Vol. 7, No. 4, December 2025

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Predicting Accounts Receivable of the Social Security Administration for Employment Using LSTM Algorithm

Ainna Khansa[1]**, Usman Ependi**[2]

[1,2]Informatics Department, Postgraduates Studies, Bina Darma University, Palembang, Indonesia

Email: [1]242420023@students.binadarma.ac.id, [2]u.ependi@binadarma.ac.id

Received: Sept 26, 2025
Revised: October 25, 2025
Accepted: Nov 14, 2025
Published: Dec 9, 2025

Corresponding Author:
Author Name*:
Usman Ependi
Email*:
u.ependi@binadarma.ac.id

DOI:
10.63158/journalisi.v7i4.1274

Abstract. This study explores the use of Long Short-Term Memory (LSTM) networks for predicting outstanding contributions from employers to the BPJS Ketenagakerjaan, Indonesia's social security agency. The research aims to address the challenges BPJS faces due to delayed or unpaid contributions, which impact the institution's operational stability and financial health. The LSTM model, a deep learning technique well-suited for time-series prediction, was applied to historical data from BPJS Ketenagakerjaan to predict overdue contributions across three different training-validation splits: 70:30, 80:20, and 90:10. The results demonstrate that the 80:20 split achieved the highest validation accuracy of 84.71%, offering the optimal balance between training data and model generalization. The model's ability to predict overdue contributions with high accuracy could significantly improve BPJS's receivables management, allowing for more proactive financial planning and risk mitigation. The study also highlights the integration of an attention mechanism within the LSTM model, enhancing its predictive capabilities by focusing on the most relevant historical data. This research contributes to the field of predictive analytics in public sector financial management, showcasing the potential of machine learning in enhancing the efficiency and effectiveness of social security programs.

Keywords: Long Short-Term Memory (LSTM), Predictive Modeling, BPJS Ketenagakerjaan, Receivables Management, Financial Risk Management

1. INTRODUCTION

The workforce in Indonesia is a strategic element in national development. To ensure the protection and welfare of workers, the government established the Social Security Administration for Employment (BPJS Ketenagakerjaan), which manages social security programs covering Work Accident Insurance, Old Age Security, Pension Insurance, and Death Insurance [1]. Despite having strong regulatory foundations, BPJS Ketenagakerjaan faces significant challenges, particularly the accumulation of outstanding contributions from employers who fail to make timely payments or accumulate arrears. This situation poses a risk to the continuity of services for program participants. Analysis of BPJS's receivables management reveals that delayed payments impact operational stability and increase administrative burdens [2], [3].

The issue of outstanding contributions can hinder the smooth processing of claims and threaten the institution's financial stability. Currently, there is no effective system to predict potential arrears from both new and existing businesses. The absence of a predictive system makes it difficult for BPJS to take anticipatory measures against the risk of default on contributions, which can affect the effectiveness of social security management [4].

This research aims to address two key research questions: a) how can historical data be utilized to predict outstanding contributions from employers at BPJS Ketenagakerjaan? and b) how can the Long Short-Term Memory (LSTM) algorithm be implemented as an accurate and practical predictive model to support the institution's receivables risk management? Previous studies have shown that LSTM is effective in handling time-series data and predicting anomalies in payment behavior, making it a strong candidate for a historical-based prediction system [5], [6]. This study is significant because it supports the digitalization and efficiency of public financial management through predictive technology systems, which have yet to be widely adopted by public institutions.

This research is situated within the domain of data science-driven solutions for managerial issues in public institutions, particularly BPJS Ketenagakerjaan. To date, studies on BPJS have largely focused on descriptive or observational analyses. For instance, a study by [7] examined the determinants of compliance among independent participants, but only in one location and without a predictive approach. Research by [8] explored the macroeconomic factors influencing membership growth, but did not address operational issues like outstanding contributions. Meanwhile, [9] identified factors influencing premium payment flow but did not offer a systematic technology-based model for mitigating these issues.

The novelty of this research lies in the application of the Long Short-Term Memory (LSTM) algorithm in the context of predicting outstanding contributions for State-Owned Enterprises (SOEs). LSTM is known for its effectiveness in analyzing temporal data due to its ability to store and process long-term historical information. Recent studies have demonstrated that LSTM, especially when combined with attention mechanisms, significantly improves prediction accuracy in various time-series contexts, such as predicting dengue cases [10], vehicle and ship movements [11], [12], and energy consumption in the public sector [13].

Furthermore, the integration of an attention layer within the model architecture enhances the efficiency of identifying important patterns, improving accuracy, and providing deeper insights compared to traditional statistical methods. Attention mechanisms allow the model to focus on the most relevant data for making predictions, as evidenced by studies across various public sectors and industries [14], [15]. This approach has not yet been widely implemented in Indonesia's social security sector, making this study a valuable contribution both scientifically and practically in advancing predictive technologies to support strategic decision-making in public institutions like BPJS Ketenagakerjaan.

2. METHODS

This research is conducted using an applied technology approach based on data science, systematically and incrementally designed to develop a predictive model for receivables using the Long Short-Term Memory (LSTM) algorithm. This approach is highly relevant in the context of data-driven decision-making in public institutions such as BPJS Ketenagakerjaan, as it allows for more accurate and preventive financial risk

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

management. The research stages follow a stepwise process, as shown in Figure 1. The development of the predictive model consists of three main sub-phases.

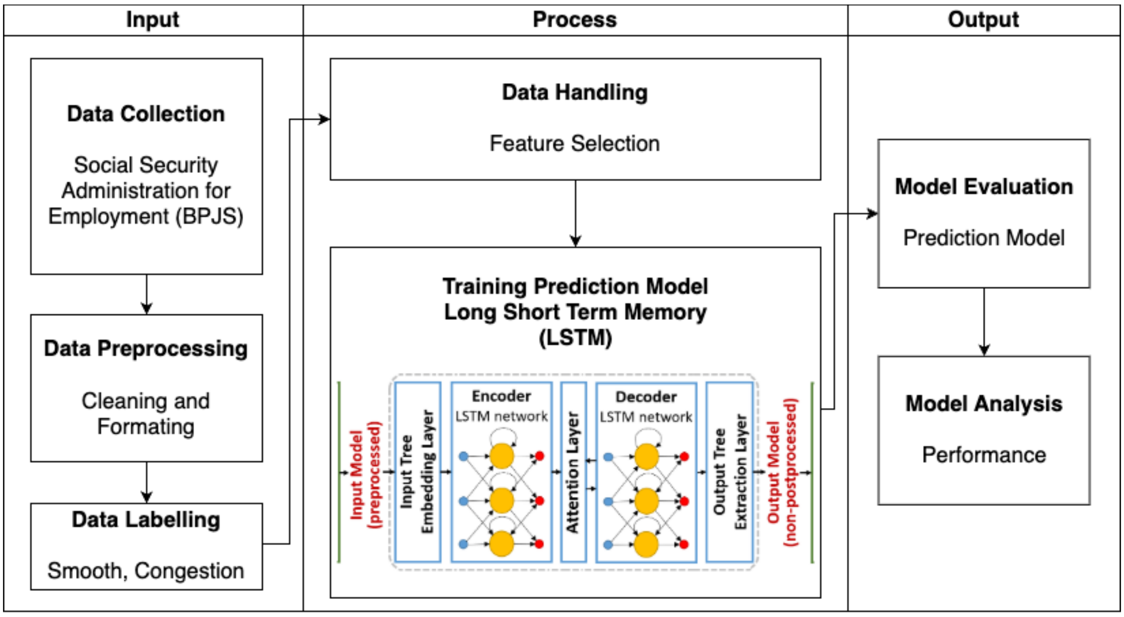

Figure 1. The Proposed Research Flow

The research flow, as illustrated in Figure 1, outlines the systematic process followed in this study. It provides a clear overview of the key stages involved, from data collection to model evaluation. Each step is carefully designed to ensure the development of an effective predictive model, addressing the research objectives and challenges associated with receivables prediction.

2.1. Input Phase: Data

The process begins with data collection, where a historical dataset is obtained from the BPJS Ketenagakerjaan Branch in Palembang. This dataset includes crucial information on receivables, employer identities, and payment behavior records. It consists of 14,580 entries spanning six months, with ten critical attributes essential for developing the predictive model. These attributes include the NPP (Employer Identification Number), work unit, company name, supervisor name, business scale, last reconciliation date, value of last reconciliation, months of arrears, criteria, and outstanding receivables.

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Following data collection, data preprocessing is performed, which involves cleaning the data, removing noise, unifying formats, and making structural adjustments to ensure compatibility with the model training process. The next step is data labeling, where each entry is annotated or classified based on the payment status, which serves as the prediction target for the model (e.g., smooth = 1 or overdue = 0). As an example, row 57409 in Table 1 represents an entry in the cleaned and balanced dataset, with the payment status labeled as "overdue" (0), and is now ready for use in predicting accounts receivable for the Social Security Administration for Employment.

Table 1. Clean and Balanced Data

Row	NPP	Date	Recon Value	Receivable	Label
1	14001586	01/09/24	1215180.0	1239481.8	1
2	14001586	03/09/24	1215180.0	1239481.8	1
4	14001586	10/09/24	1215180.0	1239481.8	1
5	14001586	11/09/24	1215180.0	1239481.8	1
6	14001586	17/09/24	1215180.0	1239481.8	1
...	...	...	...	...	...
57405	MG004400	17/02/25	37669.68	38424.0	0
57406	MG004400	19/02/25	37669.68	38424.0	0
57407	MG004400	21/02/25	37669.68	38424.0	0
57408	MG004400	25/02/25	37669.68	38424.0	0
57409	MG004400	26/02/25	37669.68	38424.0	0

2.2. Process Phase: Modelling

In this phase, the focus is on handling imbalanced data, performing feature selection, and transforming the dataset into a format that is compatible with the LSTM model. The data preprocessing steps involve addressing class imbalance, where techniques such as oversampling, undersampling, or using weighted loss functions are applied to ensure that the model can effectively learn from both the majority and minority classes. Feature selection is then performed to identify the most significant variables contributing to the model's predictions, improving the model's efficiency and accuracy.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Once the data is prepared, it is fed into the LSTM-based prediction system. The LSTM model follows an encoder-decoder architecture, which is strengthened with an attention mechanism. The attention mechanism allows the model to focus on the most relevant historical information, improving the quality of predictions by dynamically assigning different levels of importance to different time steps in the input sequence. Mathematically, the LSTM model can be represented as shown in Equation 1.

$$
\begin{gathered}
i_t = \sigma\left(W_i \cdot \left[h_{\{t-1\}}, x_t\right] + b_i\right) \\
f_t = \sigma\left(W_f \cdot \left[h_{\{t-1\}}, x_t\right] + b_f\right) \\
o_t = \sigma\left(W_o \cdot \left[h_{\{t-1\}}, x_t\right] + b_o\right) \\
c_t = f_t \cdot c_{\{t-1\}} + i_t \cdot tanh\left(W_c \cdot \left[h_{\{t-1\}}, x_t\right] + b_c\right) \\
h_t = o_t \cdot tanh(c_t)
\end{gathered} \tag{1}
$$

Where:
i_t is the input gate, f_t is the forget gate, o_t is the output gate, c_t is the cell state, h_t is the hidden state, *Wi, Wf, Wo, Wc* are the weight matrices, and b_i, b_f, b_o, b_c are the bias terms.

The attention mechanism assigns attention scores to different time steps based on their relevance to the current prediction, enhancing the LSTM's ability to focus on important features in the sequence. The attention mechanism can be expressed as shown Equation 2.

$$
\begin{gathered}
\alpha_t = softmax(W_a \cdot h_t + b_a) \\
\widehat{}\{h\}_t = \alpha_t \cdot h_t
\end{gathered} \tag{2}
$$

Where:
α_t represents the attention weight at time step t, W_a is the attention weight matrix, and $h^\wedge{}_t$ is the weighted hidden state.

In the architecture shown in Figure 1, the pre-processed input data flows through the embedding layer, then passes through the encoder LSTM network. The output of the encoder LSTM is enhanced by the attention mechanism, which allows the model to assign varying degrees of importance to the historical data at each time step. This refined information is then processed by the decoder LSTM network to generate predictions of the receivables.

2.3. Output Phase: Evaluation

The output phase involves model evaluation and analysis. Evaluation methods, such confusion metric, are applied to assess the quality of the model's predictions. This phase also involves in-depth performance analysis of the model on new data and identifying system weaknesses that can be improved in subsequent iterations.

3. RESULTS AND DISCUSSION

3.1. Experimental Setup

In this experiment, the objective was to train a Long Short-Term Memory (LSTM) model for binary classification using time-series data. The setup followed a systematic approach, involving data preprocessing, model training, and evaluation, with the data split across three different ratios: 70:30, 80:20, and 90:10 for training and testing sets. These scenarios allowed for testing the model's robustness across different amounts of training data.

The first step in the setup was to split the dataset into training and testing sets. For each scenario, different proportions of the data were allocated for training, while the remaining data was reserved for testing. The splits were as follows:

1) 70:30: 70% of the data was used for training and 30% for testing.
2) 80:20: 80% for training and 20% for testing.
3) 90:10: 90% for training and 10% for testing.

These splits were done without shuffling the data, maintaining the chronological order, which is critical for time-series problems where the sequence of data points affects predictions. Each data split was processed independently to ensure fair comparisons between the different training set sizes. After splitting the data, it was reshaped to fit the input requirements of the LSTM model. The input features were flattened to two dimensions, and the data was normalized using a MinMaxScaler to scale the features

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

between 0 and 1. Normalization is particularly important for LSTM models as it accelerates convergence and prevents the model from being biased by features with larger scales.

1) The LSTM model used in this experiment had a simple architecture consisting of:
2) An LSTM layer with 64 units, which processes the sequential data.
3) A Dropout layer with a rate of 0.2 to prevent overfitting by randomly dropping 20% of the neurons during training.
4) A Dense layer with a sigmoid activation function to output a probability value between 0 and 1 for binary classification.

The model was compiled using the Adam optimizer and binary cross-entropy loss, standard choices for binary classification tasks. Early stopping was employed to halt training if the validation loss did not improve for three consecutive epochs, preventing overfitting and saving computational resources.

The model was trained for 20 epochs, and after each epoch, the performance was evaluated using both training and validation data. The training process was monitored using accuracy and loss metrics, which were plotted for visual inspection. The model's final performance was recorded based on the evaluation metrics obtained from the test data. This setup allowed for an in-depth analysis of how the model's performance varies with different training set sizes. By comparing results across the three data splitting scenarios 70:30, 80:20, and 90:10, we could assess how the amount of training data influences the model's ability to generalize to unseen data.

3.2. Performance Evaluation

In this study, we evaluated the performance of the model across three distinct training-to-testing data ratios: 70:30, 80:20, and 90:10. This methodology allowed for a comprehensive analysis of how varying amounts of training data impact the model's accuracy and loss, shedding light on the trade-off between training data size and generalization ability. The model was trained over 20 epochs in each scenario, and key performance metrics—namely, training accuracy, validation accuracy, training loss, and validation loss—were meticulously tracked throughout the process.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

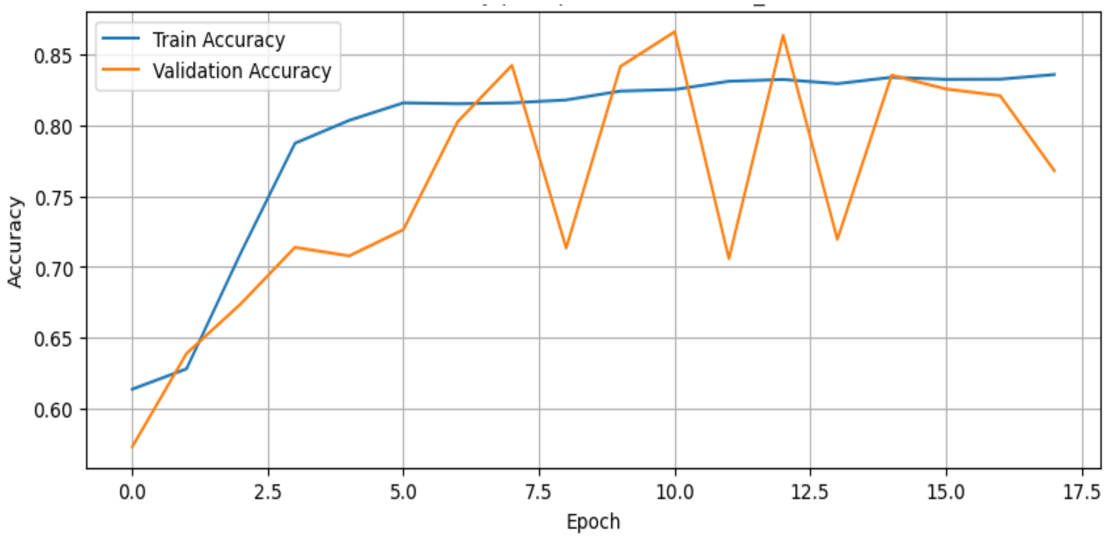

Figure 2. Accuracy Ratio 70:30

For the 70:30 data split, the model initiated training with a relatively low training accuracy of 61.28% and a training loss of 0.6691 in Epoch 1. The validation accuracy at this point was 57.31%, with a corresponding validation loss of 0.6742. As training progressed, both training and validation performance showed consistent improvement. By Epoch 20, the model achieved a training accuracy of 83.54% and a validation accuracy of 83.54%, with the validation loss decreasing to 0.3889. This steady and continuous improvement in both training and validation performance is visually depicted in Figure 2 (Accuracy Ratio 70:30) and Figure 3 (Loss Ratio 70:30), confirming the model's ability to generalize effectively over the course of training.

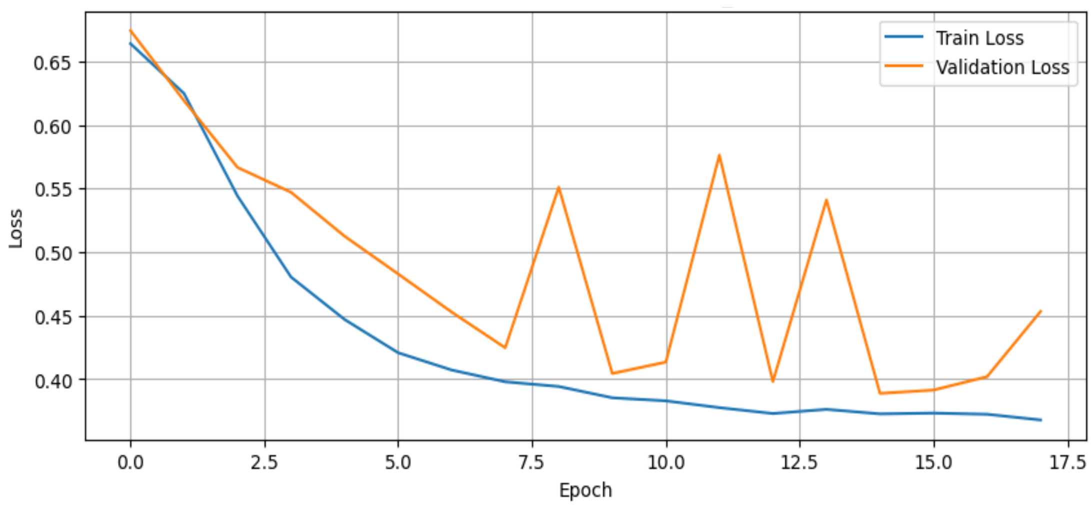

Figure 3. Lost Ratio 70:30

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The 80:20 split exhibited a similar trend, with the model starting at a training accuracy of 63.00% and a training loss of 0.6613 in Epoch 1, and a validation accuracy of 50.07% with a validation loss of 0.6987. As the model continued through the epochs, both training and validation performance exhibited significant gains. By Epoch 20, the model reached a training accuracy of 82.88% and a validation accuracy of 84.71%, with the validation loss reduced to 0.4139. These results demonstrate the model's strong ability to generalize across different data splits, as shown in Figure 4 (Accuracy Ratio 80:20) and Figure 5 (Loss Ratio 80:20).

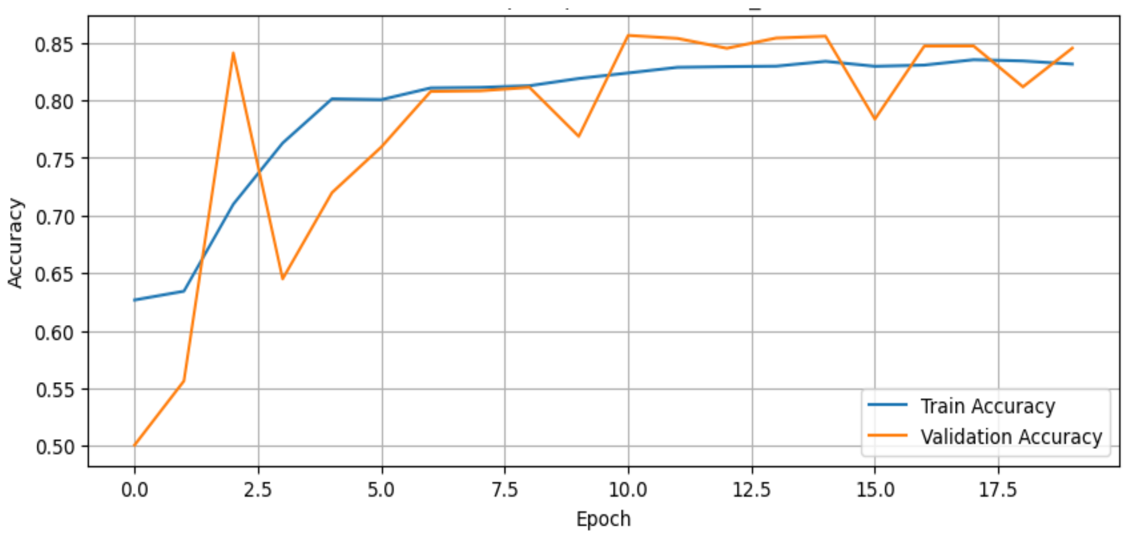

Figure 4. Accuracy Ratio 80:20

The 90:10 ratio, although providing a larger training dataset, did not result in the same level of improvement in validation performance. The model started with a training accuracy of 61.37% and a training loss of 0.6684 in Epoch 1, with a validation accuracy of 48.53% and a validation loss of 0.6859. Despite the increased training data, by Epoch 20, the model reached a training accuracy of 83.62%, but the validation accuracy was lower at 79.63%, with the validation loss rising to 0.5029. This suggests that, while additional training data can contribute to model learning, it does not necessarily guarantee better generalization, particularly if the model is unable to sufficiently adapt to the validation set. These results are illustrated in Figure 6 (Accuracy Ratio 90:10) and Figure 7 (Loss Ratio 90:10).

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

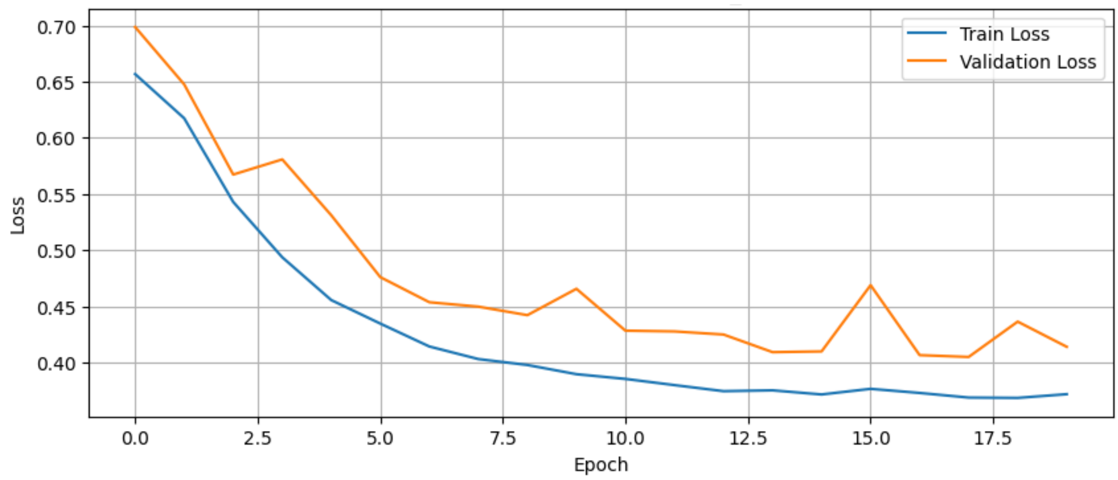

Figure 5. Lost Ratio 80:20

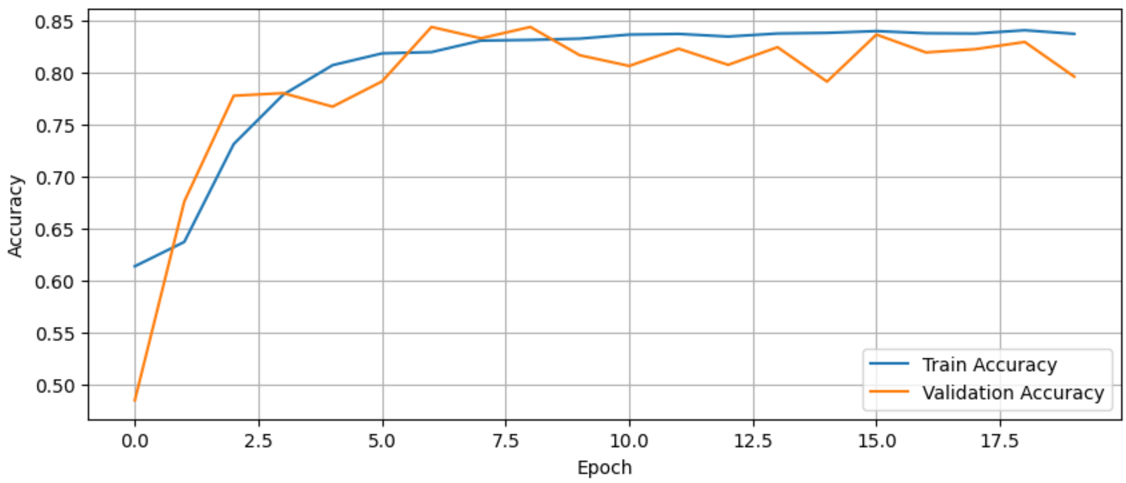

Figure 6. Accuracy Ratio 90:10

The final performance metrics across the different data split scenarios revealed the following final validation accuracy: 83.54% for the 70:30 split, 84.71% for the 80:20 split, and 79.63% for the 90:10 split. The 80:20 ratio achieved the highest validation accuracy, indicating that it strikes the most favorable balance between the amount of training data and the model's ability to generalize effectively. The 70:30 ratio also performed well, demonstrating that a smaller training set can still yield robust results, while the 90:10 ratio, despite providing more training data, led to slightly reduced validation accuracy.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

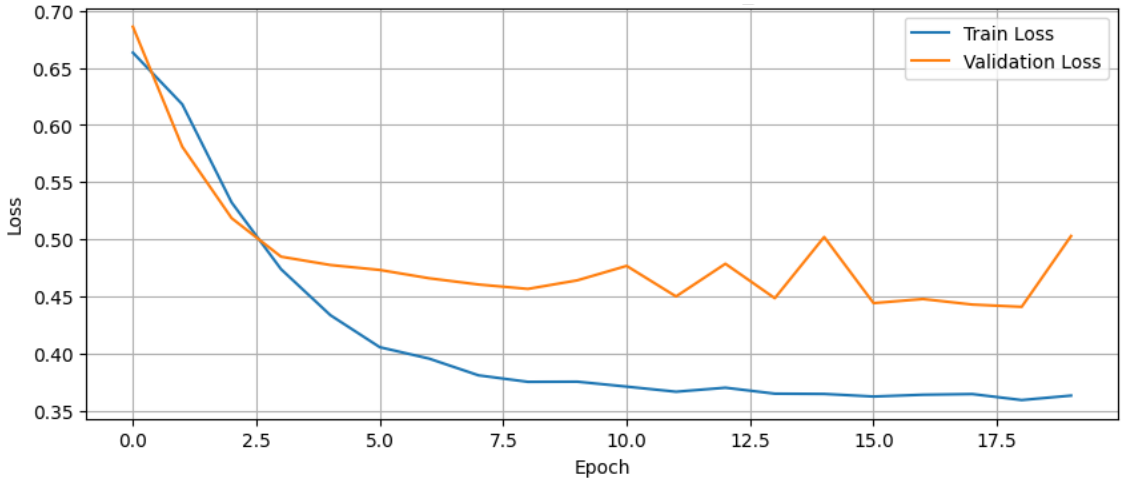

Figure 7. Lost Ratio 90:10

Figure 8 presents a comprehensive visualization of the training accuracy across all three data split scenarios. As shown, the model consistently improved over the course of training in each scenario, but the 80:20 split resulted in the highest validation accuracy and demonstrated more consistent performance throughout the epochs compared to the other splits. Finally, while increasing the amount of training data—such as in the 90:10 split—may offer a larger dataset for the model to learn from, it does not always translate into improved generalization. The 80:20 split provides the optimal balance between sufficient training data and the model's ability to generalize to unseen data, leading to the highest overall performance. These findings emphasize the importance of choosing an appropriate training-validation split for model development, as it can significantly impact the model's effectiveness and generalization capabilities.

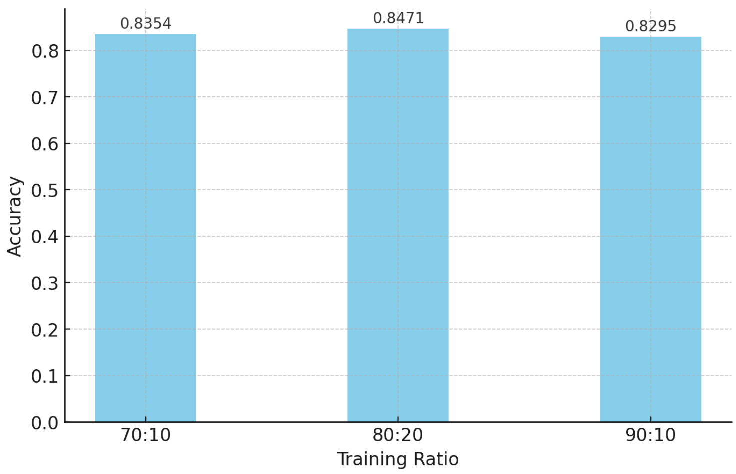

Figure 8. Tranning Accuracy for Each Scenario

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Following the analysis of training accuracy in Figure 8, the next critical step in evaluating the model's performance is to assess its ability to predict outcomes in real-world scenarios. Table 2 presents the Prediction Testing results, showcasing how the model performs on unseen data over a series of days. This table includes predictions for several entities, identified by their NPP codes, across three consecutive days. As seen in the table, the model consistently predicts "Overdue" for most NPPs across all three days, except for a few instances where the prediction varies, such as for NPP MG004392, which is predicted as "Smooth" on Day 1 and "Overdue" on the subsequent days. These results illustrate the model's ability to classify data based on patterns learned during training, and it serves as an essential indicator of how well the model can generalize from training to actual predictions in a dynamic, time-sensitive environment.

The consistency in the "Overdue" predictions suggests that the model is identifying significant patterns in the data associated with overdue statuses. However, the variation in predictions, like for MG004392, indicates that the model also recognizes potential cases where payment behaviors deviate from the usual trend, suggesting that the model has learned nuanced distinctions between different statuses. This is crucial for decision-making in systems like BPJS Ketenagakerjaan, where predicting overdue payments accurately can help prevent financial risks.

Table 2. Prediction Testing

NPP	Day 1	Day 2	Day 3
14001586	Overdue	Overdue	Overdue
14004972	Overdue	Overdue	Overdue
14007058	Overdue	Overdue	Overdue
14011986	Overdue	Overdue	Overdue
15000732	Overdue	Overdue	Overdue
...	...	...	...
GG007378	Overdue	Overdue	Overdue
GG007398	Overdue	Overdue	Overdue
GG007423	Overdue	Overdue	Overdue
MG004392	Smooth	Overdue	Overdue
MG004400	Overdue	Overdue	Overdue

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

3.3. Discussion

The results of this study demonstrate the effectiveness of using the Long Short-Term Memory (LSTM) algorithm to predict overdue payments and receivables in the context of BPJS Ketenagakerjaan. The evaluation of different training-validation split scenarios (70:30, 80:20, and 90:10) has provided valuable insights into how varying the amount of training data affects the model's ability to generalize and predict future outcomes.

The 80:20 split yielded the highest validation accuracy of 84.71%, outperforming the other ratios. This outcome suggests that the 80:20 ratio strikes an optimal balance between providing sufficient data for training while retaining enough unseen data to test the model's generalization capability. The 70:30 split also performed well, achieving a validation accuracy of 83.54%, which demonstrates that even with a smaller training dataset, the model can still perform effectively. However, the 90:10 ratio, despite offering more training data, resulted in a slightly reduced validation accuracy (79.63%), indicating that more data does not necessarily equate to better performance, especially if the model cannot generalize well on unseen data.

These findings emphasize that while increasing the amount of training data—such as in the 90:10 split—can lead to improved model training, it does not always lead to better generalization. This is crucial for public institutions like BPJS Ketenagakerjaan, where the goal is to predict overdue contributions and manage financial risks effectively. Too much training data might lead to overfitting, where the model learns specific patterns in the training data but fails to generalize to new, unseen data.

The training accuracy, as shown in Figure 8, consistently improved across all three data splits, reflecting the model's capacity to learn over time. However, the 80:20 split not only resulted in the highest validation accuracy but also demonstrated the most consistent performance, reaffirming its suitability for predicting outcomes in real-world, dynamic scenarios.

Table 2 presents the Prediction Testing results, where the model predicts "Overdue" for most NPPs across three consecutive days. The consistency of these predictions highlights the model's ability to detect significant patterns in overdue payment statuses. However, in the case of NPP MG004392, the model predicts a "Smooth" status on Day 1 and "Overdue" on subsequent days, suggesting that the model is capable of recognizing subtle variations in payment behavior. This nuanced prediction ability is crucial for BPJS Ketenagakerjaan, as it allows the institution to identify potential risks earlier and take proactive measures to address overdue payments before they become a larger issue.

This study demonstrates that the LSTM model, combined with appropriate data preprocessing techniques and a well-chosen training-validation split, can provide an effective predictive tool for managing overdue receivables at BPJS Ketenagakerjaan. The application of LSTM, particularly with the integration of attention mechanisms, offers a significant improvement over traditional methods and supports the advancement of predictive technologies in the management of public financial systems. Moreover, the findings underscore the importance of selecting an optimal training-validation split, as it can have a substantial impact on model performance and generalization.

In comparison to previous studies in the field, this research stands out by addressing the specific problem of overdue payments in public institutions, particularly focusing on BPJS Ketenagakerjaan. While past studies have primarily explored descriptive or observational analyses [7], [8], [9], this research introduces a novel approach by applying machine learning techniques, specifically LSTM, to predict arrears in contributions, which has not been widely explored in the context of social security management in Indonesia. Table 3 compares the predictive accuracy of our approach to similar studies.

Table 3. Comparison of Prediction Accuracy

Study	Model	Training Data Size	Accuracy (%)
This Study (80:20 Split)	LSTM	80%	84.71
[7]	Logistic Regression	Limited (One Location)	78.00
[8]	Regression Model	Macroeconomic Data	75.50
[9]	Descriptive Analysis	N/A	N/A

This table highlights the significant improvement in prediction accuracy using the LSTM model compared to previous works. While earlier studies focused more on macroeconomic factors or descriptive insights, our model applies time-series analysis and machine learning to predict overdue payments more effectively, achieving higher accuracy and providing actionable insights for proactive management. Finally, this study demonstrates the power of LSTM for predictive analytics in public financial management, particularly in the context of BPJS Ketenagakerjaan, offering a more data-driven, accurate, and efficient approach to managing receivables and minimizing financial risks.

4. CONCLUSION

This study successfully addresses the two primary research objectives: (1) leveraging historical data to predict outstanding contributions from employers at BPJS Ketenagakerjaan, and (2) implementing the Long Short-Term Memory (LSTM) algorithm as an accurate and practical predictive model for receivables risk management. By utilizing LSTM, the model effectively processed time-series data to predict overdue contributions, achieving high validation accuracy, particularly with the 80:20 data split, which balanced training data and model generalization. The model demonstrated its capacity to identify key patterns in payment behaviors, thereby offering BPJS Ketenagakerjaan a tool for anticipatory risk management and ensuring more effective financial planning. Moreover, the integration of the attention mechanism within the LSTM architecture allowed the model to focus on the most relevant historical data, improving prediction accuracy. The findings suggest that the LSTM-based model can serve as an innovative solution for BPJS Ketenagakerjaan, enabling better management of receivables, reducing financial risks, and enhancing the overall effectiveness of social security programs. This research contributes to the growing body of work in data science for public institutions, highlighting the potential for machine learning applications to drive operational efficiency and proactive decision-making in financial management.

ACKNOWLEDGMENT

We would like to express our deepest gratitude to the Ministry of Education, Technology, and Higher Education (Kemendiktisaintek) for funding this research through the PTM

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Grant, facilitated by LLDIKTI Region II and Universitas Bina Darma. This work has been made possible thanks to the financial support provided under the Master Contract Number: 123/C3/DT.05.00/PL/2025, and the Sub-Contract Number: 145 /LL2/DT.05.00/PL/2025, 008/SPK/DRPM/Univ-BD/VI/2025. We also extend our sincere appreciation to all the individuals and institutions involved for their continuous support and guidance throughout the research process. Without their contribution, this research would not have been successful. Thank you.

REFERENCES

[1] S. U. Adillah and A. Purnawan, "Pelaksanaan Program BPJS Ketenagakerjaan Bagi Pekerja Informal Di Bidang Peternakan Dan Pertanian Di Kecamatan Gamping Kabupaten Sleman," *Journal of Morality and Legal Culture*, vol. 1, no. 1, pp. 1–14, 2016, doi: 10.20961/JMAIL.V1I1.44757.

[2] Irna Sri Wijayanti, Sri Rahayu, and Rachmad Rachmad, "Analisis Pengelolaan Piutang Pasien BPJS Dalam Rangka Mengoptimalkan Kinerja Keuwangan Di Direktorat Keuangan RSAB Harapan Kita Jakarta," *Jurnal Manajemen dan Administrasi Rumah Sakit Indonesia (MARSI)*, vol. 1, no. 1, pp. 64–81, 2017.

[3] Ghaesani Alifa, Nur Fitriyah, and Nungki Kartikasari, "Analisis Pengelolaan Piutang Jasa Pelayanan BPJS Dalam Rangka Mengoptimalkan Kinerja Keuangan (Studi Kasus Pada Rumah Sakit Cahaya Medika Praya)," *Jurnal Riset Mahasiswa Akuntansi*, vol. 3, no. 3, pp. 12–25, 2023.

[4] D. Efriadi, R. Rahmaddeni, A. Agustin, and J. Junadhi, "Prediksi Penambahan Piutang Iuran Jaminan Sosial Ketenagakerjaan menggunakan Algoritma K-Nearest Neighbor," *Edumatic: Jurnal Pendidikan Informatika*, vol. 6, no. 1, pp. 49–57, Jun. 2022, doi: 10.29408/EDUMATIC.V6I1.5255.

[5] H. Hu, L. Tang, S. Zhang, and H. Wang, "Predicting the direction of stock markets using optimized neural networks with Google Trends," *Neurocomputing*, vol. 285, no. 4, pp. 188–195, 2018, doi: 10.1016/J.NEUCOM.2018.01.038.

[6] Ferdiansyah, S. H. Othman, R. Zahilah Raja Md Radzi, D. Stiawan, Y. Sazaki, and U. Ependi, "A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market," *2019 International Conference on Electrical Engineering and*

Computer Science (ICECOS), pp. 206–210, Oct. 2019, doi: 10.1109/ICECOS47637.2019.8984499.

[7] F. Anggraeni, "Determinant of compliance in paying contributions of independent participants in BPJS Kesehatan at RSUD Haji Kota Makassar," *Jurnal Kesehatan Masyarakat*, vol. 15, pp. 25–35, 2020.

[8] Gultom A, "Pengaruh perkembangan makroekonomi Indonesia terhadap pertumbuhan kepesertaan BPJS Ketenagakerjaan," *Jurnal Ekonomi dan Pembangunan*, vol. 14, pp. 112–125, 2016.

[9] Dennis Indah Indra Putri, Lukytawati Anggraeni, and Widyastutik, "Faktor yang Mempengaruhi Kelancaran Pembayaran Premi BPJS Ketenagakerjaan (Studi Kasus Di Kota Tangerang)," *Jurnal Ekonomi dan Kebijakan Pembangunan*, vol. 12, pp. 84–100, 2023.

[10] M. A. Majeed, H. Z. M. Shafri, A. Wayayok, and Z. Zulkafli, "Prediction of dengue cases using the attention-based long short-term memory (LSTM) approach," *Geospat Health*, vol. 18, no. 1, p. 176, 2023, doi: 10.4081/GH.2023.1176.

[11] S. Qiao, F. Gao, J. Wu, and R. Zhao, "An Enhanced Vehicle Trajectory Prediction Model Leveraging LSTM and Social-Attention Mechanisms," *IEEE Access*, vol. 12, pp. 1718–1726, 2024, doi: 10.1109/ACCESS.2023.3345643.

[12] T. Yu, Y. Zhang, S. Zhao, J. Yang, W. Li, and W. Guo, "Vessel trajectory prediction based on modified LSTM with attention mechanism," *2024 4th International Conference on Neural Networks, Information and Communication (NNICE)*, pp. 912–918, 2024, doi: 10.1109/NNICE61279.2024.10498270.

[13] Y. Chang, S. Zhang, and G. You, "Research on Electric Heating Energy Consumption Prediction Based on SSA-LSTM-Attention," *Academic Journal of Engineering and Technology Science*, vol. 7, no. 4, 2024, doi: 10.25236/AJETS.2024.070416.

[14] X. Wen and W. Li, "Time Series Prediction Based on LSTM-Attention-LSTM Model," *IEEE Access*, vol. 11, pp. 48322–48331, 2023, doi: 10.1109/ACCESS.2023.3276628.

[15] S. Sang and L. Li, "A Novel Variant of LSTM Stock Prediction Method Incorporating Attention Mechanism," *Mathematics*, vol. 12, no. 7, Apr. 2024, doi: 10.3390/MATH12070945.