BILINGUAL PUBLISHING GROUP

Artificial Intelligence Advances

https://journals.bilpubgroup.com/index.php/aia

ARTICLE

Inception Residual RNN-LSTM Hybrid Model for Predicting Pension Coverage Trends among Private-Sector Workers in the USA

***Kaixian Xu* [1], *Yiqun Cai* [2], *Alan Wilson* [3*]**

[1] *Risk & Quant Analytics, BlackRock, Jersey City, NJ 07310, USA*

[2] *University of Florida, Herbert Wertheim College, Gainesville, FL 32611, USA*

[3] *Intact Financial Corporation, Toronto, ON M5G 0A1, Canada*

ABSTRACT

Pensions are fundamental to financial security in retirement, especially in the U.S., where they play a critical role in ensuring stability for retirees and fostering broader economic benefits. However, predicting pension coverage trends poses significant challenges due to the complexity of labor markets, demographic shifts, and economic variabilities. Traditional statistical models, though foundational, often fail to handle the nonlinear patterns inherent in pension data. To address these limitations, we propose the Inception residual RNN-LSTM hybrid model, which combines the strengths of Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks with residual connections. This model captures diverse temporal patterns while mitigating vanishing gradient issues, delivering superior performance in predicting pension coverage trends. Experimental results demonstrate that our model outperforms traditional machine learning models and standalone deep learning architectures like RNN and LSTM. Its robust performance across key metrics highlights its potential as a reliable tool for forecasting complex pension trends and aiding policymakers, employers, and financial institutions in effective retirement planning.

Keywords: Pension Coverage Prediction; Machine Learning; RNN-LSTM

*CORRESPONDING AUTHOR:

Alan Wilson, Intact Financial Corporation, Toronto, ON M5G 0A1, Canada; Email: alan.wilson@intact.net

ARTICLE INFO

Received: 5 January 2025 | Revised: 15 February 2025 | Accepted: 19 February 2025 | Published Online: 28 February 2025

DOI: https://doi.org/10.30564/aia.v7i1.8704

CITATION

Xu, K., Cai, Y., Wilson, A., 2025. Inception Residual RNN-LSTM Hybrid Model for Predicting Pension Coverage Trends among Private-Sector Workers in the USA. Artificial Intelligence Advances. 7(1): 1–9. DOI: https://doi.org/10.30564/aia.v7i1.8704

1. Introduction

Pensions are a cornerstone of financial security for retirees, providing a reliable source of income after their working years[1–3]. In the United States, pension plans primarily fall into two categories: defined benefit (DB) plans, where employers guarantee a specific payout upon retirement, and defined contribution (DC) plans, such as 401(k)s, where employees contribute a portion of their income into investment accounts[4, 5]. Over the decades, there has been a marked shift from DB plans to DC plans, as private-sector employers have sought to reduce long-term liabilities. This shift has introduced greater individual responsibility for retirement planning, making it critical to understand and predict pension coverage trends, especially in the private sector.

The role of pensions in society cannot be overstated. For millions of American workers, pensions provide a safety net, ensuring financial stability and preventing poverty in old age[6, 7]. Beyond individual benefits, pensions contribute significantly to the broader economy by fostering consumer spending and stabilizing markets. However, the dynamic nature of the labor market, changing demographics, and evolving regulatory landscapes make predicting pension coverage trends a complex task. Accurate prediction models are vital for policymakers, employers, and financial institutions to ensure effective retirement planning and adapt to societal changes.

Despite their importance, predicting pension coverage trends presents unique challenges[8–10]. The private sector in the U.S. is highly diverse, encompassing industries with varying workforce compositions, benefits structures, and levels of unionization. Furthermore, economic factors, such as inflation, recessions, and wage stagnation, compound the difficulty of making accurate forecasts. Traditional statistical models, such as linear regression or time series analysis[11–14], have been extensively applied to predict similar trends. While these models can capture straightforward relationships, they often fail to handle the nonlinear patterns and complex interdependencies inherent in pension data[15–18]. For instance, shifts in worker demographics or sudden policy changes may introduce variabilities that traditional models struggle to accommodate. In recent years, artificial intelligence (AI) and machine learning (ML) have revolutionized predictive modeling across domains[19–21]. AI models have shown remarkable success in uncovering patterns and relationships within high-dimensional and time-dependent data. Their ability to adapt to nonlinearity, handle missing values, and integrate heterogeneous datasets makes them especially suitable for complex forecasting tasks like pension coverage prediction. Given these advantages, adopting AI-driven approaches is not merely an enhancement—it is a necessity for improving prediction accuracy and providing actionable insights[22, 23].

In this study, we propose a hybrid AI framework called Inception Residual RNN-LSTM to address the limitations of traditional models and advance the state of pension coverage prediction. Our model incorporates the strengths of two deep learning architectures: the recurrent neural network (RNN)[24–26] and the long short-term memory (LSTM)[27–29] network. RNNs are well-suited for capturing sequential dependencies in time-series data, while LSTMs overcome the vanishing gradient problem, enabling the model to learn long-term dependencies effectively. By integrating these architectures into an Inception-like structure, which combines parallel processing and residual connections[30, 31], we enhance the model's capacity to capture diverse temporal patterns and reduce prediction errors. The architecture of our proposed model is illustrated in the accompanying **Figure 1**. The framework begins with preprocessing the raw pension data, segmenting it into overlapping sequences to preserve temporal relationships. These sequences are then fed into the Inception residual RNN-LSTM module, which consists of multiple branches. Each branch independently processes the data using RNN and LSTM units, extracting complementary features at different temporal scales. The outputs from these branches are concatenated and passed through a dense fusion layer, where they are further refined. Residual connections are employed to stabilize training and ensure gradient flow, improving the model's convergence and robustness.

To validate the effectiveness of our framework, we benchmarked its performance against the single RNN and LSTM model as well as traditional machine learning models such as decision trees, k-nearest neighbors (KNN), and linear regression. Metrics such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R^2 were used to evaluate model accuracy. Our results demonstrate that the Inception RNN-LSTM outperforms these baseline models across all metrics, providing significantly improved predictions of pension coverage trends among private-sector workers in the U.S.

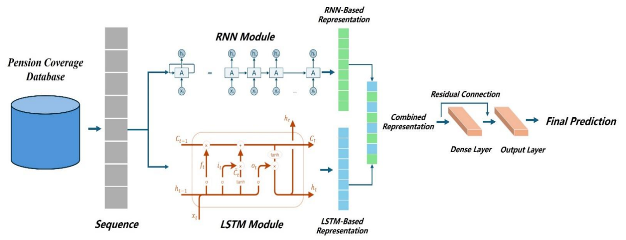

Figure 1. The workflow of the proposed inception residual RNN-LSTM hybrid model.

2. Literature Review

Recent advancements in predictive analytics have significantly enhanced the modeling of pension coverage trends. Traditional statistical methods, while foundational, often fall short in capturing the complex, nonlinear patterns inherent in pension data. To address these limitations, researchers have increasingly turned to machine learning (ML) and artificial intelligence (AI) techniques [32–35]. Rocha Salazar and Boado-Penas applied machine learning algorithms to predict early retirement decisions, utilizing data from private pension plans to assess the likelihood of individuals retiring before or after the age of 65 based on personal and macroeconomic factors [36]. A survey by Maastricht University highlighted the applicability of data science techniques in the pension industry, identifying key areas such as customer focus, organizational process optimization, and personnel management where machine learning can be effectively applied [37]. Furthermore, some other studies also reported that AI technologies can enhance pension plan governance by facilitating multi-stakeholder interactions, reducing administrative tasks, and aiding pension boards with decision-making, including the optimization of investment strategies.

3. Method

3.1. Dataset Preparation

For our study on pension coverage trends among private-sector workers in the USA, we utilized a publicly available dataset sourced from Kaggle [38]. This dataset provides valuable insights into pension coverage trends spanning 41 years, with information segmented by various demographic factors, such as race, gender, education level, and recent graduation status. The dataset focuses on employer-provided pension plans and is based on data collected from the Economic Policy Institute's State of Working America Data Library. For our analysis and prediction tasks, we focused on three key features shown in **Figure 2** from the dataset: 'all', representing the overall pension coverage rate; 'high_school', denoting pension coverage for high school graduates; 'bachelors_degree', indicating coverage among bachelor's degree holders. By selecting these features, we aimed to investigate not only the overarching pension coverage trends but also the specific impact of educational attainment on pension participation. This approach helps us understand disparities and patterns across different levels of education while considering the broader context of overall pension trends. The data was partitioned into a training set, comprising the years 1979 to 2007, and a testing set, consisting of the years 2008 to 2016. This temporal split ensures that the models are trained on historical data while being tested on more recent trends, enabling a realistic evaluation of predictive performance. A sliding window method was applied to structure the dataset into sequences, preserving temporal dependencies and preparing the data for time-series prediction tasks.

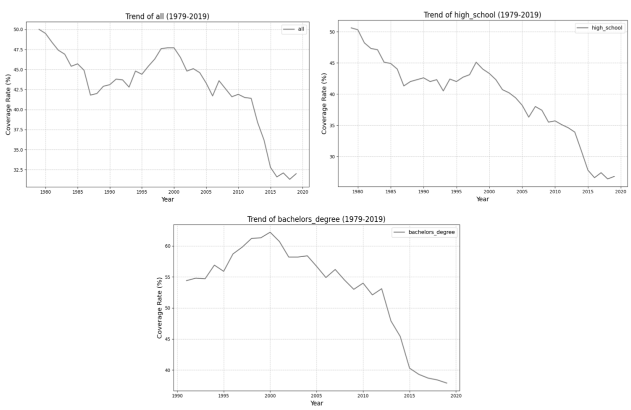

Figure 2. The trend distribution of used variables.

3.2. The Inception Residual RNN-LSTM Hybrid Model

1) Preliminaries of the RNN

Recurrent Neural Networks (RNNs) [38–40] are a class of neural networks designed to process sequential data by leveraging their inherent temporal dynamics. Unlike feedforward neural networks, RNNs have loops in their architecture, allowing them to maintain a hidden state that carries information about previous inputs. This enables RNNs to capture temporal dependencies, making them suitable for tasks such as time series prediction, language modeling, and speech recognition. By processing input one step at a time and updating their hidden state, RNNs can model sequences of varying lengths, providing a flexible framework for sequence-to-sequence tasks like machine translation.

The architecture of an RNN includes a recurrent layer where the hidden state is updated iteratively. At each timestep, the hidden state is computed as a function of the current input and the previous hidden state, typically using an activation function such as tanh. While this design allows RNNs to learn dependencies across timesteps, they struggle with long-term dependencies due to the vanishing gradient problem during backpropagation through time (BPTT). This limitation arises when gradients diminish exponentially as they are propagated backward, making it difficult for the network to learn relationships between distant timesteps. Despite these challenges, RNNs have been successfully applied to tasks like sentiment analysis, video captioning, and simple sequence generation, forming the foundation for more advanced architectures like LSTMs and GRUs. Their ability to model temporal structures has made them a cornerstone of sequence modeling in machine learning.

2) Preliminaries of the LSTM

Long Short-Term Memory (LSTM)[41–43] networks, introduced by Hochreiter and Schmidhuber in 1997, are a type of recurrent neural network (RNN) designed to overcome the vanishing gradient problem faced by traditional RNNs. This issue arises during backpropagation in deep networks when gradients diminish exponentially as they are propagated backward through many layers, making it difficult to learn correlations between distant events. LSTMs address this challenge by incorporating memory cells capable of retaining information over extended periods. Unlike standard feedforward neural networks, LSTMs include feedback connections that make them well-suited for processing sequences of data. These features make LSTMs highly effective for tasks such as time series prediction, natural language processing, and speech recognition.

The architecture of an LSTM unit includes three gates: the input gate, forget gate, and output gate, each playing a distinct role in managing information flow. The input gate determines which new values can update the memory, using a sigmoid activation layer to filter relevant inputs and a tanh layer to create candidate values for addition. The forget gate enables the model to discard irrelevant information, ensuring the network focuses only on useful data. Finally, the output gate regulates how much of the current memory state contributes to the unit's output. These gates collectively allow LSTMs to efficiently model sequences with varying intervals and lengths, enabling applications such as stock market trend prediction, text generation, and even music composition. Their ability to connect past information with current tasks, such as predicting movement in videos using previous frames, sets them apart as a powerful tool for sequence modeling.

3) Preliminaries of the residual connection

Residual connections, also known as skip connections, are a significant innovation in neural network architecture that address the challenges of training very deep networks. As networks grow deeper, issues such as vanishing gradients can hinder effective training, where gradients become too small to update parameters meaningfully. Residual connections mitigate this by providing shortcuts that allow gradients to flow directly through the network, bypassing one or more layers. This ensures that critical information is retained and gradients remain robust, making training deep architectures more feasible.

Introduced by He et al. in their groundbreaking ResNet paper[44], residual connections revolutionized deep learning by enabling the construction of much deeper networks than previously possible. The principle is straightforward yet transformative: instead of learning a direct mapping, each layer learns the residual, or difference, between its input and output. Mathematically, this is expressed as F(x)+x, where x is the input, and F(x) represents the transformation applied by the layer. By adding the transformation back to the original input, this setup creates a direct path for gradients during backpropagation, addressing the vanishing gradient problem. As a result, layers only need to learn incremental adjustments, simplifying the training process and improving network efficiency.

4) The architecture of the proposed model

Our proposed architecture combines the strengths of RNNs and LSTM networks, leveraging their respective capabilities to process sequential data effectively. The structure begins with parallel RNN and LSTM modules, each consisting of 64 neurons, which independently process the input sequence and extract diverse features. The outputs from these modules are then concatenated to form a unified feature representation, capturing both short-term and long-term dependencies. To enhance the combined representation, a dense layer with 64 neurons is applied, followed by a residual connection mechanism. This residual connection maps the RNN output directly to the combined feature representation, ensuring that critical features from the RNN branch are retained and effectively integrated. By adding the residual representation back to the processed features, the architecture ensures smoother gradient flow and improved feature learning during training.

3.3. Implementation Details

The model is developed using the TensorFlow framework, configured to train for 80 epochs with a batch size of 32. The Adam optimizer is utilized to enhance the training process by efficiently adjusting the weights based on the gradients. Mean Squared Error (MSE) is employed as the loss function, which quantifies the average squared difference between the predicted and actual values, ensuring effective optimization.

4. Results and Discussion

4.1. The Performance of the Proposed Inception Residual RNN-LSTM Hybrid Model

Figure 3 provides the training curves of the three models including Inception residual RNN-LSTM hybrid model, LSTM model, and RNN model. It illustrates the progression of loss values across 80 epochs. The Inception residual RNN-LSTM hybrid model demonstrates a rapid decline in loss during the first 10 epochs, followed by a steady decrease as it converges, reflecting its efficiency in optimizing the loss function. Similarly, the LSTM model shows a sharp drop in loss during the initial epochs but exhibits fluctuations around the 20th epoch before stabilizing. The RNN model, while also experiencing a steep decline in loss during the early stages, converges more slowly compared to the other two models. Overall, the Inception residual RNN-LSTM hybrid model achieves the most stable convergence, indicating its superior ability to effectively capture both short-term and long-term dependencies. In contrast, the LSTM and RNN models show less consistent loss reduction patterns, highlighting the potential benefits of the hybrid architecture.

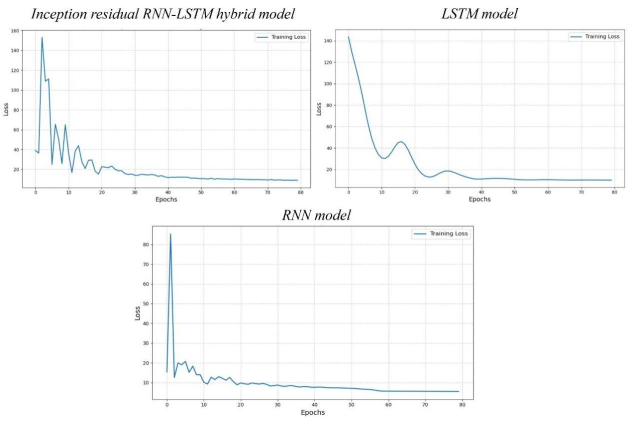

Figure 3. The training curves of different models.

We initially experimented with traditional machine learning models, including Decision Tree, K-Nearest Neighbors (KNN), and Linear Regression, to predict coverage rates (**Figure 4**). However, the results were unsatisfactory, with the models failing to accurately capture the trends in the data, particularly in more complex scenarios. The performance metrics in **Table 1** reveal high MSE, MAE, and RMSE for these models, coupled with negative or extremely low R^2 values. The poor results can be attributed to the inherent limitations of traditional machine learning algorithms in capturing temporal dependencies and nonlinear relationships present in sequential data. For example, the predicted values often deviate significantly from actual values, especially for the "bachelors_degree" and "high_school" features. To address these shortcomings, we employed deep learning models, including RNN, LSTM, and our proposed Inception residual RNN-LSTM hybrid model (**Figure 5**). The deep learning models demonstrated a notable improvement, effectively capturing both short-term and long-term dependencies in the sequential data. Among them, our hybrid model outperformed LSTM and RNN, achieving the lowest MSE, MAE, and RMSE, along with a higher R^2 value of 0.33 (**Table 1**). This superior performance can be attributed to the fusion of features from RNN and LSTM branches and the integration of residual connections, which facilitate efficient gradient flow and enhance the model's ability to learn complex patterns. The training curves (**Figure 3**) further highlight the hybrid model's stability during optimization. Compared to LSTM and RNN, it converges faster and maintains lower loss throughout the training process. The comparative visualization of model performance in **Figure 6** underscores the hybrid model's advantage, as it consistently surpasses traditional and other deep learning models across all metrics.

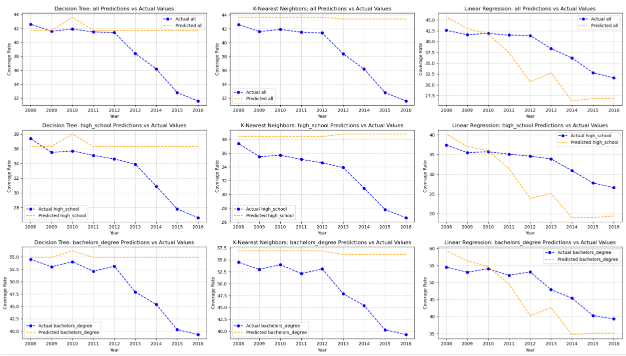

Figure 4. The predicted results compared to real values in the testing dataset based on traditional machine learning models.

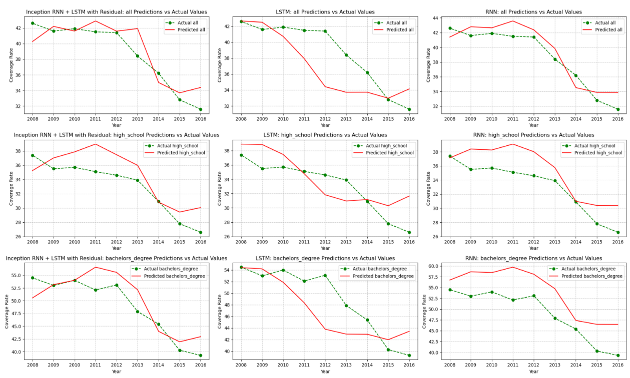

Figure 5. The predicted results compared to real values in the testing dataset based on various deep learning methods.

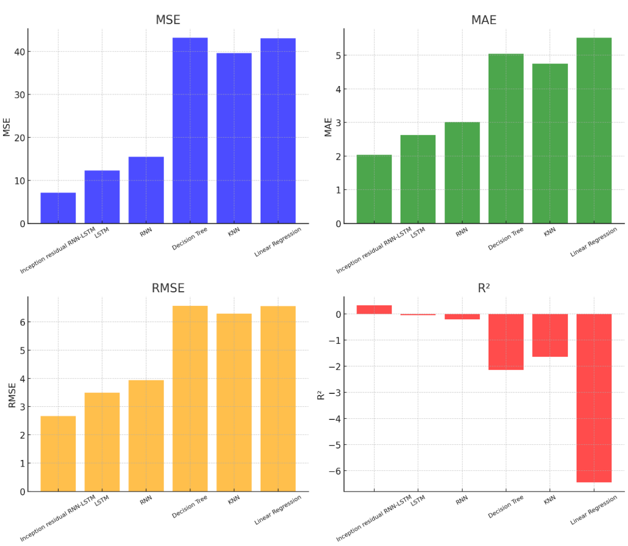

Figure 6. The visualization comparison among different models.

4.2. The Impact of the Number of Neurons in the Fused Output Layer on the Results

We investigated the impact of the number of neurons in the fused output layer of the proposed Inception residual RNN-LSTM hybrid model on its performance (**Table 2**). The study evaluated neuron configurations of 16, 32, 64, and 128, examining their effects on several metrics.

The results reveal that the choice of neuron count significantly influences the model's performance. A configuration with 64 neurons achieved the best results. This indicates that 64 neurons provide the optimal balance between model complexity and predictive capability, effectively capturing the relationships in the data without overfitting. In contrast, both smaller and larger neuron counts exhibited inferior performance. When the neuron count was reduced to 16 or 32, the metrics slightly worsened, likely due to the reduced capacity to model complex patterns. For example, the configuration with 32 neurons resulted in a higher MSE (9.84) and RMSE (3.13), along with a lower R^2 value (0.28). On the other hand, increasing the neuron count to 128 led to a dramatic degradation in performance, with an MSE of 26.04 and an R^2 of -0.55, suggesting severe overfitting. These findings highlight the importance of carefully tuning the number of neurons in the fused output layer. Selecting an appropriate configuration ensures that the model effectively generalizes unseen data while avoiding the pitfalls of underfitting or overfitting.

4.3. Ablation Study

The ablation study presented in **Table 3** examines the contributions of different components in the proposed Inception residual RNN-LSTM hybrid model. By isolating individual elements and comparing their performance, we can assess the impact of specific design choices. When comparing the LSTM and RNN models independently, their performance metrics indicate limited effectiveness in capturing the complexities of the data, with higher MSE and RMSE values. Integrating these components into an Inception RNN-LSTM hybrid model without residual connections improves performance significantly, reducing MSE to 8.63 and achieving an R^2 value of 0.37. This highlights the benefits of combining RNN and LSTM branches, as they complement each other in learning short-term and long-term dependencies. Further

Table 1. The performance of different approaches in the testing dataset.

Model Name	MSE	MAE	RMSE	R^2
Inception residual RNN-LSTM hybrid model	**7.13**	**2.04**	**2.67**	**0.33**
LSTM model	12.30	2.63	3.50	−0.04
RNN model	15.52	3.02	3.94	−0.21
Decision tree model	43.18	5.04	6.57	−2.15
KNN model	39.64	4.75	6.29	−1.64
Linear regression model	43.04	5.52	6.56	−6.45

Table 2. The performance of different numbers of neurons in the fused output layer of the proposed model.

The Number of Neurons	MSE	MAE	RMSE	R^2
16	7.38	2.08	2.71	0.30
32	9.84	2.33	3.13	0.28
64	**7.13**	**2.04**	**2.67**	**0.33**
128	26.04	4.11	5.10	−0.55

enhancing this hybrid model with residual connections leads to the final Inception residual RNN-LSTM hybrid model, which achieves the best overall performance. The residual connections improve gradient flow and allow the model to efficiently learn adjustments, reducing MSE to 7.13 and RMSE to 2.67.

Table 3. The ablation analysis of the proposed model.

Model Name	MSE	MAE	RMSE	R^2
LSTM model	12.30	2.63	3.50	−0.04
RNN model	15.52	3.02	3.94	−0.21
Inception RNN-LSTM hybrid model	8.63	2.25	2.93	**0.37**
Inception residual RNN-LSTM hybrid model	**7.13**	**2.04**	**2.67**	0.33

4.4. Discussion

While the Inception residual RNN-LSTM hybrid model demonstrates significant improvements over traditional machine learning methods and other deep learning architectures, it is not without limitations. One primary shortcoming lies in its computational complexity. The fusion of RNN and LSTM branches, coupled with the addition of residual connections, increases the model's parameters and training time. This makes the model less efficient for deployment in real-time applications or resource-constrained environments. Future work could explore methods to reduce model complexity, such as pruning or quantization, without sacrificing performance.

Another limitation is the model's reliance on extensive hyperparameter tuning. As seen in the neuron count optimization for the fused output layer, small changes in hyperparameters can significantly impact performance. Automating the hyperparameter search process, perhaps through Bayesian optimization or reinforcement learning, could make the model more adaptable and efficient. Additionally, the model's performance could be further validated using larger and more diverse datasets. The current results are promising but are limited to specific types of sequential data. Future work should evaluate the model's generalizability across domains such as speech recognition, financial forecasting, or healthcare data [45–48]. Furthermore, the model could be extended to handle multi-modal data by incorporating additional input types, such as images or categorical features, to enhance its application scope. Lastly, interpretability remains a challenge. While the residual connections improve learning efficiency, they also add to the model's complexity, making it harder to understand the contribution of individual components. Developing interpretability methods, such as attention mechanisms or feature attribution techniques, could provide insights into the model's decision-making process and improve trustworthiness for critical applications.

5. Conclusions

This study introduces the Inception residual RNN-LSTM hybrid model to enhance the accuracy of pension coverage trend predictions. By leveraging the complementary strengths of RNN and LSTM architectures within an Inception-like structure, and integrating residual connections, the model effectively addresses the limitations of traditional and standalone deep learning models. Benchmarking results validate its ability to outperform existing approaches across multiple metrics, demonstrating its applicability for sequential data prediction. Despite its computational complexity and dependence on hyperparameter tuning, the model sets a solid foundation for future advancements. Future work will focus on optimizing the model's efficiency, improving interpretability, and validating its generalizability across diverse datasets and domains.

References

[1] Barr, N., 2006. Pensions: Overview of the issues. Oxford Review of Economic Policy. 22(1), 1–4.

[2] Boldrin, M., Dolado, J.J., Jimeno, J.F., et al., 1999. The future of pensions in Europe. Economic Policy. 14(29), 288–320.

[3] Gustman, A.L., Mitchell, O.S., Steinmeier, T.L., 1994. The role of pensions in the labor market: A survey of the literature. ILR Review. 47(3), 417–438.

[4] Clark, R.L., Craig, L.A., Wilson, J.W., 2003. A history of public sector pensions in the United States. University of Pennsylvania Press. 2003 May 12.

[5] Poterba, J.M., Venti, S.F., Wise, D.A., 2007. The changing landscape of pensions in the United States. National Bureau of Economic Research: Cambridge, MA, USA.

[6] Skocpol, T., Abend-Wein, M., Howard, C., et al., 1993. Women's associations and the enactment of mothers' pensions in the United States. American Political Science Review. 87(3), 686–701.

[7] Waine, B., 2006. Ownership and security: individualised pensions and pension policy in the United Kingdom and the United States. Competition & Change. 10(3), 321–337.

[8] Salazar, J.D., del Carmen Boado-Penas, M., 2019. Scoring and prediction of early retirement using machine learning techniques: application to private pension plans. In: Anales del Instituto de Actuarios Españoles. No. 25, pp. 119–145.

[9] Vafin, E., Razumovskaya, E., Kiselev, S., et al., 2024. METHODOLOGICAL ASPECTS OF THE STUDY OF THE RESOURCE POTENTIAL OF PENSION SYSTEM: MODELING AND FORECASTING. Brazilian Journal of Law & International Relations/Relações Internacionais no Mundo. 2(44).

[10] Zhu, B., Zhang, L., Zhang, Y., 2021. Research on forecast model of elderly pension expenses based on data mining algorithm. Journal of Physics: Conference Series. 1852(3), 032051.

[11] Zhang, Z., 2024. RAG for Personalized Medicine: A Framework for Integrating Patient Data and Pharmaceutical Knowledge for Treatment Recommendations. Optimizations in Applied Machine Learning. 4(1).

[12] Wilson, A., Ma, J., 2025. MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection. Optimizations in Applied Machine Learning. 5(1).

[13] Lu, P.M., Zhang, Z., 2025. The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering. Journal of Computational Biology and Medicine. 5(1).

[14] Zhou, Z., Wu, J., Cao, Z., et al., 2021. On-Demand Trajectory Prediction Based on Adaptive Interaction Car Following Model with Decreasing Tolerance. In 2021 International Conference on Computers and Automation (CompAuto) 2021 Sep 7 (pp. 67–72). IEEE.

[15] Jihu, L., 2024. Green Supply Chain Management Optimization Based on Chemical Industrial Clusters. arXiv preprint arXiv:2406.00478.

[16] Jia, Y., Lei, J., 2024. Experimental Study on the Performance of Frictional Drag Reducer with Low Gravity Solids. Innovations in Applied Engineering and Technology. 1–22.

[17] Lei, J., 2024. Efficient Strategies on Supply Chain Network Optimization for Industrial Carbon Emission Reduction. arXiv preprint arXiv:2404.16863.

[18] Zhao, Z., Ren, P., Tang, M., 2024. How Social Media as a Digital Marketing Strategy Influences Chinese Students' Decision to Study Abroad in the United States: A Model Analysis Approach. Journal of Linguistics and Education Research. 6(1), 12–23.

[19] Zhang, H., Zhu, D., Gan, Y., et al., 2024. End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection. Journal of Information, Technology and Policy. 1–7.

[20] Zhang, G., Zhou, T., Cai, Y., 2023. CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures. Journal of Computational Methods in Engineering Applications. 1–7.

[21] Zhao, Y., Dai, W., Wang, Z., et al., 2024. Application of computer simulation to model transient vibration responses of GPLs reinforced doubly curved concrete

panel under instantaneous heating. Materials Today Communications. 38, 107949.

[22] Zhu, Y., Zhao, Y., Song, C., et al., 2024. Evolving reliability assessment of systems using active learning-based surrogate modelling. Physica D: Nonlinear Phenomena. 457, 133957.

[23] Chen, X., Wang, M., Zhang, H., 2024. Machine Learning-based Fault Prediction and Diagnosis of Brushless Motors. Engineering Advances. 4(3).

[24] Grossberg, S., 2013. Recurrent neural networks. Scholarpedia. 8(2), 1888.

[25] Medsker, L.R., Jain, L., 2001. Recurrent neural networks. Design and Applications. 5(64–67), 2.

[26] Salehinejad, H., Sankar, S., Barfett, J., et al., 2017. Recent advances in recurrent neural networks. arXiv preprint arXiv:1801.01078.

[27] Graves, A., Graves, A., 2012. Long short-term memory. Supervised sequence labelling with recurrent neural networks. 37–45.

[28] Hochreiter, S., 1997. Long Short-term Memory. Neural Computation MIT-Press.

[29] Schmidhuber, J., Hochreiter, S., 1997. Long short-term memory. Neural Comput. 9(8), 1735–1780.

[30] Szegedy, C., Ioffe, S., Vanhoucke, V., et al., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI conference on artificial intelligence 2017 Feb 12 (Vol. 31, No. 1).

[31] Jastrzębski, S., Arpit, D., Ballas, N., et al., 2017. Residual connections encourage iterative inference. arXiv preprint arXiv:1710.04773.

[32] Gan, Y., Ma, J., Xu, K., 2023. Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model. Journal of Computational Methods in Engineering Applications. 1.

[33] Gan, Y., Chen, X., 2024. The Research on End-to-end Stock Recommendation Algorithm Based on Time-frequency Consistency. Advances in Computer and Communication. 5(4).

[34] Ma, J., Chen, X., 2024. Fingerprint Image Generation Based on Attention-Based Deep Generative Adversarial Networks and Its Application in Deep Siamese Matching Model Security Validation. Journal of Computational Methods in Engineering Applications. 1–3.

[35] Xu, K., Gan, Y., Wilson, A., 2024. Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances. Innovations in Applied Engineering and Technology. pp. 1–12.

[36] Rocha Salazar, J.D., Rocha Salazar, J.D., Boado-Penas, M.D., 2019. Scoring and prediction of early retirement using machine learning techniques: Application to private pension plans. In Anales del Instituto de Actuarios Españoles 2019 Dec 15 (Vol. 25, No. 25, pp. 119–145). Instituto de Actuarios Españoles.

[37] Adekunle, O., Dumontier, M., Riedl, A., 2021. Applied Data Science in the Pension Industry: A Survey and Outlook. Netspar, Network for Studies on Pensions, Aging and Retirement.

[38] Kaggle, 2024. Pension Coverage in the USA (1979–2019). Available from: https://www.kaggle.com/datasets/asaniczka/pension-coverage-in-the-usa-1973-2022

[39] Sutskever, I., 2013. Training recurrent neural networks. Toronto, ON, Canada: University of Toronto. 2013 Jan 1.

[40] Sutskever, I., Martens, J., Hinton, G.E., 2011. Generating text with recurrent neural networks. In Proceedings of the 28th international conference on machine learning (ICML-11) 2011 (pp. 1017–1024).

[41] Sherstinsky, A., 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena. 404, 132306.

[42] Sak, H., Senior, A.W., Beaufays, F., 2024. Long short-term memory recurrent neural network architectures for large scale acoustic modeling.

[43] Danihelka, I., Wayne, G., Uria, B., et al., 2016. Associative long short-term memory. In International conference on machine learning 2016 Jun 11 (pp. 1986–1994). PMLR.

[44] He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770–778).

[45] Tang, Y., 2024. Investigating the Impact of Regional Digital Finance Development on Short-run IPO Performance: Empirical Evidence from China. Journal of Management Science & Engineering Research. 7(2), 31–43.

[46] Tang, Y., Li, C., 2023. Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises. Journal of Computational Methods in Engineering Applications. 2023 Nov 6, 1–7.

[47] Hao, Y., Chen, Z., Sun, X., et al., 2025. Planning of truck platooning for road-network capacitated vehicle routing problem. Transportation Research Part E: Logistics and Transportation Review. 194, 103898.

[48] Zhao, Y., Dai, W., Wang, Z., et al., 2024. Application of computer simulation to model transient vibration responses of GPLs reinforced doubly curved concrete panel under instantaneous heating. Materials Today Communications. 38, 107949.