embedded dual LSTM model. Then, a new embeddable Attention module is

277   designed to enhance the representation ability of financial features. Finally, a dual LASM model

278   with embedded Attention is designed to predict the financial risks of enterprises listed on the New

279   Third Board. Compared with the other four models, the results can demonstrate that the dual LSTM

280   model embedded Attention has better performance of prediction in terms of all evaluate metrics

281   and the F value is increased by 17.31%, 8.64%, 1.13% and 2.28%, respectively, where the

282   accuracy is statistically significant. In the future, we will explore to extend the financial risk

283   prediction to the all companies. Then, we will try to research on the attention to find out more

284   important point into financial situation of companies and to attended them to increase the accuracy

285   of prediction.

286

Data Availability

288   The dataset can be accessed upon request.
289

Conflicts of Interest

291   The authors declare that there are no conflicts of interest.

Funding Statement

293   This work is funded by "the Construction Plan of Scientific Research and Innovation Platform of
294   Wuhan College", the project number is KYP202001; And "and supported by the Research and
295   Innovation Team of Wuhan College" , the project number is KYT201903. We would like to thank

296 Research Center for Value Evaluation and Creation of Private Enterprises and Audit Value
297 Innovation Research Team in the New Era which contribute to this study greatly.

Acknowledgments

299 I would like to thank the anonymous reviewers whose comments and suggestions helped improve
300 this manuscript.

301

Reference

303 [1] Kingma D P, Ba J. Adam. A Method for Stochastic Optimization.Computer Science,
304   2014,(v1).
305 [2] Huang C, Han T T, Wu F, et al. Financial Crisis Prediction based on Biorthogonal Wavelet
306   Mixed Kernel KPCA-SVM Model. Journal of system management, 2015, 24(1). (in Chinese)
307 [3] Chun W D, X Y. Research on SVM early warning of potential systemic financial risk in the
308   period of supply chain structural reform. FORECASTING, 2018, 37(5). (in Chinese)
309 [4] Ni Z W, Xue Y J, Ni L P, et al. Research on multi-core SVM financial early warning method
310   by Manifold Learning. System engineering - theory & Practice, 2014, 34(10). (in Chinese)
311 [5] Dixon M, Klabjan D, Bang J H, et al. Implementing Deep Neural Networks for Financial Market
312   Prediction on the Intel Xeon Phi. High Performance Computational Finance,2015.
313 [6] OuYang H B, Huang K, Yan H J. Financial time series prediction based on LSTM neural
314   network. Chinese Management Science, 2020, 28(4). (in Chinese)
315 [7] Marso S, Merouani M E. Predicting Financial Distress Using Hybrid Feedforward Neural
316   Network with Cuckoo Search Algorithm. Procedia Computer Science, 2020, (170).
317 [8] Niu FangQi. Research on Financial Risk Prediction Model of Small and Medium-sized
318   Enterprises Based on LSTM Model. China Academy of financial Sciences, 2021. (in Chinese)
319 [9] Zaremba W, Sutskever I , Vinyals O . Recurrent Neural Network Regularization. Eprint Arxiv,
320   2014.
321 [10] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation, 1997,
322   9(8):1735-1780.
323 [11] Yan B H, Ma G Q. Empirical Analysis of Financial Risk of Chinese Real Estate Companies
324   based on Z-Score Model. Finance, 2011(5):5. (in Chinese)
325 [12] Huang L W. Fisher STEPWISE discriminant Analysis Based on Optimal Variables. System
326   science and Mathematics, 2021, 41(08):2338-2348. (in Chinese)
327 [13] Wang Liguo, Jung H, et al. CPAT: Coding-Potential Assessment Tool using an alignment-
328   free logistic regression model. Nucleic acids research, 2013.
329 [14] Ahmad, M., Shabbir, S., Roy, S. K., Hong, D., Wu, X., Yao, J., ... & Chanussot, J. (2021).
330   Hyperspectral image classification—Traditional to deep models: A survey for future
331   prospects. IEEE Journal of Selected Topics in Applied Earth Observations and Remote
332   Sensing, 15, 968-999.

333   [15] Ahmad, M., Ghous, U., Hong, D., Khan, A. M., Yao, J., Wang, S., & Chanussot, J. (2022). A
334   Disjoint Samples-Based 3D-CNN With Active Transfer Learning for Hyperspectral Image
335   Classification. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-16

Figure 1

The framework of the LSTM basic cell

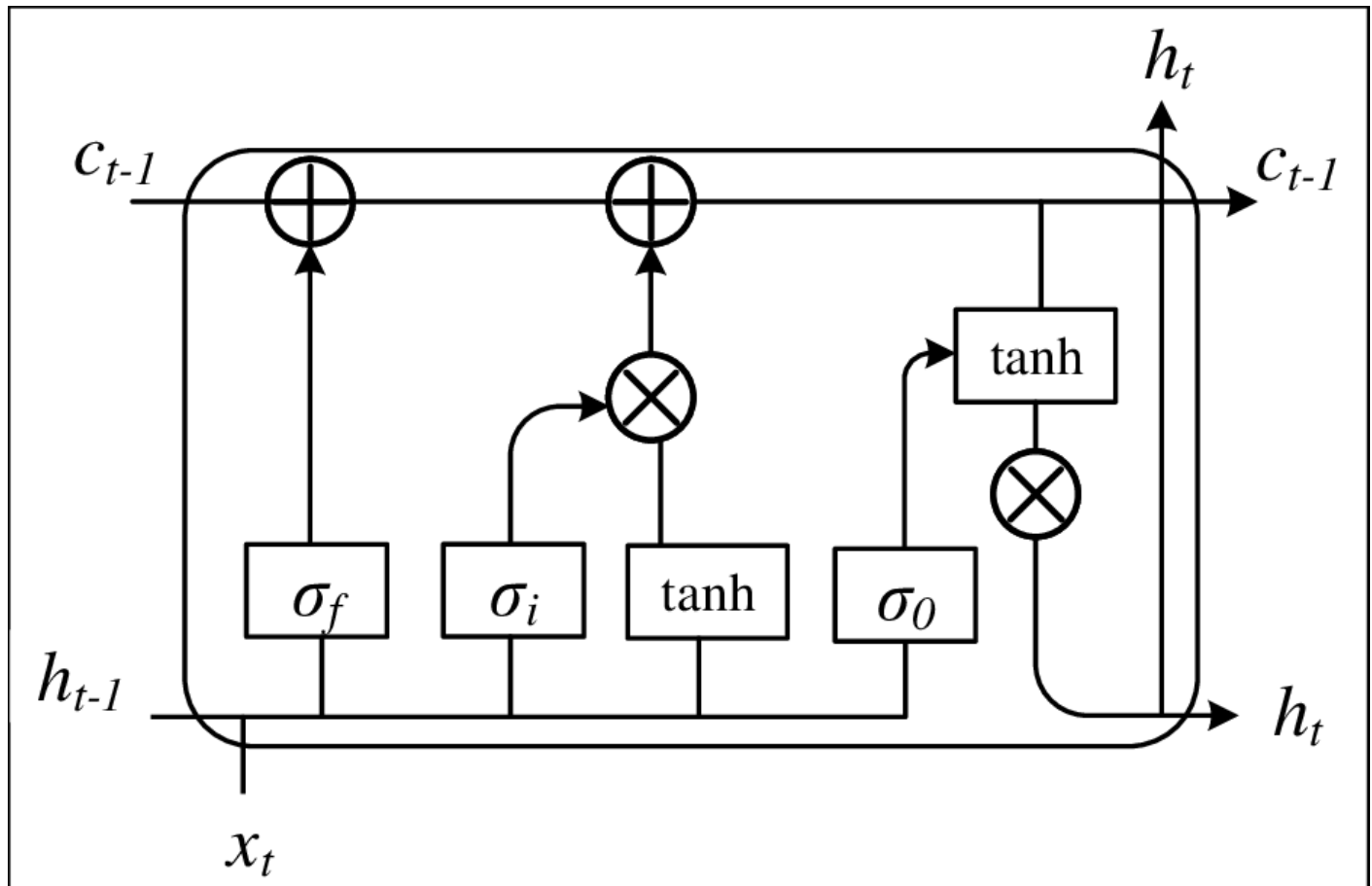



Figure 2

The framework of the LSTM basic cell

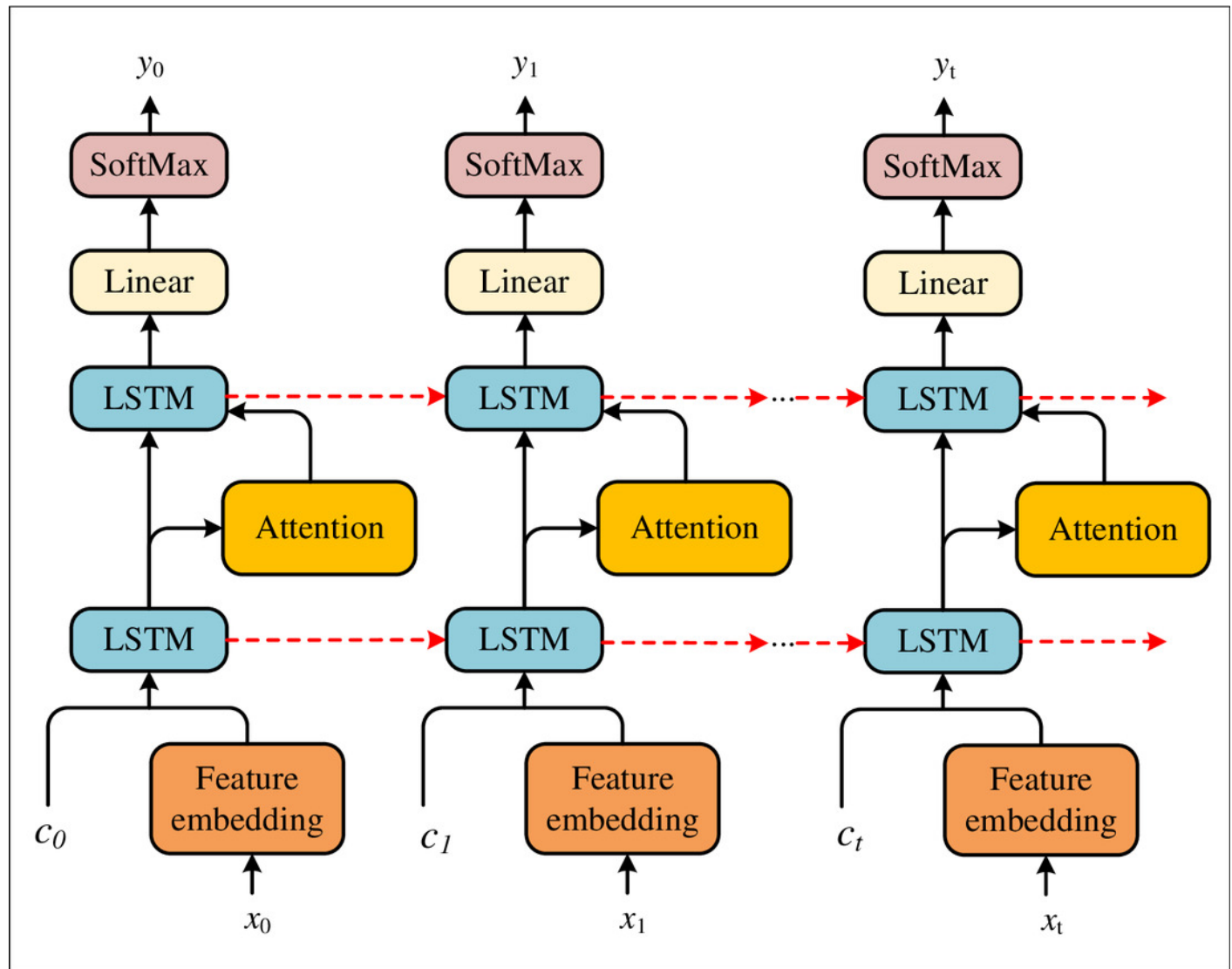



Figure 3

The accuracy of our method with a different dimension of LSTM hidden layer.

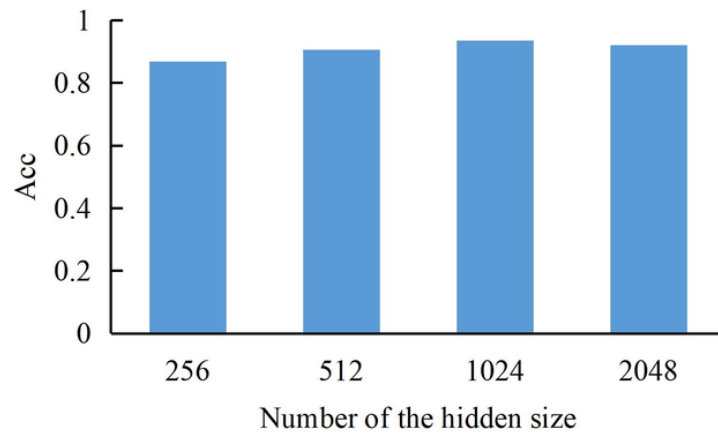



Figure 4

The accuracy of our method with a different number of LSTM hidden layers.

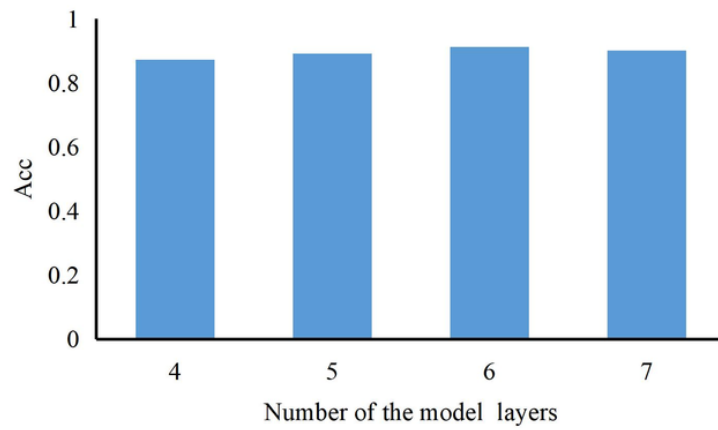
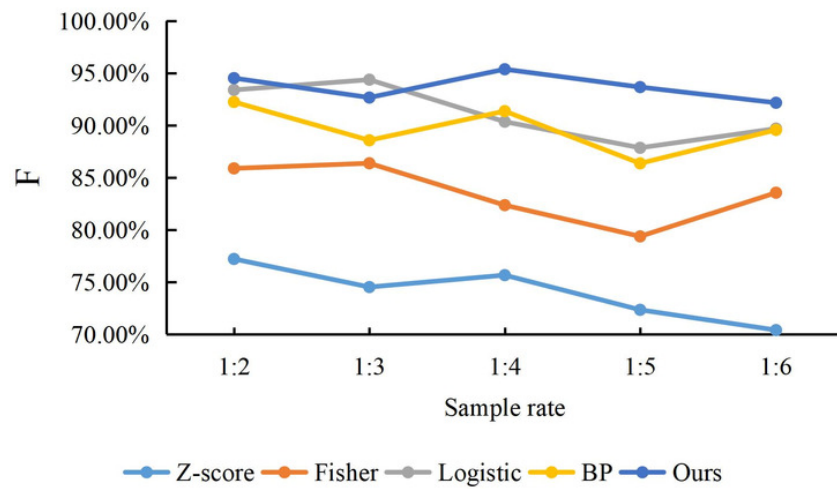



Figure 5

The accuracy of methods with different sample rate

Financial indicators

Profitability	Total assets compensation
	Net asset income
	Net profit from sales
	Gross sales profit
Solvency	Asset liability ratio
	Equity ratio
	Tangible assets / net debt
Cash flow	Cash received from selling goods and providing services
	Cash from operating activities
	Net profit (cash)

Training setting

parameter	The values
Learning rate	0.002
Dropout	0.3
Model optimizer	Adam
Maximum sequence length	6
Number of Epoch	200
Batch size	50

Table 3 Comparative experiments with methods based on the deep learning framework

Table 3 Comparative experiments with methods based on the deep learning framework

Method	R	P	F
Z-score	77.94%	76.47%	77.20%
Fisher discriminant	82.35%	89.71%	85.87%
Logistic	92.65%	94.12%	93.38%
BP Neural Network	91.30%	93.18%	92.23%
Ours	93.33%	95.36%	94.51%