Bulletin of Electrical Engineering and Informatics
Vol. 14, No. 3, June 2025, pp. 2004~2016
ISSN: 2302-9285, DOI: 10.11591/eei.v14i3.8653

BiLSTM OptiFlow: an enhanced LSTM model for cooperative financial health forecasting

Evi Maria[1], Teguh Wahyono[2], Kristoko Dwi Hartomo[2], Purwanto[1], Christian Arthur[3]
[1]Department of Information System, Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
[2]Department of Informatics Engineering, Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
[3]School of Transdisciplinary, University of Technology Sydney, Ultimo, Australia

Article Info

Article history:

Received May 1, 2024
Revised Nov 10, 2024
Accepted Nov 19, 2024

Keywords:

Cooperative
Deep learning
Financial health
Financial ratio
Forecasting

ABSTRACT

This paper presents bidirectional long short-term memory (BiLSTM) OptiFlow, an optimized deep learning model designed to predict the financial health of cooperatives using key financial ratios: debt to equity ratio (DER), net profit margin (NPM), and return on equity (ROE). By leveraging a BiLSTM architecture combined with an optimal decayed learning rate, this model aims to enhance forecasting accuracy. The proposed model was tested against three established methods—recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU)—and evaluated using mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared error (MSE) metrics. Results indicate that BiLSTM OptiFlow outperforms the other models across all key indicators. This research offers a robust approach to cooperative financial forecasting, with significant implications for decision-making processes in cooperative management.

Corresponding Author:

Teguh Wahyono
Department of Informatics Engineering, Faculty of Information Technology
Satya Wacana Christian University
Diponegoro Street No.52-60, Sidorejo, Salatiga City, Central Java, Indonesia
Email: teguh.wahyono@uksw.edu

1. INTRODUCTION

Cooperatives are a unique business model because they serve economic goals and social purposes. Cooperatives are not solely driven by profit but also prioritize cooperation, social justice, and the welfare of their members [1]. A cooperative allows its members to control the business since the venture has no shareholders. Consequently, cooperative activities' economic and social benefits remain as large as possible in the community where the cooperative was established [2]. Trust is social capital that is believed to increase the participation of cooperative members. In 2016, there were 3 million cooperatives around the world that experienced rapid growth [3]. The world's 300 largest cooperatives collectively reported a turnover of US$2.018 trillion. This turnover reflects that cooperatives contribute to a country's economic growth [4].

In its development, cooperatives experience various challenges ranging from managerial problems to capital problems [5]. Cooperative health is the focus of attention of both the government and academia because of its strategic role as an essential pillar in a country's economy. This condition encourages research on cooperative health predictions to be essential. Research in accounting and finance has found empirical evidence of the effectiveness of financial ratios in predicting the health condition of organizations both in the private sector as suggested in the previous research [6]-[11]. In line with that, there are similar studies in public sectors [12]-[15]. There are also studies investigating financial ratios in predicting the health condition of cooperatives [16], [17]. These studies reflect that financial ratios could serve as tools for management and financial report users such as investors, creditors, employees, and governments to better understand business entities' business health condition.

Cooperative financial performance refers to the extent to which the cooperative can generate profits, manage debt and equity, and meet financial obligations promptly. The health of the cooperative can be seen from the financial aspect by looking at the value of debt to equity ratio (DER), net profit margin (NPM), and return on equity (ROE). DER is one of the indicators to measure the level of solvency, while NPM and ROE measure profitability. High DER indicates that the cooperative has a high financial risk because it has an enormous debt to pay creditors. In contrast, low DER indicates that the cooperative is more stable because its business activities obtain funding from the capital of its members. Meanwhile, the high NPM and ROE ratio value means that cooperatives are more efficient and profitable, and vice versa [16].

This research addresses the challenge of predicting DER, NPM, and ROE values to assess cooperative health, leveraging the bidirectional long short-term memory (BiLSTM) approach tailored for the complex patterns of cooperative financial data. Unlike public entities, cooperatives do not publicize their financial statements, only reporting to members at the annual member meeting. The study introduces the BiLSTM OptiFlow model, which combines BiLSTM architecture with an optimal decayed learning rate, and compares its performance against other deep learning methods like recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent units (GRU). Utilizing data from Indonesian cooperatives from 2000-2021, sourced from the Indonesian Ministry of Cooperatives and SMEs, the research seeks to determine the most effective deep-learning method for forecasting financial health in cooperatives.

2. LITERATURE REVIEW

Previous studies have addressed financial forecasting problems using various statistical approaches. Few used seasonal autoregressive integrated moving average (SARIMA) and autoregressive integrated moving average (ARIMA) to predict stocks [18], [19]. The ARIMA model is explicitly used for short-range predictions and has weaknesses because it is challenging to confirm long-term investment performance. Moreover, financial data is volatile and influenced by government policies, global economic conditions, and other disasters. Research by Devi *et al.* [20] found that those who studied forecasting with ARIMA got an accuracy of up to 38%. The next model that still uses statistics is the Bayesian model, often used as a comparison model with the latest artificial neural network (ANN) developments. The study [21] used Bayesian models to predict and yield an accuracy of up to 78%. Research with statistical approach models that have been mentioned often produces incorrect predictions caused by underfitting and overfitting conditions, so the model needs to be more reliable to predict financial problems.

Since the statistics approach models often get inaccurate results, ANN models are often used for forecasting [22]. Lakshminarayanan *et al.* [23] used the Elliot wave indicator and technical analysis alone to predict 5 Stocks. The results produced with ANN are more accurate than statistical models, which is 93.83%. These findings suggest tuning the inputs and models may also improve model performance. Many types of architectures and tuning were developed to improve the performance of this model, such as convolutional neural network (CNN), LSTM, and GRU [24]-[26]. Cao and Wang [26] proposed a new model with a combination of CNN+SVM to predict stock indexes from various countries. They focus on tuning the model on historical data input with down sampling and produce a model that maximizes its ability even though it does not include the number of prediction results.

The LSTM and GRU models proposed by [24] forecast CNPC's stock price. Results from both models show that LSTM has better capabilities with lower mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) values than the GRU model. Although LSTM is better than other models, researchers revealed that GRU has a faster prediction time than LSTM due to the number of internal gating and fewer parameters. Another form of development of LSTM is BiLSTM, which is a sequence processing model consisting of 2 LSTMs that function to receive forward and backward inputs. The strength of this model is that the model better understands the context from the past and the future so that the prediction results can be more relevant [27], [28]. Bi-LSTM is widely used for forecasting, one of which is research on stock price forecasting from Bhutan tourism corporation limited (BTCL), which uses a comparison of LSTM, ConvLSTM, and GRU models [29]. That research reveals that with a small amount of model tuning, Bi-LSTM performs best with an R-squared value of 95.93%.

The OptiFlow method optimizes the learning rate by using a decay strategy. This approach gradually reduces the learning rate over time to avoid overshooting the global minimum while ensuring faster convergence. The process begins with a large initial learning rate (e.g., 0.1) and applies the decay function, reducing the rate based on changes in the loss function. Improper learning rates can prevent model training from being slow or difficult to converge. In some cases, a learning rate that is too high can cause the model to jump to the global minimum, while a very low learning rate can keep the model stuck at the local minimum [30]. A study experimented with adjusting the learning rate on VGG and RESNET models for detecting the CIFAR-10 photo data classification [31]. They found that tuning the learning rate with the decay learning rate technique could help achieve faster convergence and better results in the task.

3. METHOD

The study introduces BiLSTM OptiFlow, a model that uniquely combines a LSTM bidirectional architecture with an optimal learning rate search method featuring a decayed rate. This approach focuses on identifying the optimal rate by tracking the lowest loss value during training, a significant departure from conventional fixed-rate training. Figure 1 illustrates the process flow, starting from data transformation of a feature-rich dataset into training and testing divisions. The training phase aims to determine and apply the best learning rate while training the model, which is then evaluated using the testing data.

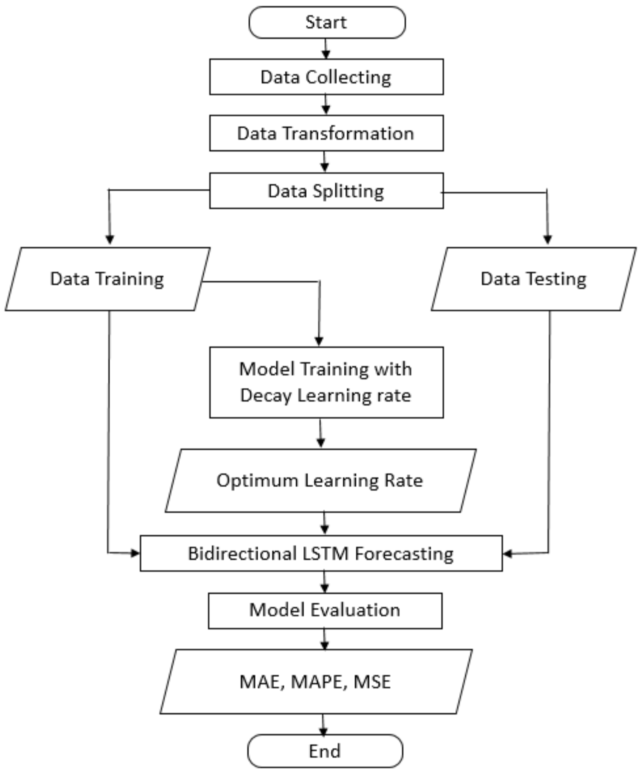

Figure 1. Proposed method

3.1. Data collecting

This study used official data from the Ministry of Cooperatives and SMEs of the Republic of Indonesia in charge of cooperatives and small and medium enterprises. Data is available on https://kemenkopukm.go.id/data-koperasi website; the available data is in the form of an aggregate table of each province in Indonesia from year to year from 2000 to 2021.

3.2. Data transformation

After collecting raw data from each province in Indonesia, a transformation process was undertaken to analyze the health performance of cooperatives using three key indicators. The first, DER, measures the total debt, assumed here as external capital, against own capital. The second, NPM, indicates how much net income is generated for every Rp. 100 in sales. The third, ROE, evaluates how effectively cooperatives use their capital to generate net income. These indicators help members understand their strategic role in the cooperative's development and assess its management.

$$DER = \frac{Total\ External\ Capital}{Equity\ Capital} x\ 100 \quad (1)$$

$$NPM = \frac{Net\ Income}{Business\ Volume} x\ 100 \quad (2)$$

$$ROE = \frac{Net\ Income}{Equity\ Capital}\ x\ 100 \quad (3)$$

3.3. Data splitting

The data splitting process is critical in model building and model performance evaluation in data science and machine learning. The goal is to divide the dataset into two subsets: training and testing data. Larger datasets are used to train the model, while smaller datasets are used to test how much the built model can generalize to never-before-seen data. Data splitting is essential to avoid the problem of overfitting, where the model might memorize the training data without understanding the underlying pattern. Separating the test data can objectively measure the model's performance and identify whether the model can generalize well.

3.4. Model training with decayed learning rate

Model training with a decayed learning rate is used to find the best learning rate in training models in data science and machine learning. The main goal is to determine the optimal learning rate so that the model can achieve convergence quickly without causing divergence or stagnation in the training process. This process starts by initializing the initial learning rate at an enormous value, for example, 10e-1. Then, this learning rate is gradually lowered, considering specific decay schemes, such as reducing the learning rate every few epochs or based on changes in the loss function. These steps help explore different learning rates and find the most appropriate value for the trained model. In the early stages of training, a small learning rate allows the model to understand patterns in the data.

The Figure 2 shows that the learning model (epoch increases) over time, and the learning rate increases to optimize the training process and achieve faster convergence. However, if the learning rate is too large, the model may trip or fail to reach the global minimum. Therefore, decay measures help prevent this problem. One thing to note is that the picture above is significant. Nevertheless, the results may not show the learning rate starting with the number 1 because it is influenced by the optimizer used.

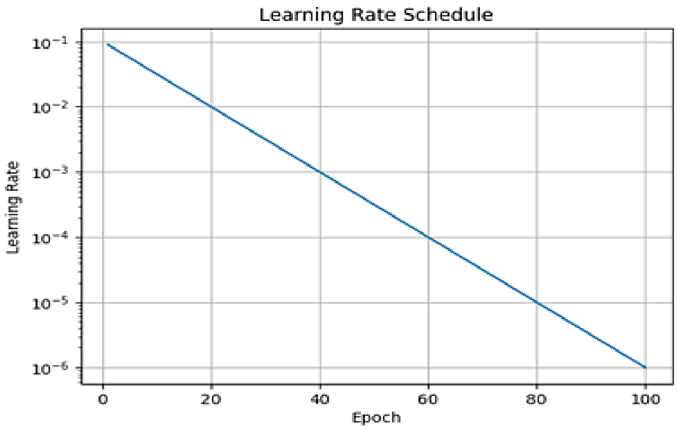

Figure 2. Learning rate decay strategy

3.5. Bidirectional long short-term memory forecasting model

BiLSTM is a recurrence neural network model used in various applications, including time series forecasting. LSTM was developed to address the problem of long-term dependency in time sequence data, where vital information from the past can be blurred or lost in ordinary neural networks. Using a gate mechanism, LSTM allows the model to remember long-term and short-term information, making it suitable for analyzing time series data with complex temporal relationships.

The neural network architecture illustrated in Figure 3 is designed to effectively capture both temporal and spatial features of data for predictive analysis. It begins with a Lambda layer that processes data without adjustable parameters, maintaining the integrity of the input. Following this is a bidirectional layer, consisting of sub-layers with outputs of (None, 2, 64) and (None, 128), enabling learning from data sequences in both directions to better understand contextual relationships. Subsequently, two dense layers output shapes of (None, 16) and (None, 1), pivotal for connecting neurons and formulating predictions by integrating previously learned features. The model concludes with another parameter-free Lambda layer, mirroring the first to finalize data processing, ensuring each component efficiently contributes to the model's predictive capabilities as depicted in Figure 3.

```
Layer (type)                 Output Shape              Param #
===============================================================
lambda (Lambda)              (None, 2, 1)              0

bidirectional (Bidirectiona  (None, 2, 64)             8704
l)

bidirectional_1 (Bidirectio  (None, 128)               66048
nal)

dense (Dense)                (None, 16)                2064

dense_1 (Dense)              (None, 1)                 17

lambda_1 (Lambda)            (None, 1)                 0

===============================================================
```

Figure 3. Stacked Bi-LSTM model

3.6. Model evaluation

The process of evaluating forecasting models using the MAE, mean absolute percentage error (MAPE), and MSE metrics is a critical step to understanding how well the model performs predictions and measures the quality of prediction results. This evaluation compares the model's predicted value with the observed data's actual value. Those metrics help identify how our model can make accurate estimates in line with the original data.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y| \quad (4)$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}|\frac{y_i - y}{y_i}| \quad (5)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - y)^2 \quad (6)$$

where, y_i is actual value, y is predicted value, and N is data size.

4. RESULTS AND DISCUSSION

4.1. Data collecting

By downloading various cooperative recapitulation files in Indonesia from 2000–2021 from the web https://kemenkopukm.go.id/data-koperasi, researchers get 11 features aggregated per province. The study focuses only on transforming the data in (1)-(3); therefore, features other than those used in the formula will be removed. The results of this process can be seen in Table 1, which displays cooperative recapitulation data in Indonesia with examples in 2021 in million rupiah.

Table 1. Indonesian cooperative recapitulation data for 2021

No	Province	Equity capital	External capital	Business volume	Net income
1	Aceh	1.401.963,26	199.155,16	1.564.088,47	242.810,03
2	North Sumatra	6.423.080,40	1.804.598,89	6.269.050,03	387.281,66
3	West Sumatra	2.816.986,32	1.524.268,57	4.008.031,30	209.190,14
4	Riau	1.841.508,85	1.478.263,54	2.645.162,41	177.276,19
...	...		...	...	...
33	Papua	246.542,21	281.523,22	237.775,70	17.384,49
34	West Papua	44.459,70	281.523,22	40.883,11	6.682,22

4.2. Data transformation

After getting raw data about the recapitulation of cooperatives in Indonesia, the data will be converted into several indicators so that information when forecasting becomes insightful. The features, as shown in Table 1, equity capital, external capital, business volume, and net income, will be included in (1)-(3). After entering (1)-(3), DER, NPM, and ROE indicators per province will be obtained. The results of this transformation can be seen in Table 2.

Table 2. Get DER, NPM, and ROE from 2021 data

No	Province	DER	NPM	ROE
1	Aceh	14,20544787	15,52405984	17,31928624
2	North Sumatra	28,09553637	6,177677	6,029531562
3	West Sumatra	54,10990317	5,219274	7,426026123
4	Riau	80,27458244	6,701901907	9,626681403
5	Jambi	64,19308581	4,666580815	14,36266699
...	...	...	...	...
33	Papua	11,10201373	7,311298001	7,05132399
34	West Papua	47,55468435	16,34469589	15,02983601

After transforming Table 2 to all years, it is time to aggregate the three indicators, i.e., DER, NPM, and ROE per year. After aggregating these three indicators per year, time-series data from DER, NPM, and ROE from 2000 to 2021 are ready for forecasting. Table 3 shows the time series data used for the time series.

Table 3. DER, NPM, and ROE from 2000 to 2021

Year	DER	NPM	ROE
2000	182.976312	3.003612	10.187870
2001	139.509966	8.093034	26.790250
2002	172.412072	3.478805	11.536596
2003	158.592808	5.908170	19.871860
...	...	...	...
2020	114.094680	4.151563	9.110025
2021	116.106779	3.936990	7.837024

4.3. Data splitting

The data in Table 3 will be divided into 2: training and testing data. The training data will take 2000 to 2016 to find the optimal learning rate and train the model. Data testing on this model ranges from 2017 to 2021. Testing data that the model has never seen will then be used to evaluate the model.

4.4. Model training with decayed learning rate

This process will find the best learning rate based on the learning rate scheme seen in Figure 2. The BiLSTM model will conduct training with training data from 2000-2016 in Table 3, carried out in as many as 100 epochs so that the learning rate with a minor loss will be considered the best. Figure 4 shows the search for learning rate using the decayed learning rate of the DER indicator. The training process showed that the learning rate from 10-2 to 10-4 enabled the model to learn data patterns well with a constant loss of less than 50. The learning rate after 10-4 to 10-6 shows a loss that increases exponentially until the loss value reaches almost 125. This high loss value was caused by a learning rate that was too small. It showed that the learning model learned too slowly, and the predicted value needed closer to the actual value. This feature's most optimal learning rate value was 0.0025, with a loss value of 9.46.

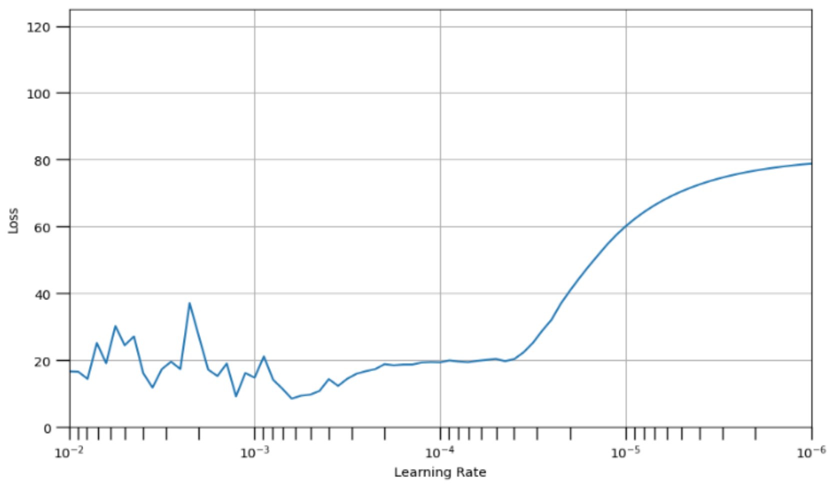

Figure 4. Decayed learning rate of the DER indicator

The search for learning rate using decayed learning rate is shown in Figure 5, where the model will learn with a decreased learning rate. The difference from the learning rate in the DER feature is that the smaller loss range is between 0-10. The smaller loss range occurs because the range of NPM features is also tiny compared to DER. Learning rates at 10-3 to 10-5 have a stable loss value and the possibility of an optimal learning rate in this range. In the following range, 10-5 to 10-6, loss values keep increasing and are unstable. The increase in the DER indicator is due to the learning rate value being too small, so it cannot find the value. It even gets stuck in the local minima, so the prediction must still be closer to the original value. Based on the graph, it is known that the most optimal learning rate is 0.000178, with a loss value of 0.4954.

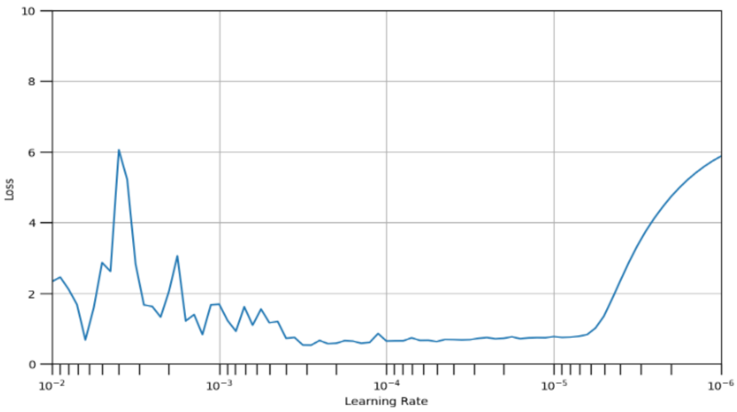

Figure 5. Decayed learning rate of NPM indicator

Figure 6 shows the training results with a decayed learning rate from the ROA feature. Almost the same as the NPM feature, the resulting loss value range is higher than the DER feature because the ROA feature range is higher than DER. However, there is a difference between the two previous features; they are unstable loss values from the largest to the most minor learning. The unstable loss occurs due to the batch values that do not match the data, or in other words, they are inconsistent data patterns. This process produces the most optimal learning rate value for the ROA indicator, which is 0.00013 with a loss value 2.3786.

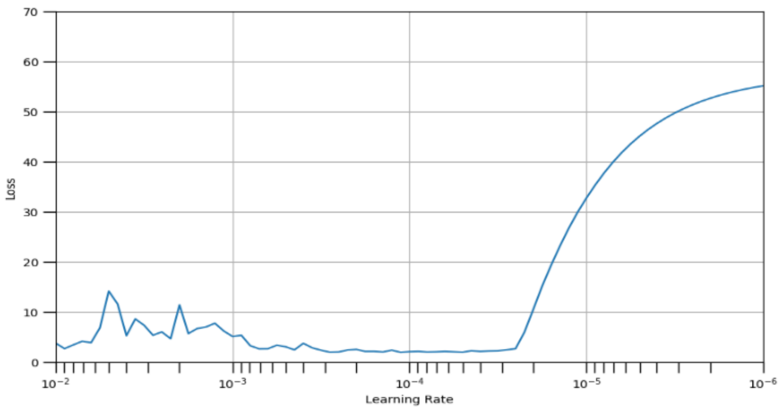

Figure 6. Decayed learning rate of ROA indicator

4.5. Bidirectional long short-term memory forecasting

This process involves using LSTM methods with a bidirectional approach, allowing the model to process data forward and backward simultaneously. Previously, the model has searched for optimal and decayed learning rates to improve model performance. Through this experiment, accurate prediction results can provide valuable insights for stakeholders in making strategic decisions to improve cooperative performance in the future.

Figure 7 is the result of forecasting the DER indicator with the proposed model compared to other RNN models. In the training phase (2002-2016), it was seen that the proposed model, LSTM, and GRU had prediction results closer to the original value compared to RNN, whose prediction results increasingly reached the original value. In the testing phase (2017-2021), the proposed model performed better than the other three algorithms. The RNN algorithm does predict closer values, but the model underfits so that the resulting predictions are only like horizontal lines. The LSRM and GRU models experience overfitting, as shown in the training phase. The results are always almost equal to the original value, but in the testing data, the performance decreased because the prediction was far from the actual value.

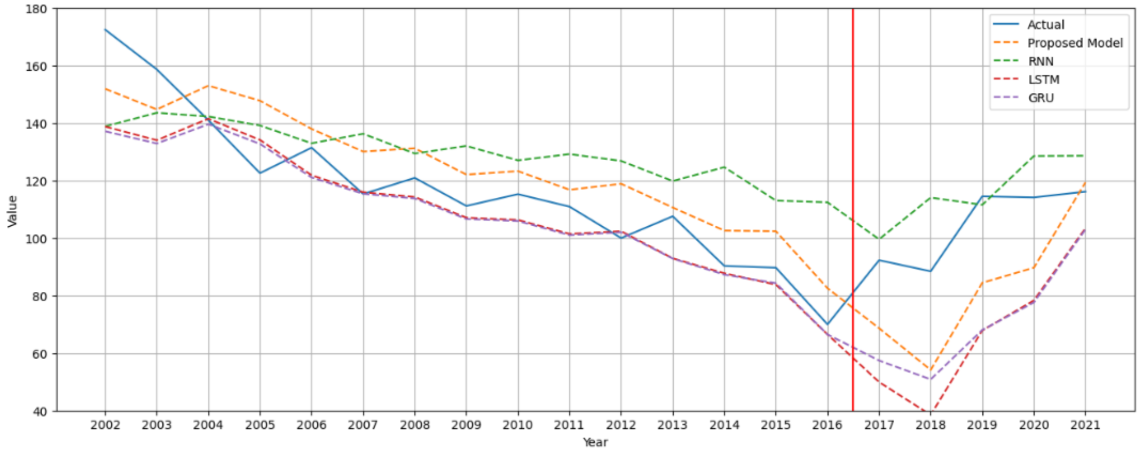

Figure 7. DER forecasting

Figure 8 shows the predicted NPM indicator value; the prediction results show that all models show underfitting. The underfitting occurs because of data that has no pattern or is too fluctuating and too little data. Although all models do not fit well, the proposed model still can predict the NPM value well, as evidenced by specific years, i.e., 2003, 2008, 2014, and 2016, closer to the original value than the other three models.

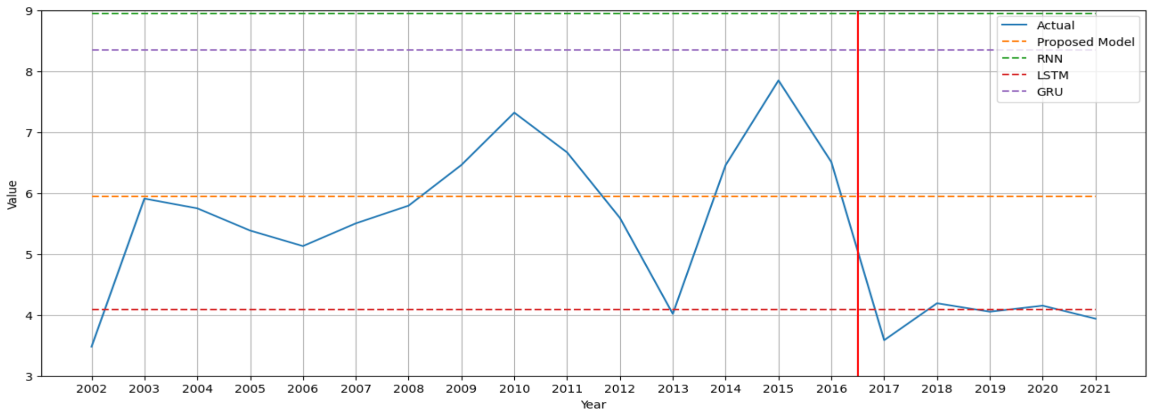

Figure 8. NPM forecasting

The predicted value of the ROA indicator with the proposed model compared to 3 other models is shown in Figure 9. At this value, the proposed model, LSTM, and GRU show more or less the same performance, so the exact results of which model is the best will be carried out in the following process: model evaluation. The RNN model seems unable to understand data patterns, with predicted results tending to be a straight line from training data to testing data.

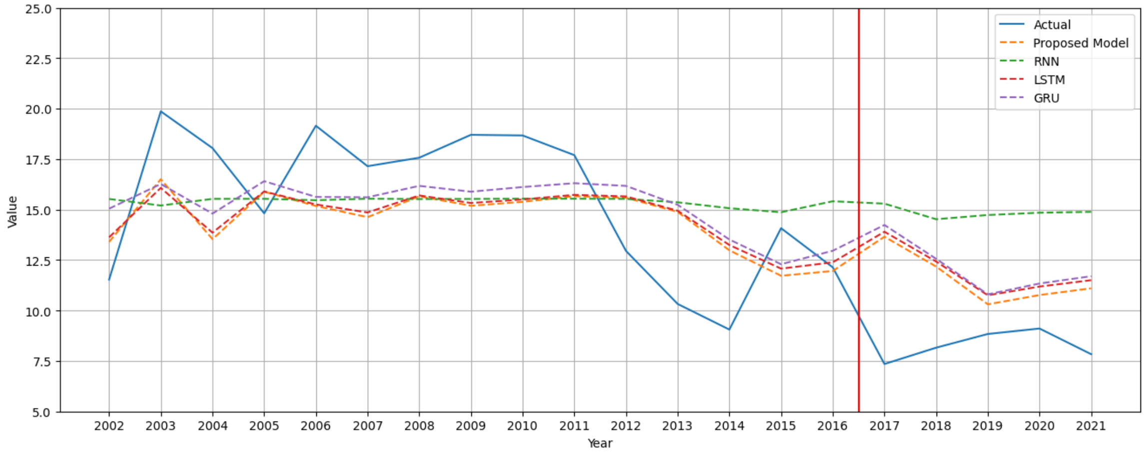

Figure 9. ROA forecasting

4.6. Model evaluation

This process involves using LSTM methods with a bidirectional approach, allowing the model to process data forward and backward simultaneously. Previously, the model has searched for optimal and decayed learning rates to improve model performance. Through this experiment, accurate prediction results can provide valuable insights for stakeholders in making strategic decisions to improve cooperative performance in the future.

This process will produce the previous prediction process's MAE, MAPE, and MSE values. The results of the predictions in Figures 7-9 show that visualized performance differs depending on data patterns and model capabilities. The process uses (4)-(6) to determine the best model for each indicator.

Table 4 shows the results of forecasting evaluation of DER indicator values that illustrate the performance comparison of four different models, proposed model, RNN, LSTM, and GRU, in three main evaluation matrices: MAE, MAPE, and MSE. The evaluation results show that the proposed model has the lowest MAE of 15.155, indicating that the average prediction error between the actual value and the value predicted by the model is relatively lower than other models. In addition, while the proposed model's MAPE of 14.019% is not the lowest, it is comparable to other models, with the GRU model showing a slightly better MAPE of 13.758%. Regarding MSE, the proposed model also showed better performance with a value of 17.405, indicating that the error spread between the prediction and the actual value is relatively lower than other models.

Table 4. Evaluation metrics of the DER indicator

Type	MAE	MAPE	MSE
Proposed model	**15.155**	14.019	**17.405**
RNN	17.512	17.047	20.648
LSTM	16.303	14.697	22.775
GRU	15.555	**13.758**	21.144

Although there are variations in performance between the LSTM, GRU, and proposed model, the proposed model consistently performs better in forecasting the DER indicator based on the MAE and MSE evaluations. However, the proposed model's performance in terms of MAPE is comparable to GRU, which shows a slightly better result. Overall, the evaluation results suggest that the proposed model is a reliable choice for predicting the value of the DER indicator.

Table 5 shows the evaluation results of four forecasting models—proposed model, RNN, LSTM, and GRU—against NPM indicator values. While the proposed model performs well, it does not consistently outperform all other models in every metric. The proposed model has the lowest MAE of 1.116, indicating that its predictions have the smallest average error compared to the other models. However, in terms of MAPE, the proposed model's value of 24.637% is slightly higher than that of the LSTM model, which has the lowest MAPE at 23.472%, indicating a smaller prediction error percentage for LSTM.

In terms of MSE, the proposed model shows a favorable result with a value of 1.368, although LSTM follows closely with a value of 1.802. The RNN and GRU models demonstrate lower performance across all metrics, with notably higher prediction errors. Despite LSTM showing competitive results, particularly in MAPE, the proposed model still shows strong performance overall, making it a reliable option for forecasting the NPM indicator, though LSTM could also be considered depending on the specific metric of interest.

Table 5. Evaluation metrics of NPM indicator

Type	MAE	MAPE	MSE
Proposed model	**1.116**	24.637	**1.368**
RNN	3.561	75.828	3.774
LSTM	1.435	**23.472**	1.802
GRU	2.966	64.135	3.218

Table 6 presents a performance comparison of four models—proposed model, RNN, LSTM, and GRU—using evaluation metrics MAE, MAPE, and MSE. The proposed model outperforms the others with a MAE of 2.922, signifying a lower average error in predicted versus actual values, and a MAPE of 24.861%, indicating fewer percentage errors compared to the RNN and LSTM models. It also maintains a competitive MSE of 3.238, showing a lower variance in prediction errors. While the LSTM and GRU models demonstrate competitive performance in certain metrics, the proposed model consistently leads in MAE and MAPE, establishing it as the most reliable for predicting ROE values among the models tested.

Table 6. Evaluation metrics of the ROA indicator

Type	MAE	MAPE	MSE
Proposed model	**2.922**	**24.861**	**3.238**
RNN	3.920	36.830	4.417
LSTM	2.999	25.960	3.314
GRU	2.986	26.704	3.320

In various experiments that have been conducted, it is seen that the performance of the proposed model consistently outperforms other models in Cooperative Financial Health Forecasting. The MAE results show that the model has a relatively small error in predicted values versus actual values. The MAPE value also shows that the average prediction error of the model is still acceptable. Meanwhile, the MSE results show that the model is quite consistent in the predictions produced without any major fluctuations. Thus, the proposed model offers a powerful approach to cooperative financial forecasting, with significant implications for the decision-making process in cooperative management.

5. CONCLUSION

The study benchmarks the Bi-LSTM model against other models, including RNN, standard LSTM, and GRU, using estimation metrics such as MAE, MAPE, and MSE. The results clearly demonstrate that the Bi-LSTM model consistently outperforms the others, achieving the lowest values for MAE, MAPE, and MSE. This superior performance underscores the model's ability to accurately predict the financial health of cooperatives, ensuring their viability and sustainability in the long run. Moreover, it would be valuable to explore the applicability of the BiLSTM OptiFlow model across other financial sectors, while also considering alternative optimization techniques that utilize adaptive learning rate schedules based on Hessian or gradient magnitudes. The data used in this study was drawn from Indonesian cooperatives spanning the years 2000 to 2021. Although these findings are directly relevant to this dataset, future research should investigate the model's potential to generalize to cooperatives in other regions or countries, particularly those with differing financial structures and reporting practices.

ACKNOWLEDGEMENTS

We thank the Ministry of Education, Culture, Research, and Technology Republic Indonesia for funding this research through the Leading Applied Research in Higher Education Program under grant number 198/SPK/D.D4/PPK.01.APTV/VI/2023 (21 June 2023), 001/LL6/PPM-V/AL.04/2023 (22 June 2023), and 054/SPK-PTUPT/RIK/6/2023 (23 June 2023).

FUNDING INFORMATION

This research was funded by Ministry of Education, Culture, Research, and Technology Republic Indonesia through the Leading Applied Research in Higher Education Program under grant number 198/SPK/D.D4/PPK.01.APTV/VI/2023 (21 June 2023), 001/LL6/PPM-V/AL.04/2023 (22 June 2023), and 054/SPK-PTUPT/RIK/6/2023 (23 June 2023).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Evi Maria	✓	✓			✓	✓		✓	✓	✓		✓	✓	
Teguh Wahyono	✓	✓				✓		✓	✓	✓		✓	✓	✓
Kristoko Dwi Hartomo	✓		✓	✓			✓			✓	✓		✓	✓
Purwanto	✓				✓	✓	✓	✓		✓	✓			
Christian Arthur	✓		✓	✓	✓	✓	✓	✓		✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The conducted research is not related to either human or animals use.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] T. Mazzarol, R. A. Simmons, and E. A. M. Limnios, "A Conceptual Framework for Research into Co-Operative Enterprise," *SSRN Electronic Journal*, pp. 22–50, 2012, doi: 10.2139/ssrn.2015641.
[2] M. L. Lode, T. Coosemans, and L. R. Camargo, "Is social cohesion decisive for energy cooperatives existence? A quantitative analysis," *Environmental Innovation and Societal Transitions*, vol. 43, pp. 173–199, Jun. 2022, doi: 10.1016/j.eist.2022.04.002.
[3] D. Gambetta, *Trust: Making and Breaking Cooperative Relations*. B. Blackwell, 1988.
[4] C. Karakas, *Cooperatives: Characteristics, activities, status, challenges*. 2019.
[5] M. Sánchez-Robles, J. R. Saura, and D. Ribeiro-Soriano, "Overcoming the challenges of cooperative startups businesses: insights from a bibliometric network analysis," *Review of Managerial Science*, vol. 18, no. 6, pp. 1617–1648, Jun. 2024, doi: 10.1007/s11846-023-00670-9.
[6] D. Liang, C. C. Lu, C. F. Tsai, and G. A. Shih, "Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study," *European Journal of Operational Research*, vol. 252, no. 2, pp. 561–572, Jul. 2016, doi: 10.1016/j.ejor.2016.01.012.
[7] H. Waqas and R. Md-Rus, "Predicting financial distress: Importance of accounting and firm-specific market variables for Pakistan's listed firms," *Cogent Economics and Finance*, vol. 6, no. 1, pp. 1–16, Jan. 2018, doi: 10.1080/23322039.2018.1545739.
[8] S. A. Balasubramanian, G. S. Radhakrishna, P. Sridevi, and T. Natarajan, "Modeling corporate financial distress using financial and non-financial variables: The case of Indian listed companies," *International Journal of Law and Management*, vol. 61, no. 3–4, pp. 457–484, Oct. 2019, doi: 10.1108/IJLMA-04-2018-0078.
[9] M. Karas and M. Reznakova, "Cash flows indicators in the prediction of financial distress," *Engineering Economics*, vol. 31, no. 5, pp. 525–535, Dec. 2020, doi: 10.5755/j01.ee.31.5.25202.

[10] Y. N. Isayas, “Financial distress and its determinants: Evidence from insurance companies in Ethiopia,” *Cogent Business and Management*, vol. 8, no. 1, pp. 1–6, Jan. 2021, doi: 10.1080/23311975.2021.1951110.
[11] M. Ugur, E. Solomon, and A. Zeynalov, “Leverage, Competition, and Financial Distress Hazard: Non-Monotonic Effects in the Presence of Agency Costs,” *Journal of Banking and Finance*, pp. 1–54, 2020.
[12] I. M. García-Sánchez, B. Cuadrado-Ballesteros, J. V. Frías-Aceituno, and N. Mordan, “A New Predictor of Local Financial Distress,” *International Journal of Public Administration*, vol. 35, no. 11, pp. 739–748, Sep. 2012, doi: 10.1080/01900692.2012.679173.
[13] J. M. Trussel and P. A. Patrick, “Predicting fiscal distress in special district governments,” *Journal of Public Budgeting, Accounting and Financial Management*, vol. 25, no. 4, pp. 589–616, 2013, doi: 10.1108/JPBAFM-25-04-2013-B001.
[14] G. Turley, G. Robbins, and S. McNena, “A Framework to Measure the Financial Performance of Local Governments,” *Local Government Studies*, vol. 41, no. 3, pp. 401–420, 2015, doi: 10.1080/03003930.2014.991865.
[15] E. Gorina, C. Maher, and M. Joffe, “Local Fiscal Distress: Measurement and Prediction,” *Public Budgeting and Finance*, vol. 38, no. 1, pp. 72–94, 2018, doi: 10.1111/pbaf.12165.
[16] Z. Shamsuddin, S. Mahmood, P. L. Ghazali, F. Salleh, and F. A. M. Nawi, “Indicators for Cooperative Performance Measurement,” *International Journal of Academic Research in Business and Social Sciences*, vol. 8, no. 12, pp. 577–585, 2018, doi: 10.6007/ijarbss/v8-i12/5056.
[17] M. S. Henock, “Financial sustainability and outreach performance of saving and credit cooperatives: The case of Eastern Ethiopia,” *Asia Pacific Management Review*, vol. 24, no. 1, pp. 1–9, Mar. 2019, doi: 10.1016/j.apmrv.2018.08.001.
[18] S. Khanderwal and D. Mohanty, “Stock Price Prediction Using ARIMA Model,” *International Journal of Marketing & Human Resource Research*, vol. 2, no. 2, pp. 98–107, 2021.
[19] D. S. Kumar, B. C. Thiruvarangan, C. S. N. Reddy, A. Vishnu, A. S. Devi, and D. Kavitha, “Analysis and Prediction of Stock Price Using Hybridization of SARIMA and XGBoost,” in *2022 International Conference on Communication, Computing and Internet of Things, IC3IoT 2022 - Proceedings*, IEEE, Mar. 2022, pp. 1–4, doi: 10.1109/IC3IOT53935.2022.9767868.
[20] B. U. Devi, D. Sundar, and P. Alli, “An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50,” *International Journal of Data Mining & Knowledge Management Process*, vol. 3, no. 1, pp. 65–78, Jan. 2013, doi: 10.5121/ijdkp.2013.3106.
[21] L. S. Malagrino, N. T. Roman, and A. M. Monteiro, “Forecasting stock market index daily direction: A Bayesian Network approach,” *Expert Systems with Applications*, vol. 105, pp. 11–22, Sep. 2018, doi: 10.1016/j.eswa.2018.03.039.
[22] O. Bustos and A. Pomares-Quimbaya, “Stock market movement forecast: A Systematic review,” *Expert Systems with Applications*, vol. 156, p. 113464, Oct. 2020, doi: 10.1016/j.eswa.2020.113464.
[23] S. Lakshminarayanan, G. R. Weckman, A. Snow, and J. Marvel, “Stock market hybrid forecasting model using neural networks,” in *IIE Annual Conference. Proceedings*, 2006, pp. 1–6.
[24] Y. Liu, “Stock Prediction Using LSTM and GRU,” in *Proceedings - 2022 6th Annual International Conference on Data Science and Business Analytics, ICDSBA 2022*, IEEE, Oct. 2022, pp. 206–211, doi: 10.1109/ICDSBA57203.2022.00054.
[25] J. Kavinnilaa, E. Hemalatha, M. S. Jacob, and R. Dhanalakshmi, “Stock Price Prediction Based on LSTM Deep Learning Model,” in *2021 International Conference on System, Computation, Automation and Networking, ICSCAN 2021*, IEEE, Jul. 2021, pp. 1–4, doi: 10.1109/ICSCAN53069.2021.9526491.
[26] J. Cao and J. Wang, “Stock price forecasting model based on modified convolution neural network and financial time series analysis,” *International Journal of Communication Systems*, vol. 32, no. 12, Aug. 2019, doi: 10.1002/dac.3987.
[27] S. Prakash, A. S. Jalal, and P. Pathak, “Forecasting COVID-19 Pandemic using Prophet, LSTM, hybrid GRU-LSTM, CNN-LSTM, Bi-LSTM and Stacked-LSTM for India,” in *2023 6th International Conference on Information Systems and Computer Networks, ISCON 2023*, IEEE, Mar. 2023, pp. 1–6, doi: 10.1109/ISCON57294.2023.10112065.
[28] D. Park and D. Ryu, “Forecasting stock market dynamics using bidirectional long short-term memory,” *Romanian Journal of Economic Forecasting*, vol. 24, no. 2, pp. 22–34, 2021.
[29] Y. Loday, P. Apirukvorapinit, and P. Vejjanugraha, “Stock Price Prediction Using Modified Bidirectional Long Short-Term Memory and Deep Learning Models: A Case Study of Bhutan Tourism Corporation Limited Stock Data,” in *2023 8th International Conference on Business and Industrial Research (ICBIR)*, IEEE, May 2023, pp. 645–650, doi: 10.1109/ICBIR57571.2023.10147486.
[30] R. Pascanu, M. Tomas, and B. Yoshua, “On the difficulty of training recurrent neural networks,” *Journal of Machine Learning Research*, vol. 8, no. 2, pp. 157–166, 2002.
[31] J. Li and X. Yang, “A Cyclical Learning Rate Method in Deep Learning Training,” in *2020 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2020, pp. 1–5, doi: 10.1109/CITS49457.2020.9232482.

BIOGRAPHIES OF AUTHORS

Evi Maria received her Ph.D. degree in accounting from Gadjah Mada University, Yogyakarta in 2019. Currently, she is an associate professor in accounting in the Department of Information Systems, Satya Wacana Christian University, Indonesia. Her research interests include financial health in non-profit organizations and accounting fraud detection in local government. She can be contacted at email: evi.maria@uksw.edu.

Teguh Wahyono is lecturer at the Faculty of Information Technology, Satya Wacana Christian University. He received his bachelor's degree in Informatics Engineering from UKDW Yogyakarta (1998), Master of Computer Science from Gadjah Mada University (2011) and Doctor of Computer Science from Bina Nusantara University (2020). His teaching interest are information system analysis, software engineering, and artificial intelligence. His research interest are spatial data mining, machine learning, and deep learning. He can be contacted at email: teguh.wahyono@uksw.edu.

Kristoko Dwi Hartomo received the Ph.D. degree in computer science from the Faculty of Science, Gadjah Mada University, Yogyakarta, in 2019. He has been active on research, since 2008, until now on geography information system. He is a Professor with the Faculty of Information Technology, Satya Wacana Christian University, Central Java, Indonesia. He has published his articles on international journals. Moreover, he had five copyrights and written some reference books on computer science. He can be contacted at email: kristoko@uksw.edu.

Purwanto obtained a Master's degree in Computer Science from the Faculty of Information Technology, Satya Wacana Christian University. He obtained his bachelor's degree in Accounting from Satya Wacana Christian University. He specializes in information systems and accounting research for non-profit organizations and small, micro and medium enterprises. He can be contacted at email: purwanto@uksw.edu.

Christian Arthur received the bachelor's degree in computer science from Brawijaya University, Indonesia. He is currently pursuing the Master of Data Science and Innovation degree with UTS, Australia. During his academic journey, he focused on the dynamic field of deep learning, particularly its application to forecasting. His primary research interest includes advancing forecasting techniques by combining theoretical knowledge with practical applications to improve predictive modeling. He is keen on implementing and improving deep learning across various domains and remains enthusiastic about contributing to innovations in predictive analytics and deep learning methodologies. He can be contacted at email: Christian.Arthur@student.uts.edu.au.