ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20221110173532
Original scientific paper

Internet Financial Credit Risk Assessment with Sliding Window and Attention Mechanism LSTM Model

Menggang LI, Zixuan ZHANG, Ming LU, Xiaojun JIA*, Rui LIU, Xuan ZHOU, Yingjie ZHANG

Abstract: With the accelerated pace of market-oriented reform, Internet finance has gained a broad and healthy development environment. Existing studies lack consideration of time trends in financial risk, and treating all features equally may lead to inaccurate predictions. To address the above problems, we propose an LSTM model based on sliding window and attention mechanism. The model uses sliding windows to enable the model to effectively exploit the contextual relevance of loan data. And we introduce the attention mechanism into the model, which enables the model to focus on important information. The result on the Lending Club public desensitization dataset shows that our model outperforms ARIMA, SVM, ANN, LSTM, and GRU models.

Keywords: attention mechanism; credit risk; internet finance; LSTM; sliding window; time-series prediction

1 INTRODUCTION

With the rapid development of Internet finance, Small Medium and Micro enterprises or individuals at a lower cost were allowed to obtain loans. The huge scale of Internet finance has had a wide-ranging impact on the financial market and the real economy [1]. But compared with the traditional financial industry, Internet finance has more complex credit risk issues. Accurately assessing the reimbursement ability of borrowers, can drop financial risks of lenders, reduce credit loan fraud, and ensure the security of Internet financial transactions [2].

The methods of analyzing and predicting financial markets mainly include traditional statistical methods and machine learning methods. The traditional statistical models comprise logistic regression (LR), discriminant analysis, risk index models, and conditional probability models among others, which model, analyze and predict default risk by finding an optimal linear combination of input variables. Luo demonstrates the advantages of multivariate discriminant analysis and logistic regression over neural networks in medium horizon bankruptcy prediction of Tunisian firms [3]. Serrano-Cinca used the Partial Least Squares Discriminant Analysis (PLS-DA) model to predict the US banking crisis and obtained results close performance to SVM [4]. A LR model was applied to the evaluation of the credit risk of P2P lending platform Paipaidai borrowers [5]. Due to their strong interpretability, Traditional statistical models are welcomed by some well-known international rating agencies (such as S&P, Moody's and Fitch) [6], but this approach often ignores the complex interactions between financial variables. Traditional statistical model is a method with high interpretability and widely used in various fields [7, 8].

The combination of machine learning and financial credit risk prediction is a cutting-edge idea in recent years. By analyzing the historical data of online lending, machine learning explores the internal logic behind the data, and then accomplishes the evaluation and prediction of credit risk. It mainly includes Decision Tree(DT) [9], Boosting Tree [10], Random Forest(RF), Gradient Boosting Decision Tree(GBDT) [11], LightGBM algorithm, BP Neural Network [12] and Deep Convolution Neural Network. The study proposes an RF-based classification method to predict borrower status and outperforms FICO credit score [13]. Ma used the LightGBM algorithm to predict the default rate, which performed better than the XGBoost algorithm. And he found that the vital factor for effect to P2P online borrowers is the loan details [14]. Wu et al. target the behavioral patterns of borrowers in China and employ a deep multiple kernel classifier (DMKC) for credit score prediction [15]. Xia et al. argue that macroeconomic variables, including interest rates, inflation, and unemployment rates, can directly affect credit risk. Therefore, they take macroeconomic variables into account in their model and evaluate credit scores by combining a deep learning model and an integrated tree-based model [16].

However, because of the huge and complex data, it is difficult to obtain the characteristics of borrowers through recent data mining and traditional machine learning [17]. Therefore, Kim proposes a deep dense convolutional network for repayment prediction in social lending [18]. Gunnarsson et al. compared the performance of LR, DT, RF, XGBoost, MLP and deep belief networks models. The results show that XGBoost is the best-performing credit scoring method [19]. These new classifiers are prior to C4.5 decision trees, ensemble voting classifiers, random forests, and k-order neighbor classifiers. Wang used the Long Short Term Memory (LSTM) model for the first time to analyze P2P sequence data, which has greater potential than traditional time series models [20].

In other fields, the LSTM model shows a great result. During the soaring of the Internet of Things, Cloud Computing and Artificial Intelligence, the world is moving into Industrialization 4.0. Transition to Industry 4.0 is expected to cause formidable structural changes, productivity increments and competitiveness in the manufacturing industry all over the world [21]. The combination of deep study and manufacturing, consulting the schedule for job shop problems [22, 23], optimization management of storehouse [24], to bring up productivity and descend inventory costs. During the electrical business workshop, Wu and Wang designed and achieved an intelligent algorithm based on the LSTM model [25].

In stock forecasting, LSTM model can improve the accuracy of stock return forecasts by 12.9% [26]. Better performance in stock market price prediction by combining LSTM model and Empirical Mode Decomposition [27]. In e-commerce, LSTM models can effectively predict the future behavior of customers [28].

The above references provide ideas for this research, but lack consideration of the following two aspects: First, borrower feature extraction is the key to credit risk assessment. How to more effectively extract features from loan data is an urgent problem to be solved. Second, some existing research have not delved into time series information in credit risk forecasting. Financial time series are inherently complex, noisy, dynamic, nonlinear, nonparametric and chaotic [27]. Therefore, the study of financial series plays an important role in the prediction of credit risk.

This paper constructs an Internet financial credit risk evaluation model based on sliding window and attention mechanism LSTM (SL-ALSTM). The model can effectively perform financial time series forecasting. The contributions of this paper are as follows:

1. We highlight the temporal trend of financial market risk, and relative to a model that treats all features equally, we believe that some features are the main reasons for loan defaults.
2. We develop a new method SL-ALSTM, which captures the temporal correlation features of financial markets and calculates the importance of different features.
3. We evaluate the proposed method on a real-world dataset. Experiment results show that compared with ARIMA, SVM, ANN, LSTM and GRU models, the accuracy of the SL-ALSTM model predictions has been significantly improved.

The structure of this paper is as follows. Section II introduces the techniques used in the proposed method, including the Sliding Window, Attention Mechanism and LSTM. Section III introduces the SL-ALSTM model proposed in this paper. Section IV presents the experiments and analysis of the results. The last section summarizes our research.

2 THEORY AND METHOD

2.1 Sliding Window

The Sliding Window is a method, on the whole time series to choose a part of specified unit length, and calculate the chosen time series. Assuming that a certain loan information X has a total of n data sets in time series, the time series is constructed as $X(n) = [x(1), x(2), ..., x(m), ...x(n)]$. $x(m)$ represented loan default rates of the loan information X among m-th time. If m-th time is the current moment, thus time series of loan information X in prior time t, show as $X(m) = [x(m-t+1), ..., x(m-1), x(m)]$, and t is the size of sliding window. When making monthly new loan default rate forecasts, according to the t value, the historical monitoring data of the loan information X is filled in the sliding window in turn as the input of the LSTM prediction model, and then we can get the predicted value $x(m + 1)$ at time $(m + 1)$.

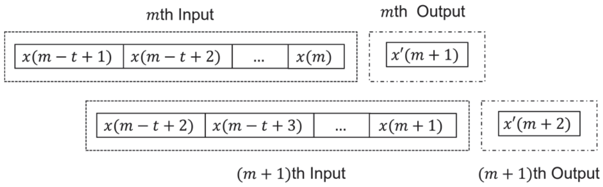


Figure 1 The structure of the sliding window

The structure diagram of the sliding window is shown in Fig. 1.

2.2 Attention Mechanism

The attention mechanism is a model that simulates the attention of the human brain. It can be used as a resource allocation scheme, which is the main means to solve the problem of information overload. In the case of limited computing power, it can process more important information with limited computing resources [29]. The essential idea of the attention mechanism is derived from the encoder-decoder structure, that is, the input sequence $\{x_1, x_2, x_3, ..., x_m\}$ is encoded by the encoding module and transformed into a semantic encoding C. Then it is input to the decoding module for decoding and transformed into the output sequence $\{y_1, y_2, y_3, ..., y_n\}$. However, as the length of the sequence increases, the previous information will be overwritten by the later input. The attention mechanism can solve this problem fundamentally. In the decoding process, it needs to calculate the degree of influence of different inputs on the predicted value, and then obtain energy score e. After that, the energy score e is mapped by the attention distribution function g to get the attention weight α. Different weights are assigned to different inputs to obtain the context vector c. The mathematical form of the attention mechanism can be given by the following Eqs. (1) to (3).

$$e = f(x_1, x_2, x_3, ..., x_m) \tag{1}$$

$$\alpha = g(e) \tag{2}$$

$$c = \sum \alpha_i x_i \tag{3}$$

where f is the score function.

2.3 LSTM

The Long short Term Memory neural network (LSTM) is a variation of recurrent neural network. Though the standard Recurrent Neural Network (RNN) may dig out time series information in data, if the time series processed is too long, the gradient will disappear or the gradient will explode, which will cause the model unable to train. The LSTM model solves the problem that RNN cannot handle long-term sequences by using input gate i_t, forget gate f_t and output gate o_t, so that the model only remembers the useful information hidden in the long-term sequence. Input gate i_t controls the size of new memory contents added to memory. The forget gate f_t determines the amount of storage that needs to be forgotten. The output gate o_t is used to adjust the amount of output memory content. The mathematical form of the LSTM can be given by the following Eqs. (4) to (9).

$$i_t = \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i) \tag{4}$$

$$f_t = \sigma(x_t W_{xf} + h_{t-1} W_{hf} + b_f) \tag{5}$$

$$o_t = \sigma(x_t W_{xo} + h_{t-1} W_{ho} + b_o) \quad (6)$$

$$\tilde{c}_t = \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c) \quad (7)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (8)$$

$$h_t = o_t \odot \tanh(c_t) \quad (9)$$

where σ denotes sigmoid activation function, and W_{xi}, W_{xf}, W_{xo}, W_{xc} are weight parameter, b_i, b_f, b_o, b_c are the bias value, c_t represents the cell state, h_t represents the final output. The internal structure of the LSTM model is shown in Fig. 2.

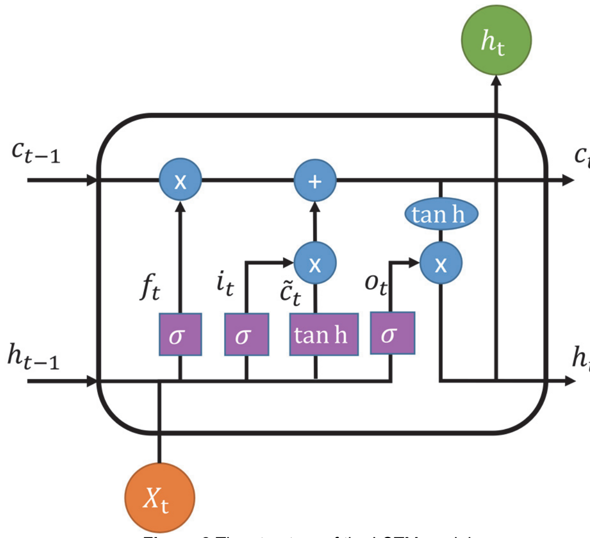


Figure 2 The structure of the LSTM model

3 PROPOSED MODELS

In this section, our proposed model SL-ALSTM will be described in detail.

SL-ALSTM consists of four main layers: sliding window layer, attention layer, LSTM layer and dense layer. In the sliding window layer, the data needs to be segmented by sliding window, and the length of the sliding window is the length of the input data for each model; the Attention layer is mainly responsible for extracting the key information in the loan data; the LSTM layer processes the financial time series; the dense layer outputs the final prediction results. The model structure diagram is shown in Fig. 3.

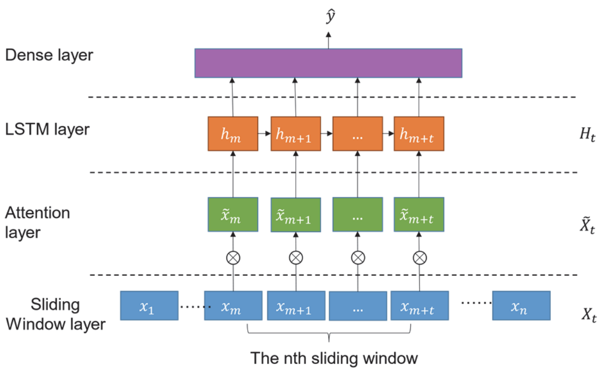


Figure 3 The structure of the SL-ALSTM model

3.1 Sliding Window Layer

To clearly describe the model structure, let $D = \{(x_1, y_1), (x_2, y_2), ..., (x_t, y_t)\}$ be a set of t training samples, where $x_i = (x_i^1, x_i^2, ..., x_i^n)$ describing the input vector of the ith example, describing n features of the monthly loan status. The historical time series data $X_t = \{x_1, x_2, ..., x_t\}$ is used for loan default rate forecasting. We assume that the length of the sliding window is t, and we want to predict the default rate at the future $(t + k)$-th time. First sliding window, the data from the initial time to tth time are input to the model to predict the default data at $(t + 1)$-th time in the future. Second sliding window, the data from the second time to $(t + 1)$-th time are input into the model to predict the default data at the future $(t + 2)$-th time, until it predicts the future $(t + k)$-th time data. In the sliding prediction process, if the real loan default rate has been obtained statistically and the predicted data is covered in time, the latest real data is used to replace the predicted value as the model input data.

3.3 Attention Layer

The attention mechanism can focus on key features to reduce the influence of non-key features on the prediction, and it is considered as a fully connected layer and a softmax function. The working process of the attention mechanism in SL-ALSTM is detailed below.

The output x_t of the previous stage is used as input to obtain u_t through a perceptron layer. u_t represents the degree of influence at represent output for the lending market state at a time t. The equation is as follows.

$$u_t = \tanh(W_i x_t + b_i) \quad (10)$$

Then, it uses the softmax function to get the normalized weight α_t. α_t is formulated as follows.

$$\alpha_t = \frac{\exp(u_t^T \mathrm{u}_w)}{\sum_{j=1}^{T} \exp(u_t^T u_w)} \quad (11)$$

This step measures the importance of each feature by calculating the similarity of u_t to the context vector u_w. Where T is the size of the sliding window. $\exp(\cdot)$ is the exponential function. u_w is a learnable parameter that is randomly initialized and during the training process.

After that, the $\tilde{x}_t$ is the part of the output of the attention layer, and it can be expressed as:

$$\tilde{x}_t = \alpha_t * x_t \quad (12)$$

Through continuous learning and optimization of the corresponding weights, the key information in the input features is highlighted, and the nonlinear feature between variables can be further explored in depth.

3.2 LSTM Layer and Dense Layer

LSTM is specifically used for sequence modeling. In this part, we use the output of the attention process to capture the nonlinear relationships in the borrowing and lending time series through the LSTM model. The end of the proposed model is a predictor, which is a single layer neural network used to obtain the final result $\hat{y}$. The predictor for default rate value refers to Eqs. (13) to (14).

$$h_t = LSTM(\tilde{x}_t) \tag{13}$$

$$\hat{y} = W_d h_t + b_d \tag{14}$$

where W_d and b_d are learnable parameters. h_t is the output of the LSTM model.

4 EXPERIMENT RESULTS

4.1 Lending Club Dataset

In this paper, the desensitization data published on the official website of Lending Club is selected for experiments. The time range is from 2009 to 2018, and the data set has a total of 2 260 701 samples. The data includes personal information, loan information, credit information and so forth. The Loan Status variable is the main focus of this article. It has seven states: Fully Paid, Charge Off, Default, Current, In Grace Period, Late (16-30 days), Late (31-120 days). This paper believes that except for the Fully Paid state, the rest of the states should be regarded as defaults.

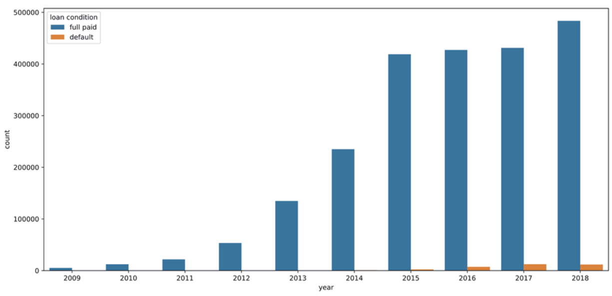


Figure 4 Total borrowings and defaults by year

As shown in Fig. 4, we count the total borrowings and defaults for each month in the raw data.

4.1 Data Per-Processing

Data preprocessing mainly deals with missing values and outliers in the original data. First, variables with more than 50% missing data were removed, and then the data were filled by simple mode and mean imputation. Second, the average default rate of new loans per month is calculated as the dependent variable. Some features were calculated as independent variables, including the average monthly loan interest rate, the percentage of loan grades per month, and the percentage of verification status per month and so on.

In order to eliminate the adverse effects caused by singular sample data and improve the training speed of the model, the data is normalized and processed as a decimal between $(0,1)$. The normalization formula is shown in Eq. (15).

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{15}$$

where X_{norm} is the data after normalization, X represents the original data, $X_{\max}$ and $X_{\min}$ denotes the maximum value and the minimum value in the original data.

4.2 Evaluation Indices of Model Performance

Models built on time series data must make accurate forecasts on two-dimensional metrics: forecast accuracy and forecast trend accuracy. Therefore, we choose root mean square error (RMSE), mean absolute error (MAE) and direction accuracy (DA) as the evaluation metrics in this paper. Among them, RMSE and MAE are used to evaluate the prediction accuracy of the model predictions. The formulas of the three statistical indicators are shown in Eq. (16), Eq. (17) and Eq. (18):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_l)^2} \tag{16}$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|(y_i - \hat{y}_l)| \tag{17}$$

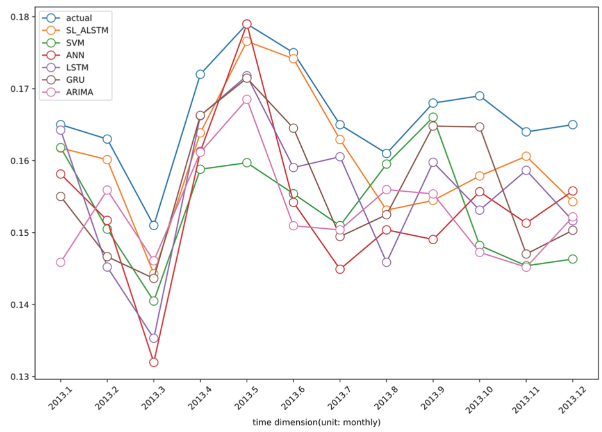
$$DA = \frac{1}{m}\sum_{i=1}^{m} a_i, \quad a_i = \begin{cases} 1, & (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0 \\ 0, & (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) \le 0 \end{cases} \tag{18}$$

4.3 Result Analysis

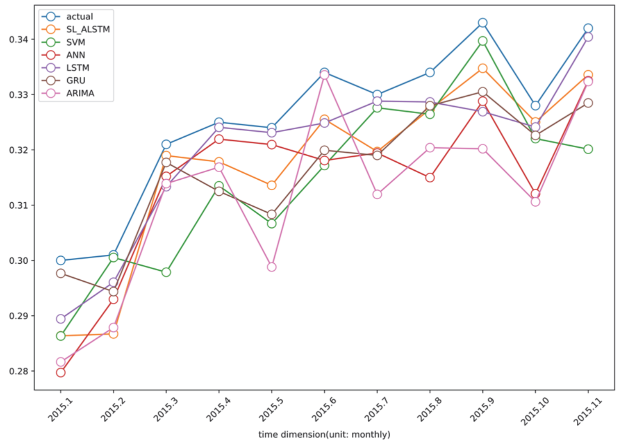
In order to prove the advantages of the SL-ALSTM method, we make a detailed comparative analysis. We selected five comparison methods, including the following methods:

1. Autoregressive Integrated Moving Average (ARIMA): ARIMA model is a traditional time series forecasting method. The parameters of the ARIMA model are selected by stationary test, autocorrelation test and partial autocorrelation test.
2. Support vector machine (SVM): SVM splits the data by finding the most spaced supporting hyperplanes in the feature space.
3. Artificial neural network (ANN): We adopt a simple ANN model with only one hidden layer and update the weights through backpropagation.
4. LSTM: The standard LSTM model is used, and the state vector at the last moment is used as the output of the model.
5. Gate recurrent unit (GRU): We adopt GRU which is easier to train than LSTM, and the state vector at the last moment is used as the output of the model.

To verify the performance of the model, SL-ALSTM is compared with ARIMA, SVM, ANN, LSTM, GRU. During the experiment, in order to ensure the accuracy and objectivity of the experiment, each model was tested 10 times on the same training set and test set, and the average value was taken as the final result. Fig. 5 shows the numerical prediction of the loan default rate of each model.



(a) Prediction of loan default rates in 2013



(b) Predicts loan default rates in 2015

Figure 5 Prediction of loan default rates by various models

Tab. 1 shows the prediction accuracy of the model for the loan default rate from 2009 to the fourth quarter of 2018. The best results have been bolded.

Table 1 Comparison of prediction accuracy between SL-ALSTM and baseline models

	ARIMA	SVM	ANN	LSTM	GRU	SL-ALSTM
RMSE	0.111	0.106	0.077	0.058	0.058	**0.053**
MAE	0.102	0.092	0.063	0.052	0.048	**0.046**
DA	0.454	0.0545	0.634	0.636	0.729	**0.734**

From the table, the values of the three metrics of SL-ALSTM are 0.053, 0.046 and 0.734, respectively. Fig. 6 shows the performance of the SL-ALSTM model and baseline models on the three metrics. In the overall view, the RMSE value and MAE value of the SL-ALSTM model are smaller than baseline models. The DA value is 73% and higher than baseline models, which shows the relatively accurate prediction ability of the model in forecasting.

As can be seen from the table, deep learning models outperform ARIMA, SVM models. This is because the ALSTM model has high data requirements and requires time-series data to be stable, while the SVM model uses the traditional artificial feature extraction method that requires enough professional knowledge to select appropriate features.

Among the four deep learning models, the LSTM, GLU, and SL-ALSTM models outperform the ANN models. This shows that these models can find time-varying laws in the field of P2P online lending credit risk assessment through user online operation behavior data.

As seen from the table, compared with standard LSTM, the RMSE value of SL-ALSTM is reduced by 8.6% and the MAE value is reduced by 11.5%. Therefore, the performance of SL-ALSTM is better than the standard LSTM, indicating that paying attention to the contextual correlation of loan data and considering the importance of different sequence features will help improve the accuracy of prediction results.

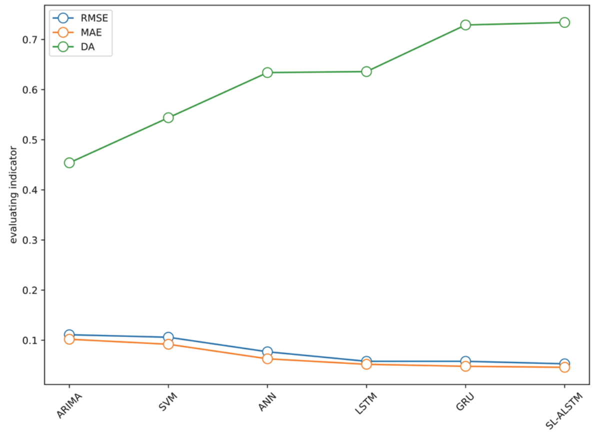


Figure 6 Comparison of the prediction accuracy of the SL-ALSTM model and baseline models

In summary, the experimental results show that the performance of our model is better than other baseline models. And the DA value of the SL-ALSTM model is around 70%, showing that the model in this paper has relatively accurate prediction ability.

5 CONCLUSION

In this study, we apply the SL-ALSTM model to predict the average default rate of new loans issued each month using data from the Lending Club P2P lending platform from 2009 to 2018. The conclusions are as follows: (1) Sliding window processing on the data can effectively utilize the contextual correlation of loan data; (2) Using the attention mechanism, the importance difference between loan features can be effectively utilized to improve the accuracy of prediction; (3) Comparing the SL-ALSTM model with ARIMA, SVM, ANN, LSTM, and GRU models, it is found that SL-ALSTM has significantly improved prediction accuracy and trend accuracy.

Given the powerful learning and generalization capabilities of LSTM-like models, their application in Internet finance will have a broader and broad prospect. In addition, this case enriches the application of deep learning models in P2P time series data, and provides useful references for further research by other scholars in this field.

Internet finance is constantly growing. In addition to the lending club dataset, there are data available from other platforms. The data provided by different platforms have different attributes. The periodicity and spatio-temporal correlation of time series data have been considered in the field of intelligent transportation system [30, 31] and medical imaging [32], and there have been a large number of deep learning application cases. In the future, we will further improve the deep learning framework to consider the periodicity and spatio-temporal correlation of internet financial data. And we also need to consider the interpretability of the model to enhance the reliability of the model prediction results.

Acknowledgements

This research is funded by the R&D Program of Beijing Municipal Education Commission (Grant No. KJZD20191000401). This research is also supported by Beijing Laboratory of National Economic Security Early-warning Engineering.

6 REFERENCES

[1] Luo, C., Li, M., Peng, P., & Fan, S. (2018). How Does Internet Finance Influence the Interest Rate? Evidence from Chinese Financial Markets. *Dutch Journal of Finance and Management*, *2*(1). https://doi.org/10.20897/djfm/89590

[2] Serrano-Cinca, C., Gutierrez-Nieto, B., & Lopez-Palacios, L. (2015). Determinants of Default in P2P Lending. *PLoS One*, *10*(10), e0139427. https://doi.org/10.1371/journal.pone.0139427

[3] Kouki, M. & Elkhaldi, A. (2011). Toward a predicting model of firm bankruptcy: evidence from the Tunisian context. *Middle Eastern Finance and Economics*, *14*, 26-43.

[4] Serrano-Cinca, C. & Gutiérrez-Nieto, B. (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction. *Decision Support Systems*, *54*(3), 1245-1255. https://doi.org/10.1016/j.dss.2012.11.015

[5] Chen, Y. (2017). Research on the credit risk assessment of chinese online peer-to-peer lending borrower on logistic regression model. *DEStech Transactions on Environment, Energy and Earth Science*, 216-221. https://doi.org/10.12783/dteees/apees2017/7689

[6] Chen, N., Ribeiro, B., & Chen, A. (2015). Financial credit risk assessment: a recent review. Artificial Intelligence Review, 45(1), 1-23. https://doi.org/10.1007/s10462-015-9434-x

[7] Xiao, M. (2022). Supervision Strategy Analysis on Price Discrimination of E-Commerce Company in the Context of Big Data Based on Four-Party Evolutionary Game. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/2900286

[8] Xiao, M. & Zhai, C. (2022.). Research on Subsidy Strategy of Shared Accommodation Platform under the Background of Big Data Based on Evolutionary Game. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/9066215

[9] Bai, B., Yen, J., & Yang, X. (2008). False financial statements: characteristics of China's listed companies and CART detecting approach. *International journal of information technology & decision making*, *7*(02), 339-359. https://doi.org/10.1142/S0219622008002958

[10] Xia, Y., Liu, C., & Liu, N. (2017). Cost-sensitive boosted tree for loan evaluation in peer-to-peer lending. *Electronic Commerce Research and Applications*, *24*, 30-49. https://doi.org/10.1016/j.elerap.2017.06.004

[11] Liu, W. A., Fan, H., & Xia, M. (2022). Multi-grained and multi-layered gradient boosting decision tree for credit scoring. *Applied Intelligence*, *52*(5), 5325-5341. https://doi.org/10.1007/s10489-021-02715-6

[12] Cui, D. (2015). Financial Credit Risk Warning Based on Big Data Analysis. *Metallurgical & Mining Industry*, (6).

[13] Malekipirbazari, M. & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, *42*(10), 4621-4631. https://doi.org/10.1016/j.eswa.2015.02.001

[14] Ma, X., Sha, J., Wang, D., Yu, Y., Yang, Q., & Niu, X. (2018). Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning. *Electronic Commerce Research and Applications*, *31*, 24-39. https://doi.org/10.1016/j.elerap.2018.08.002

[15] Wu, C.-F., Huang, S.-C., Chiou, C.-C., & Wang, Y.-M. (2021). A predictive intelligence system of credit scoring based on deep multiple kernel learning. *Applied Soft Computing*, *111*, 107668. https://doi.org/10.1016/j.asoc.2021.107668

[16] Xia, Y., Guo, X., Li, Y., He, L., & Chen, X. (2022). Deep learning meets decision trees: An application of a heterogeneous deep forest approach in credit scoring for online consumer lending. *Journal of Forecasting*. https://doi.org/10.1002/for.2891

[17] Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big Data: Deep Learning for financial sentiment analysis. *Journal of Big Data*, *5*(1). https://doi.org/10.1186/s40537-017-0111-6

[18] Kim, J. Y. & Cho, S. B. (2019). Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network. *Expert Systems*, *36*(4), 2019. https://doi.org/10.1111/exsy.12403

[19] Gunnarsson, B. R., Vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). "Deep learning for credit scoring: Do or don't? *European Journal of Operational Research*, *295*(1), 292-305. https://doi.org/10.1016/j.ejor.2021.03.006

[20] Wang, Y. & Ni, X. S. (2020). Risk Prediction of Peer-to-Peer Lending Market by a LSTM Model with Macroeconomic Factor. *Proceedings of the 2020 ACM Southeast Conference*. https://doi.org/10.1145/3374135.3385287

[21] Sarı, T., Güleş, H., & Yiğitol, B. (2020). Awareness and readiness of Industry 4.0: The case of Turkish manufacturing industry. *Advances in Production Engineering & Management*, *15*(1), 57-68. https://doi.org/10.14743/apem2020.1.349

[22] Tian, W. & Zhang, H. (2021). A dynamic job-shop scheduling model based on deep learning. *Advances in Production Engineering & Management*, *16*(1). https://doi.org/10.14743/apem2021.1.382

[23] Ren, J., Ye, C., & Yang, F. (2020). A novel solution to JSPS based on long short term memory and policy gradient algorithm. *International Journal of Simulation Modelling*, *19*(1), 157-168. https://doi.org/10.2507/IJSIMM19-1-CO4

[24] Vukićević, A., Mladineo, M., Banduka, N., & Mačužić, I. (2021). A smart Warehouse 4.0 approach for the pallet management using machine vision and Internet of Things (IoT): A real industrial case study. *Advances in Production Engineering & Management*, *16*(3). https://doi.org/10.14743/apem2021.3.401

[25] Wu, P. & Yang, D. (2021). E-Commerce Workshop Scheduling Based on Deep Learning and Genetic Algorithm. *International Journal of Simulation Modelling (IJSIMM)*, *20*(1). https://doi.org/10.2507/IJSIMM20-1-CO4

[26] Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. *2015 IEEE international conference on big data (big data)*, *2015: IEEE*, 2823-2824. https://doi.org/10.1109/BigData.2015.7364089

[27] Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, *519*, 127-139. https://doi.org/10.1016/j.physa.2018.11.061

[28] Molina, L. E., Bhulai, S., Reader, V., de Jeu, R., & Tjepkema, J. (2018). *Understanding user behavior in e-commerce with Long Short-term Memory (LSTM) and Autoencoders*. Masters thesis, Vrije Universiteit Amsterdam.

[29] Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48-62, 2021. https://doi.org/10.1016/j.neucom.2021.03.091

[30] Zhang, D., Peng, Y., Zhang, Y., Wu, D., Wang, H., & Zhang, H. (2021). Train time delay prediction for high-speed train dispatching based on spatio-temporal graph convolutional network. *IEEE Transactions on Intelligent Transportation Systems*, *23*(3), 2434-2444. https://doi.org/10.1109/TITS.2021.3097064

[31] Zhang D. et al. (2022). An Interpretable Station Delay Prediction Model Based on Graph Community Neural Network and Time-Series Fuzzy Decision Tree. *IEEE Transactions on Fuzzy Systems*. https://doi.org/10.1109/TFUZZ.2022.3181453

[32] Li, X., Wang, L., Chen, Q., Hu, Y., Du, J., Chen, X., Zheng, W., Lu, J., & Chen, N. (2020). The reorganization of insular subregions in individuals with below-level neuropathic pain following incomplete spinal cord injury. *Neural Plasticity*. https://doi.org/10.1155/2020/2796571

Contact information:

Menggang LI
1) National Academy of Economic Security, Beijing Jiaotong University,
Beijing 100044, China
2) Beijing Laboratory of National Economic Security Early-warning Engineering,
Beijing Jiaotong University, Beijing 100044, China
3) Beijing Center for Industrial Security and Development Research,
Beijing Jiaotong University, Beijing 100044, China
E-mail: mgli1@bjtu.edu.cn

Zixuan ZHANG
Business School, The University of Hong Kong, Hong Kong
E-mail: u3589687@connect.hku.hk

Ming LU
School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, China
E-mail: 21120526@bjtu.edu.cn

Xiaojun JIA
(Corresponding author)
1) National Academy of Economic Security, Beijing Jiaotong University,
Beijing 100044, China
2) Beijing Center for Industrial Security and Development Research,
Beijing Jiaotong University, Beijing 100044, China
E-mail: xjjia@bjtu.edu.cn

Rui LIU
Graduate School, Lyceum of the Philippines University Manila,
Manila 1002, Philippines
E-mail: lui.rui@lpunetwork.edu.ph

Xuan ZHOU
1) National Academy of Economic Security, Beijing Jiaotong University,
Beijing 100044, China
2) Beijing Laboratory of National Economic Security Early-warning Engineering,
Beijing Jiaotong University, Beijing 100044, China
E-mail: xuanzhou@bjtu.edu.cn

Yingjie ZHANG
School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, China
E-mail: zhangyingjie@bjtu.edu.cn