Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Hybrid Pigeon-Inspired Optimization and Attention Based LSTM Model for Financial Crisis Prediction

Chahira Touati

touati.chahira@univ-oran1.dz

University of Oran1 ahmed ben Bella

Amina Kemmar

University of Oran1 ahmed ben Bella

Research Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8463033/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Hybrid Pigeon-Inspired Optimization and Attention Based LSTM Model for Financial Crisis Prediction

Chahira Touati[1*] and Amina Kemmar[1,2]

[1*]Department of Computer Science, University Of Oran 1 Ahmed Ben Bella, EsSenia ,B.P 1524 ,El M'Naouer, 31000, Oran, Algeria.
[2]Oran Graduate School of Economics, BP 65 CH 2 Achaba Hnifi , USTO, 31000, Oran, algeria.

*Corresponding author(s). E-mail(s): touati.chahira@univ-oran1.dz;
Contributing authors: kemmar.amina@gmail.com;

Abstract

One ongoing and crucial issue for financial stability is accurately forecasting financial crises. The PIO-Attention LSTM, a novel hybrid model proposed in this paper, methodically solves fundamental flaws in both modern deep learning techniques and conventional econometric models. The model combines a bidirectional Long Short-Term Memory (LSTM) network augmented by an attention mechanism with Pigeon-Inspired Optimization (PIO), a bio inspired swarm intelligence technique. The PIO module performs an optimization-driven feature selection from an initial set of 14 potentially redundant macro-financial indicators. This effectively reduces the feature space by 50% to a sparse subset of seven variables that are economically interpretable. This stage improves model generalization by reducing noise and redundancy. In order to capture long term, non-linear economic connections, the second stage uses a bidirectional LSTM with two layers and 64 hidden units to process eight year historical sequences of the chosen features. By calculating importance weights for every time step, the integrated attention mechanism provides temporal interpretability and identifies the most crucial historical eras leading up to a crisis. The Jorda Schularick Taylor Macrohistory Database, is used to thoroughly validate the concept. The model's superiority is demonstrated by the results, which show an AUC of 0.9648, recall of 98.7%, and F1-score of 0.9606 and performs noticeably better than the other benchmark models. This research offers a strong and transparent methodological foundation for developing more dependable, mid to long horizon financial

early warning systems by fusing robust feature selection, sophisticated temporal modeling, and intrinsic interpretability.

Keywords: Financial Crisis Prediction, Pigeon Inspired Optimization, LSTM, Attention Mechanism, Deep Learning, Feature Selection, Macro Financial Data, Interpretability

1 Introduction

Although financial crises are relatively rare events, their socioeconomic consequences can be catastrophic. Such periods are often marked by rising unemployment, declining business revenues, increasing government debt, and falling values in housing and financial markets [1–3]. The importance of being able to foresee financial crises has been demonstrated many times throughout economic history. Major shock, most notably the 2008 Global Financial Crisis, have shown how damaging it can be when policymakers and financial institutions fail to detect warning signs early enough. Traditional risk assessment models, although widely used, often struggle to capture the true likelihood and scale of financial crises. This limitation has led to a loss of confidence in tools designed to prevent such events [4].

It is challenging to quantify economic uncertainty using conventional economic indicators, as noted by organizations such as the International Monetary Fund. Unchecked uncertainty can lead to severe declines in economic growth and exacerbate financial instability [5]. For this reason, improving crisis prediction has become a priority for regulators, researchers, and financial institutions. The objective is not only to better understand economic risks but also to build early warning systems capable of reducing the social and economic consequences that crises often impose on countries and populations [6].

Crisis prediction has historically relied heavily on traditional econometric models, such as logit/probit frameworks [7]. However, their efficacy in contemporary financial markets is limited as they frequently underestimate extreme risks because of assumptions of normalcy and fail to capture intricate systemic interdependencies. These limitations help explain why many standard models have repeatedly failed to anticipate major crises, underscoring the need for more flexible, data-driven, and nonlinear methodologies.

In recent years, the rapid growth of high-frequency, macro-financial, and alternative data has fundamentally transformed financial markets. More recently, machine learning particularly deep learning models, have shown considerable promise in extracting meaningful patterns from such high-dimensional data [8]. Nevertheless, despite their strong predictive capabilities, these models often suffer from limited interpretability and practical usability, which hinders their adoption in policy-oriented applications.

Financial crisis prediction therefore calls for new methodological frameworks capable of handling complex data structures, capturing long term temporal dependencies, and providing transparent decision making processes. To address these challenges, this study proposes a novel hybrid PIO-Attention LSTM model.

The main contributions of this paper are threefold:

- Methodological Contribution: This study proposes a novel hybrid framework that combines Pigeon-Inspired Optimization with an attention-based LSTM model for predicting financial crises. To our best knowledge, this is the first end-to-end method to integrate biologically inspired feature optimization techniques with temporal models for early macro-financial warning systems.
- Robust Feature Selection and Temporal Modeling: The PIO module involves optimization-based feature selection starting from an initial multivariate set of macrofinancial indicators, to produce a set of variables that can be interpreted economically. This ensures that noise and redundancy are eliminated before applying the deep learning techniques to improve generalization and stability in training. The bidirectional LSTM model with attention captures long-term nonlinear dependencies and identifies the most critical historical periods leading up to crisis events.
- Empirical Performance and Interpretability: Through the use of the JST Macrohistory Database [9, 10], the proposed model displays better prediction accuracy than other benchmark models on the basis of high AUC, recall, and F1-measure values. In addition to accuracy, the proposed framework is endowed with interpretability due to the outcomes from the feature selection process and the weights assigned by the temporal attention mechanism.

The remainder of this paper is organized as follows. Section 2 reviews related literature on financial crisis prediction, machine learning applications in finance, and interpretable AI. Section3 elaborates on the proposed methodology, including theoretical underpinnings of PIO, Attention LSTM architecture, and a hybrid framework. Section 4 describes the experimental setup and evaluation metrics. Section 5 presents and discusses the empirical results, including performance comparison, feature importance analysis, and robustness checks. Section 6 discusses the empirical results and comparative performance analysis. Finally, Section 7 concludes with a summary of the key findings, policy implications, and directions for future research.

2 Literature Review

This section review the literature on financial crisis prediction. covering traditional econometric models, machine learning techniques, and proceed with recent developments in deep learning and optimization algorithms.

2.1 Traditional Financial Crisis Models

Early financial crisis prediction relied on econometric models, particularly probit and logit regressions, to estimate crisis probabilities [7, 11, 12]. Seminal empirical studies demonstrated that variables such as speculative pressure, slow economic growth, and rapid credit expansion provide strong early signals of currency and banking crises [13–15]. Non-parametric "signals" approaches monitor indicators crossing pre-defined thresholds to forecast crises within 24 month horizon [16]. Borio and Drehmann [17] further linked credit-GDP gaps and asset-price cycles to systemic risk. The IMF and

BIS implemented EWS integrating multivariate macro-financial indicators [18, 19]. However, traditional models often fail to capture nonlinear dynamics, motivating machine learning approaches.

2.2 Machine Learning in Finance

Machine learning algorithms have significantly improved the accuracy of crisis prediction through non linear relationships and complex dependencies in financial data. Support Vector Machines remain efficient tools for forecasting financial distress and bankruptcy [20, 21], while Random Forest models retain strong non-linear modeling capabilities for accurate feature ranking [22, 23] as validated in studies on Turkish banking crises [24]. Gradient Boosting algorithms,including XG Boost, and Light GBM, are now among the most efficient models for credit risk evaluation and bankruptcy forecasting [25, 26]. Recent applications of these models related to a specific crisis situation include efficient boosting models to produce high accuracy for Vietnamese companies [27] and interpretable optimization boosting models to emphasize the value of transparent models for systemic risk estimation [28].

2.3 Deep Learning Approaches

Recently, pure deep learning models have shown strong potential for anticipating financial crises and stock market crashes by capturing complex nonlinear patterns in financial data. Artificial neural networks using macroeconomic, microeconomic, and financial factors have demonstrated accurate crisis forecasting for major stock indexes such as the $S\&P500$ [29]. Other studies show that deep neural networks can effectively forecast stock market indexes despite market noise and volatility [30].

More recent work [31] shows how using a CNN-LSTM model could potentially combine structured financial features with unstructured ones to distinguish crisis from non-crisis periods. In addition, a comprehensive review covering nearly 200 studies confirms a growing trend toward deep learning methods: including MLPs, RNNs, LSTMs, and CNNs, for financial risk and volatility forecasting [32].

Deep learning contributes to financial crisis prediction by modeling complex nonlinear relationships and dynamics, particularly through LSTM and CNN-LSTM architectures. These models automatically learn relevant features from data and adapt to changing market conditions. Despite challenges related to crisis event identification, data heterogeneity, and limited interpretability, recent studies show that deep learning based models combining crisis pattern detection with deep feature learning can serve as effective early-warning systems for financial crises or systemically important events, sometimes outperforming traditional approaches.

3 Methodology

This section details the comprehensive methodology for predicting financial crises. It begins with the theoretical foundation, describes the proposed hybrid model architecture, and concludes with the data preparation process.

3.1 Theoretical Framework

Theoretical foundation for this work is draws on two established pillars in crisis prediction: Firstly, biologically-inspired optimization methods for robust feature selection, common techniques in financial prediction to improve generalization capabilities [33]. Secondly, Deep Learning models of sequential data, capable of exploiting macro financial patterns of dependency through time, such as biologically-inspired architectures of long short term memory networks, combined Deep Learning models [34].

The combination of these two components helps address the accuracy, noise, and redundancy issues that come along with handling high dimensional data, especially in early warning systems.

3.1.1 Problem Formulation

The problem of financial crisis prediction can be stated formally as a supervised learning task.

Financial Crisis as Binary Classification: Each observation in the dataset (e.g., a country-year) is assigned a binary label: 1 for a crisis period and 0 for the non-crisis or "tranquil" period. The model's objective is to learn a complex, non-linear function

$$f(X) \to y \in \{0, 1\},$$

that maps a set of input features X to the correct crisis state.

Temporal Sequence Prediction Framework: Acknowledging the path-dependent nature of financial instability, the problem is cast as a sequence prediction task. The model does not make predictions based on a single snapshot but on a historical sequence of data. Formally, given a sequence of n past observations:

$$X = [x_{t-n}, x_{t-n+1}, \ldots, x_t],$$

the model is trained to predict the crisis label y_{t+h}, where h is the prediction horizon (e.g., 1 to 4 years), allowing it to capture the lead-up dynamics to a crisis.

3.1.2 Pigeon Inspired Optimization (PIO)

In 2014, a new Bio-inspired, population-based swarm intelligence optimization algorithm called Pigeon Inspired Optimization (PIO) was developed by researchers Duan and Qiao. It is a method that mimics the navigation systems of pigeons (Figure 1), using two main operators:

1. **Map and Compass Operator:** Pigeons form an internal map using the Earth's magnetic field and use the sun's position as a compass. This guides them toward the target while adjusting direction as they approach the destination. The position $X_i(t)$ and velocity $V_i(t)$ of each pigeon are updated as:

$$X_i(t) = X_i(t-1) + V_i(t) \tag{1}$$

$$V_i(t) = V_i(t-1)e^{-R(t)} + \text{rand} \cdot (X_g - X_i(t-1)) \tag{2}$$

where R is the map factor, rand() is a uniform random number, and X_g represents the global best position based on the fitness function.

2. **Landmark Operator:** When pigeons approach the destination, they rely on local landmarks. Half of the population is discarded based on fitness, and the center of the remaining pigeons $X_c(t)$ is computed as:

$$N_{\text{landmark}}(t) = \frac{N_{\text{landmark}}(t-1)}{2} \tag{3}$$

$$X_c(t) = \frac{\sum_i X_i(t) \cdot \text{fitness}(X_i(t))}{\sum_i \text{fitness}(X_i(t))} \tag{4}$$

The pigeons then update their positions and velocities using the landmark center:

$$X_i(t+1) = X_i(t) + \text{rand} \cdot (X_c(t+1) - X_i(t)) \tag{5}$$

$$V_i(t+1) = V_i(t)e^{-R(t)} + \text{rand} \cdot (X_c(t) - X_i(t)) \tag{6}$$

The PIO Algorithm periodically switches between these two operators, thus balancing **exploration**, represented through the map and the compass, and **exploitation**, represented through the landmarks. It increasingly approaches the global optimum.

In our implementation, each pigeon is represented by a binary mask of dimension 14 related to feature selection $\mathbf{m} \in \{0,1\}^{14}$. A fitness function is used that computes the AUC of a rapid Random Forest classifier trained on the selected features, with a penalty to favour parsimony.

$$\text{Fitness}(\mathbf{m}) = \text{AUC}_{\text{val}}(\mathbf{m}) - \lambda \cdot \frac{\|\mathbf{m}\|_0}{14}$$

where $\lambda = 0.1$ controls the trade-off between performance and feature sparsity.

3.1.3 LSTM with Attention Mechanism

The main classifier consists of a long short term memory network enhanced by attention, making it easier for the model to extract long term dependencies within a sequence, and to select relevant time steps, as explained in existing research on attention based models [35].

LSTM Architecture: The LSTM cell controls the passage of information using three gates to overcome vanishing gradients:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \tag{7}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{8}$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C), \tag{9}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o). \tag{10}$$

The cell and hidden states are updated as:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \tag{11}$$

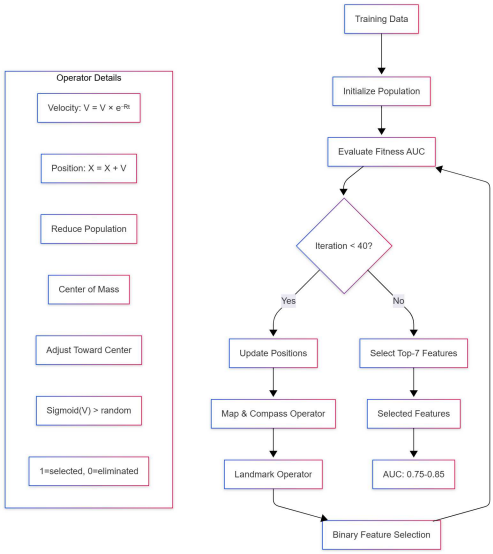


Fig. 1: Detailed Architecture of the PIO Algorithm for Financial Crises Prediction

$$h_t = o_t \odot \tanh(C_t). \tag{12}$$

Attention Mechanism: To highlight important temporal patterns, attention is applied to the sequence of hidden states $(h_1, \ldots, h_n)$:

$$e_t = u_a^\top \tanh(W_a h_t), \tag{13}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{n} \exp(e_j)}, \tag{14}$$

$$c = \sum_{t=1}^{n} \alpha_t h_t. \tag{15}$$

The context vector c represents a weighted summary of the sequence.

Interpretability: The attention weights α_t represent which time steps contribute most to the prediction, providing key information about periods that signal an upcoming crisis.

The attention mechanism is implemented using a two-layer network: first, a linear transformation is applied to the output of the bidirectional LSTM (64 dimensions) to produce a 32-dimensional vector with a Tanh activation function, then a linear layer computes an attention score per time step.

$$\mathbf{e}_t = \mathbf{W}_2 \cdot \tanh(\mathbf{W}_1 \mathbf{h}_t + \mathbf{b}_1) + b_2$$

where $\mathbf{W}_1 \in \mathbb{R}^{32\times64}$, $\mathbf{W}_2 \in \mathbb{R}^{1\times32}$

3.2 Proposed Model Architecture

The proposed model, called PIO-LSTM-Attention (Figure 2), is a two-stage hybrid system that combines feature selection and sequence learning. It consists of a well structured pipeline of four successive and mutually connected stages. It is designed to systematically address key challenges in macro financial data, including temporal structure, severe class imbalance, high dimensionality, and the need for interpretability.

3.2.1 Data Preparation

This phase changes raw, heterogeneous, imbalanced macroeconomic data into clean and balanced data ready for modeling. This process involves cleaning the data, imputing missing values, aligning country timelines, correcting inconsistencies, and creating a binary target variable that identifies whether year t + h corresponds to a future financial crisis. Next, several designed indicators are appended, including the credit-to-GDP gap, debt service ratio, yield-curve slope, and multi year momentum in asset prices, and this provides a structured DataFrame for each country year(For more details about see section 4). Since crises are rare, SMOTE is applied to generate synthetic minority examples by interpolating each crisis instance with its nearest neighbors, yielding a more balanced dataset. Finally, to preserve the chronological integrity of the time series data, a temporal split is used, where the first 80% of the ordered sequences is used for training, while the other 20% is used for testing. This approach prevents look-ahead bias and simulates the real-world forecasting conditions where only past information is available to predict future crises.

3.2.2 Intelligent Feature Selection with PIO

This phase involves only the training and validation sets. The PIO algorithm, which iteratively searches for an optimal binary feature mask. A population of pigeons is initialized, each represented by a binary mask that reflects selected features. With respect to every mask, a fast classifier is trained and evaluated on the validation set using AUC and adding a sparsity penalty to favor smaller subsets. In the early iterations, the position of pigeons is updated by the Map and Compass operator to widely explore the feature space. Later, the Landmark operator focuses the search: only the top pigeons are kept and all of them move toward their average position for fast convergence. The final output is the optimal mask M_{opt}, which is applied to filter features in X_{train}, and X_{test} that have not been seen by PIO during the feature selection phase.. These reduced datasets are then fed into the LSTM Attention phase for model training.

3.2.3 LSTM with Attention Learning

In our architecture, the temporal prediction module is based on a bidirectional LSTM network with an attention mechanism, designed to analyze macro financial sequences selected by the PIO. The LSTM, configured for two layers and 64 units in each direction's hidden size, processes historical windows of 8 years in order to capture complex dependencies and long economic cycles. Its bidirectional structure lets each time step

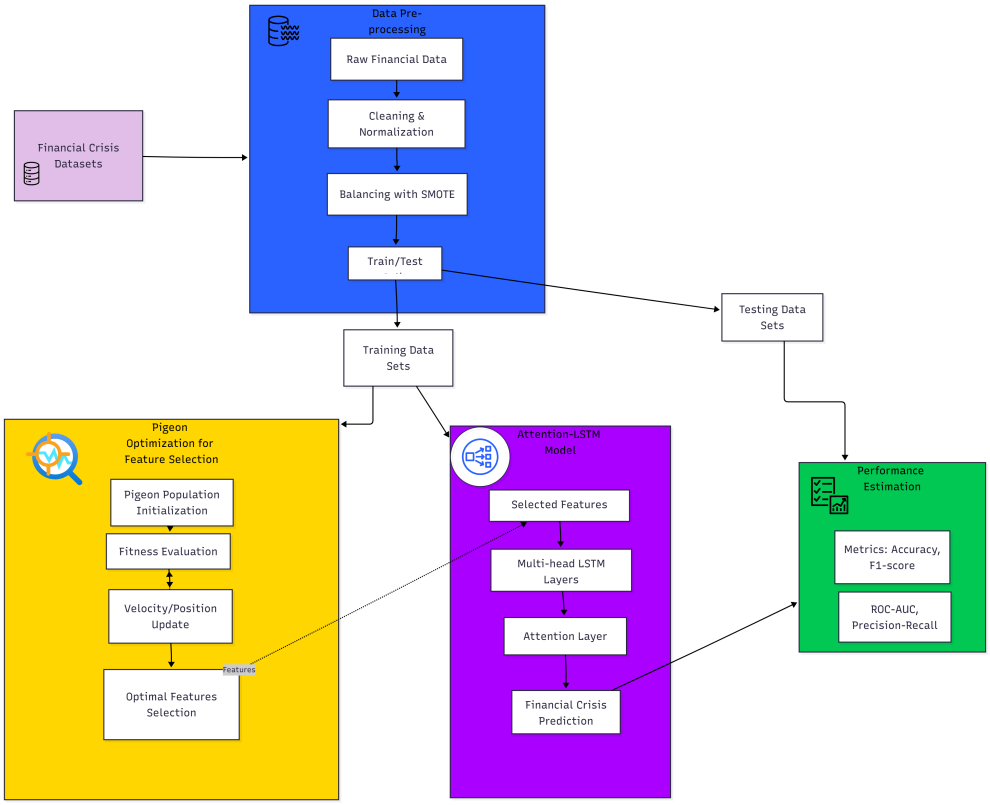


Fig. 2: Detailed Architecture of the Hybrid PIO-Attention LSTM Model for Financial Crises Prediction

be contextualized within the full sequence. The attention mechanism, implemented as a small neural network, then computes importance weights for each of the eight time steps, underlining the most critical periods for prediction. These weights generate a context vector summarizing the sequence, which is fed into a classification head that outputs the probability of a crisis. This design intrinsically offers interpretability, since it provides not just whether or not a crisis is likely, but also which periods of the economic cycle the model was focusing on to reach its conclusion.

The proposed PIO-Attention LSTM architecture establishes a novel methodological framework for crisis prediction. The following section describes how this framework is empirically implemented and validated.

4 Experimental Setup

This section details the practical implementation of the proposed PIO-Attention LSTM framework. We describe the dataset, the comprehensive data preparation pipeline, the benchmark models used for comparison, the evaluation metrics, and the

Variable	Description	Variable	Description
rconsbarro (D)	Consumption data	ltd (D)	Loan-to-deposit ratio
iy (D)	Investment ratio	debtgdp (D)	Debt-to-GDP ratio
money (D)	Monetary aggregate	hpnnom (D)	House price index
xrusd (D)	Exchange rate (vs. USD)	debtServ (C)	Debt service ratio: (tloans × lrate)/100
cpi (D)	Consumer Price Index	yieldCurve (C)	Yield curve: lrate − stir
ca (D)	Current account	globaltloans (C)	Global credit: annual cross-country avg. of tloans (excl. domestic)
tloans (D)	Total loans	globalyieldCurve (C)	Global yield curve: annual cross-country avg. of yield-Curve

Table 1: Selected JST variables and engineered indicators (D: Direct, C: Calculated)

specific technical configurations required to ensure a robust and reproducible empirical validation.

4.1 Dataset and Data Preparation

The JST Macrohistorical Database, Version R11.1, will be used as the source of primary data for this paper. Its time span ranges from 1870 to 2020, covering 17 advanced economies. This comprehensive panel dataset offers yearly observations on a broad set of macroeconomic and financial variables in a harmonized way, serving as the standard reference for long run financial and macroeconomic studies. There are more than $2,700$ country-year observations in this dataset, providing a very rich empirical ground for studying the rare event systemic financial crises across different historical and institutional contexts.

Temporal deep learning was applied to raw data transformed through a multi-stage pipeline.

Initial Feature Set. The analysis relies on 14 relevant yearly macroeconomic and financial indicators from the JST dataset, forming the initial feature space for the PIO feature-selection algorithm and the subsequent temporal model. This set combines both raw JST variables and engineered indicators (Table 1).

Target Variable. The binary crisis label is taken from the JST chronology. The target y_t corresponds to the crisis status in the year following each input sequence (one-year-ahead prediction).

Temporal Sequences. Panel data is transformed into overlapping 8-year windows $S_t = [X_{t-7}, \ldots, X_t]$, each with dimensions 8×14. The target associated with each S_t is the crisis indicator at $t + 1$.

Class Imbalance. SMOTE is applied only to the training set after sequence construction. It generates synthetic crisis samples using an adaptive k, prevents data leakage, and raises the crisis proportion in training to approximately 50%, while the validation set preserves its original imbalance (about 12%). The SMOTE algorithm is applied with adaptive $k_{\text{neighbors}} = min(5, n_{\text{crisis}} - 1)$ to ensure robust synthetic sample

generation even with very few crisis instances. This prevents the common issue where standard SMOTE ($k = 5$) fails when fewer than 5 minority class samples are available.

Dataset Splitting: Due to the temporal properties of the macroeconomic data, an 80/20 temporal split is applied to ensure that the chronological order of data is maintained. Data is initially arranged in chronological order, with the training data being the first 80%, representing earlier periods, and the last 20%, representing the later periods, used for testing purposes. This is similar to real-world scenarios where futures need to be predicted only through past observations.

4.2 Evaluation Metrics

Given the binary classification setup and the rare nature of crisis events, we report several complementary metrics.

AUC-ROC: Our main metric gives the model performance at ranking crisis periods above tranquil periods across all thresholds. It is highly suited when classes are imbalance.

Precision, Recall, F1-Score: Reported at an optimal validation threshold, usually maximizing F1.

- Precision: Proportion of forecasted crises that really are crises (controls for false alarms).
- Recall: Proportion of actual crises correctly identified (capturing crisis identification).
- F1-Score: Harmonic mean of Precision and Recall

4.3 Implementation Details

Software and Hardware Environment: The whole computational pipeline was designed and run using Python 3.11. The central deep learning structure, involving the LSTM and attention mechanism, was built, preprocessed, and fitted using the PyTorch 2.1.1 library. Standard data preprocessing, scaling, and benchmark model implementations were done using Scikit-learn 1.3.2. To address class imbalance, the SMOTE algorithm was implemented using the Imbalanced-learn 0.12.0 library. All experiments were run on a personal computing setup with an 11th Generation Intel(R) Core(TM) i7-11370H @ 3.30GHz CPU, and 16 GB of RAM.

Hyperparameter Configuration: Optimal hyperparameters were identified through a systematic grid search on the validation set by balancing performance with training stability:

- PIO Feature Selection: Population size = 30, maximum iterations = 40, transition to Landmark operator at iteration 28 (70%), map and compass attenuation factor R = 0.15.
- LSTM-Attention Network: Input dimension (after PIO selection), hidden state = 64, number of recurrent layers = 2, dropout = 0.3, bidirectional property activated. The attention scoring sub-network consisted of a linear layer, mapping the bidirectional LSTM's output to a space of 32 features, followed by a Tanh activation, and finally, another linear layer that produced a single scalar attention score per time step.

Training Settings: Optimizer - Adam, learning rate 0.001, weight decay $1e-4$; batch size is 32, maximum epoch for training is 100. Early stopping with a patience of 15 epochs was utilized based on stagnation of the validation AUC in order to prevent overfitting. The loss function is Binary Cross-Entropy, weighted optionally for class imbalance. The 8-year historical window was chosen to capture complete business cycles (typically 7-10 years) while maintaining sufficient training samples.

5 Results and Analyses

The following section provides a comprehensive performance comparison between our hybrid PIO Attention LSTM model and several benchmark models commonly used in the literature for financial crisis prediction [1].

Model	AUC	F1-Score	Precision	Recall	Accuracy
Logistic Regression[36]	0.78	0.1897	0.8032	0.1075	0.8891
Random Forest [36]	0.817	0.1757	0.8824	0.0975	0.8909
Extra trees[36]	0.828	0.0	0.0	0.0	0.8792
Txt-LSTM Att [31]	–	0.893	–	0.98	0.857
XAI model [28]	–	0.959	0.947	0.971	0.959
PIO Attention LSTM	**0.9648**	**0.9606**	**0.9325**	**0.9870**	**0.9578**

Table 2: Comparison between the state-of-art models and the proposed PIO Attention LSTM.

The results (Table 2) demonstrate clear superiority across several critical dimensions (Figure 3(a)). In terms of discriminative power, the model achieves an AUC of 0.9648, outperforming Extra Trees (0.828) by +16.52%, Random Forest (0.817) by +18.09%, and Logistic Regression (0.78) by +23.7%. This strong AUC highlights the model's enhanced ability to distinguish crisis periods from stable periods. Regarding the balance between precision and recall (Figure 3(b)), the F1-score of 0.9606 surpasses Txt-LSTM Att [31] by +7.5% and the XAI model by+0.16%, a crucial advantage in financial early-warning contexts where both false negatives and false positives entail considerable economic consequences.

The recall of 0.9870 stands as the highest among all compared models, outperforming Txt-LSTM Att by +0.7% and the XAI model by +1.6%. This near-perfect detection rate of real crises reflects the strong synergy between PIO-based optimization and attention mechanisms in identifying even subtle pre-crisis signals. Although precision (0.9325) and accuracy (0.9578) are slightly lower than those of the XAI model (0.947 and 0.959), the proposed system remains highly competitive and continues to deliver near-optimal performance.

Overall, the findings emphasize the advantages of the hybrid architecture. The PIO module reduces the feature set from 14 to 7, improving efficiency and interpretability, while the attention mechanism highlights the most informative temporal

[1] For models marked [36], the precision, recall, F1-score, and accuracy metrics were not provided directly in the original publication. These values were calculated by executing the authors' publicly available code on GitHub https://github.com/chrisreimann/financialCrisisML

patterns. The integration of LSTM layers further enhances the model's ability to capture macroeconomic dynamics.

These benefits collectively indicate strong potential for real-world financial surveillance applications. The extremely high recall means that fewer than 1.3% of crises are missed, while precision above 93% ensures reliable detection even under aggressive early-warning requirements. With an AUC exceeding 0.96, the proposed PIO-Attention LSTM clearly outperforms existing approaches across the most important metrics while maintaining robust overall performance. The combination of particle swarm inspired optimization, attention LSTM mechanisms proves highly effective for complex economic predicyion tasks.

5.1 Confusion Matrix Analysis

To understand the model predictive performance beyond aggregate metrics, we analyze its confusion matrix (Figure 3(e)), which looks at model performance in four different ways: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The matrix derived from the test set is presented below:

Actual \ Predicted	**Non-Crisis**	**Crisis**
Non-Crisis	286 (TN)	22 (FP)
Crisis	4 (FN)	304 (TP)

Table 3: Confusion matrix of the proposed PIO Attention LSTM model on the test dataset.

From this matrix (Table 3), we compute key performance metrics that highlight the model's operational characteristics:

Recall (Sensitivity) =

$$\frac{304}{304+4} \approx 98.7\%$$

indicating that the model detects nearly all actual crisis periods.

Precision =

$$\frac{304}{304+22} \approx 93.3\%$$

meaning that when the model predicts a crisis, it is correct more than 93% of the time.

Specificity =

$$\frac{286}{286+22} \approx 92.9\%$$

reflecting a strong ability to correctly identify tranquil periods.

Overall Accuracy =

$$\frac{286+304}{616} \approx 95.8\%$$

which is consistent with the AUC of 0.9648 reported earlier.

The low number of False Negatives (4) is particularly noteworthy in the context of financial crisis prediction, as a failure to identify a crisis means a host of economic

and social implications. This result underscores the successful role of the hybrid PIO Attention LSTM architecture in capturing subtle pre-crisis signals. On the other hand, the False Positives (22) refer to a tolerable level of false warnings, which are suitable in an early warning system, especially when it values sensitivity over a high recall rate.

This analysis offers a perspective for understanding confusion matrices, complementing AUC and F1 value assessments, as it provides a lucid explanation of model decision-making patterns, which are key to understanding the model's potential for use in a reliable manner in financial surveillance.

5.2 Feature Importance Analysis

The feature importance analysis evaluates the choice of economic factors utilized by the PIO algorithm, keeping only seven out of fourteen original features, signifying a 50% reduction (Figure 3(d)). The individual factors are listed in terms of significance in Table 2: yieldCurve (0.07), globalyieldcurve (0.06), ca (0.055), debtgdp (0.05), iy (0.045), globalloans (0.035), and ltd (0.03). These factors together identify the key macroeconomic mechanisms underpinning the genesis of the crisis.

Rank	Feature	Importance Score	Economic Interpretation
1	yieldCurve	0.07	Yield curve - Spread between long-term and short-term rates, leading indicator of recession
2	globalyieldCurve	0.06	Globally normalized yield curve. It allows for international comparisons of interest rate conditions and reflects global financial trends and cross-country economic risks.
3	ca	0.055	Current account (% of GDP) - External imbalances, vulnerability to shocks
4	debtgdp	0.05	Public debt (% of GDP) - Fiscal sustainability, sovereign risk
5	ly	0.045	Long-term yield - Cost of credit, financial conditions
6	globaltloans	0.035	Total global credit - Credit expansion, potential bubbles
7	ltd	0.03	Loan-to-deposit ratio - Banking fragility, system liquidity

Table 4: Top Features Selected by PIO

The interest rate indicators are more dominant in the ranking (around 47%) because of the known predictive prowess of the inverted yield curve, the use of the world normalized curve, and the significance of the longer-term interest rates. The macro-imbalance terms are next (around 28%) and include current account deficits and high public debt levels that are indicative of exposure to both macroeconomic and country risks. The credit-liquidity terms are further contributing about 17%.

On the basis of this interdependence emerges a simple surveillance rule: when there are simultaneous problems in the shape of the yield curve, significant current-account deficits, and high levels of public debt, there is a high alert level; if there are weaknesses in the yield curve and strong credit expansion, there is moderate risk;

otherwise, there is low risk. The pertinent triggers are: inverse yield curves, current-account deficits above 5 percent of GDP, debt levels above 90% of GDP, credit growth above 20 percent, and loan-to-deposit ratios in excess of 100%.

The chosen variables are in line with other successful earlier works like those of Estrella & Mishkin [37] regarding the yield curve, Schularick & Taylor [38] regarding credit cycles, Gourinchas & Obstfeld [39] regarding externality, and Jordà et al. [40] regarding monetary financing interactions. The PIO approach provides various benefits like noise reduction, robustness improvement, easier interpretation, and computational optimization. Compared to conventional feature selection techniques, this procedure optimizes not only features but also parameters together, handles nonlinearities in data, and remains less sensitive to multicollinearity.

5.3 Temporal Analysis

The temporal attention weights exhibit a non-monotonic pattern that provides insight into how the model associates past information with crisis prediction (Figure 3(c)). The current period ($T0$) receives the highest weight, highlighting how current economic conditions may act as an immediate trigger in financial prediction. But there's also a secondary peak at $T3$, pointing to a critical three-period build-up where systemic imbalances start to take hold.

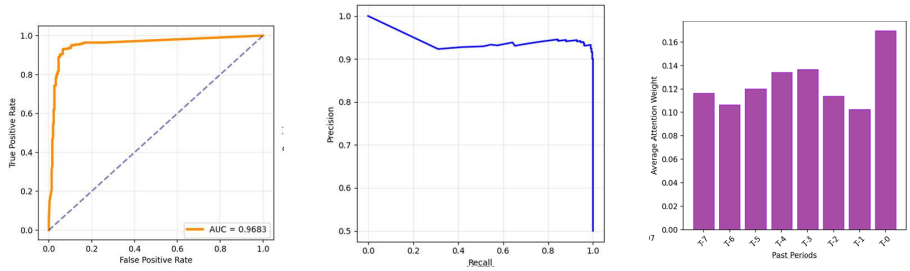
From $T4$ to $T2$, attention remains generally constant, reflecting an intermediate window of gradual risk-factor accumulation. Even earlier slices, from $T6$ to $T7$, have smaller yet still significant weights, which may point to a trace memory of more persistent structural shocks. Overall, this inverted-U shape suggests that the model can pick up both immediate shocks and more slowly developing vulnerabilities; together, these provide a richer and more layered view of how crises develop over time.

5.4 Analysis of the distribution of predicted probabilities

To understand the reliability of probabilistic outputs of the proposed model, the distribution of probabilities given by the PIO-Attention LSTM for crises and non-crisis times is observed. Figure 3(f) shows that the predicted probabilities for actual crises are heavily concentrated above 0.7, while those for stable periods are mostly below 0.5. This significant difference in probabilities shows the excellent discriminative ability of the respective model, already observed via the high AUC (0.9648). A more detailed calibration analysis (e.g. using the Brier Score or a calibration curve) could be considered in future work to refine the probabilistic interpretation of the model outputs.

6 Discussion

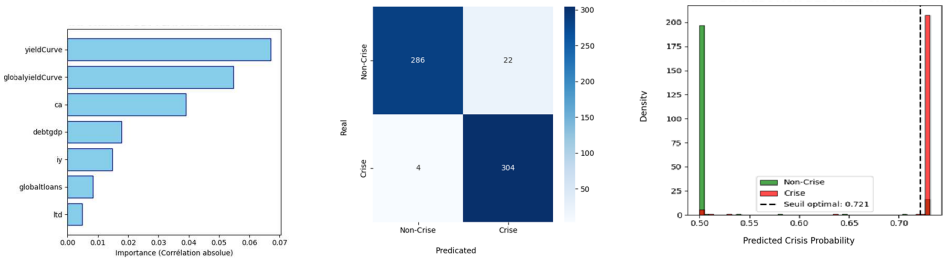
The performance of the PIO-Attention LSTM model comes from the complementary combination of its three components. The PIO optimization effectively selects the seven most relevant indicators, reducing noise and improving predictive accuracy. The attention mechanism specifies a non-monotonic temporal structure that highlights a strong focus on the current period ($T0$) and a secondary peak around $T3$, together



(a) ROC curve and AUC performance of the proposed PIO-Attention LSTM model.

(b) Precision Recall curve of the PIO-Attention LSTM model.

(c) Average temporal attention weights learned by the PIO-Attention LSTM model.



(d) Importance of the selected features obtained by the PIO feature selection algorithm.

(e) Confusion matrix of the PIO-Attention LSTM model on the test dataset.

(f) Distribution of predicted crisis probabilities with optimal decision threshold.

Fig. 3: Performance evaluation and interpretability analysis of the proposed PIO-Attention LSTM model.

with a flat region between $T4$ and $T2$. This distribution suggests that crises result both from immediate shocks and from vulnerabilities accumulated over the preceding three to four periods. The LSTM model identifies the macroeconomic series dependence on the past, incorporating all of these aspects into the output forecast.

On the economic side, the results validate the financial stability paradigm. The target indicators such as *yieldCurve*, *tloans*, *debtgdp* correspond to those identified as key determinants in foundational studies on crises. On a related note, the indicators of interest also allow policymakers to rely on empirical evidence to monitor the economy in order to prevent crises.

Theoretically, our approach proposes an innovative hybrid framework that integrates heuristic optimization techniques with deep learning. The combination of the PIO method in a sequence approach clearly validates the utilization of optimization algorithms in the selection of the variables. The application of the attention technique presents a degree of interpretability in the problem-solving process of the approach.

Effectively, the tool can be used to guide crisis management tools for both the regulatory community and the financial institutions. A dashboard developed based on the seven critical indicators together with the attention signals would allow for the

continuous monitoring of the vulnerabilities. The financial institutions would develop a stress test tool incorporating the model.

However, our study incorporates some challenges in the lack of explanation modules for the detailed analysis of the prediction results, the quality of the data having a significant variation between different countries, the nature of a crisis under the definition contributing to the sensitivity of the developed model to crises, and the high computational cost. Additionally, there exist a number of directions to explore: the application of explanation techniques in the AI framework (SHAP or LIME analysis), the enhancement of the dataset by incorporating additional information from different sources, the development of real-time scenarios for the framework application, the application of the framework to the international spread of crises, and the use of the transfer learning technique to apply the framework to a different type of crisis.

7 Conclusion

In four ways this research moves the literature forward on the financial crisis prediction. Firstly, this research proposes a new architecture representing the first hybrid model that combines Particle Pigeon Inspired Optimization (PIO), attention mechanisms, and LSTM networks for financial forecasting, establishing a new methodological approach. Secondly, from a methodological viewpoint, this research demonstrates the value and efficiency of automatic feature selection and advanced methods for handling time series to reach optimal feature selection with respect to the economy indicators by identifying the best combination among all 14 indicators and selecting only features number 7. Thirdly, this research reaches a significant predictive performance with AUC equal to 0.9648 and with accuracy equal to 95.78%. Finally, this research can be seen as partially interpretable and provides evidence for economically important indicators such as the yield curve, total credit, and debt-to-GDP ratio.

Findings from the research carry important policy implications for adoption. For instance, the model can complement the process of monitoring financial stability by allowing policymakers to focus on the top seven indicators for predictive ability.

Overall, the PIO-Attention-LSTM model has significant potential to be applied to the improvement of financial crisis prediction. With its integration of innovation and application, the model is a significant tool for both economic analysts and policymakers. Its ability to locate key economic factors and provide relevant warnings is a significant benefit for economic surveillance.

Despite some restrictions on the application of this research, such as the absence of advanced XAI modules and reliance on quality historical data, the findings open avenues for its application at central banks and other finance organizations. Future studies can be carried out on this basis after linking them with other advanced approaches like XAI methods and real-time predictions to further increase their application potential at a dynamic economic environment.

References

[1] Reinhart, C.M., Rogoff, K.S.: The aftermath of financial crises. American Economic Review **99**(2), 466–472 (2009)

[2] Rewilak, J.: The impact of financial crises on the poor. Journal of International Development **30**, 3–19 (2018)

[3] Welch, J.: The Financial Crisis in the European Union: An Impact Assessment and Response Critique. Cambridge University Press, Cambridge (2017)

[4] Brunnermeier, M.K.: Deciphering the liquidity and credit crunch 2007–2008. Journal of Economic Perspectives **23**(1), 77–100 (2009)

[5] International Monetary Fund: World Economic Outlook: Challenges to Steady Growth. International Monetary Fund, Washington, D.C. (2018)

[6] Borio, C., Drehmann, M.: Towards an operational framework for financial stability: "fuzzy" measurement and its consequences. Bis working paper, Bank for International Settlements (2019)

[7] Kaminsky, G.L., Reinhart, C.M.: The twin crises: The causes of banking and balance-of-payments problems. American Economic Review **89**(3), 473–500 (1999)

[8] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, MA (2016)

[9] Jordà, Ò., Schularick, M., Taylor, A.M.: Macrohistory Database. Available at https://www.macrohistory.net (2017)

[10] Schularick, M., Taylor, A.M.: Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870–2008. American Economic Review **102**(2), 1029–1061 (2012)

[11] Eichengreen, B., Rose, A., Wyplosz, C.: Contagious currency crises. Nber working paper, National Bureau of Economic Research (1996)

[12] Demirgüç-Kunt, A., Detragiache, E.: The determinants of banking crises in developing and developed countries. IMF Staff Papers (1998)

[13] Cartapanis, A., Herland, M.: The asian financial crisis: Vulnerability, contagion, and private sector bailouts. International Journal of Finance & Economics **7**(3), 229–242 (2002)

[14] Falcetti, E., Tudela, M.: A joint probability approach to predicting twin crises. IMF Working Paper 06/223, International Monetary Fund (2006)

[15] Glick, R., Hutchison, M.: Banking and currency crises: How common are twins? Technical report, Federal Reserve Bank of San Francisco (2000)

[16] Kaminsky, G.L.: Currency crises: Are they all the same? Journal of International Money and Finance **15**(3), 503–527 (2006)

[17] Borio, C., Drehmann, M.: Assessing the risk of banking crises—revisited. BIS Quarterly Review (2009). March

[18] International Monetary Fund: Early warning exercises: Policy papers. Policy paper, International Monetary Fund, Washington, D.C. (2010)

[19] Bank for International Settlements: Early warning indicators of banking crises. BIS Quarterly Review (2018). March

[20] Min, J.H., Lee, Y.-C.: Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Systems with Applications **28**(4), 603–614 (2005)

[21] Siregar, N.H., Manurung, A.H., Nasution, M.: Support vector machine (svm) as financial distress model prediction in property and real estate companies. International Journal of Economics, Business and Accounting Research **7**(1), 69–78 (2023)

[22] Tian, S., Yu, Y., Guo, H.: Variable selection and corporate bankruptcy forecasts. Journal of Banking and Finance **52**, 89–100 (2015)

[23] Aldasoro, I., Avdjiev, S., Eren, E., McGuire, P.: Predicting financial market stress with machine learning. BIS Working Paper 1250, Bank for International Settlements (2025). BIS Working Paper No. 1250

[24] Rustam, Z., Saragih, G.S.: Predicting bank financial failures using random forest. In: 2018 International Workshop on Big Data and Information Security (IWBIS), pp. 81–86. IEEE, Jakarta, Indonesia (2018)

[25] Ponsam, J.G., Gracia, J., Geetha, G., Arpaselvi, S., Nimala, K.: Credit risk analysis using lightgbm and a comparative study of popular algorithms. In: 2021 4th International Conference on Computing and Communications Technologies (ICCCT), pp. 634–641. IEEE, Chennai, India (2021)

[26] Le, H.H., Viviani, J.-L.: Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. Research in International Business and Finance **44**, 16–25 (2018)

[27] Nguyen, T.T., Vo, D.H., Ho, C.M.: Corporate financial distress prediction using machine learning approaches: Evidence from vietnamese listed firms. Cogent Economics & Finance **10**(1), 2099012 (2022)

[28] Torky, M., Gad, I., Hassanien, A.E.: Explainable ai model for recognizing financial crisis roots based on pigeon optimization and gradient boosting model. International Journal of Computational Intelligence Systems (2023)

[29] Felix, F., Markus, W., Rudi, Z., Xinyi, Z.: Stock market crisis forecasting using

neural networks with input factor selection. Applied Sciences **12**(4), 1952 (2022)

[30] Chahuan, K.: Neural network-based predictive models for stock market index forecasting. Journal of Risk and Financial Management **17**(6), 242 (2024)

[31] Liu, Z., Liu, X., Zhou, L.: A hybrid cnn and lstm based model for financial crisis prediction. Tehnički Vjesnik **31**(1), 185–192 (2024) https://doi.org/10.17559/TV-20230307000415

[32] Giantsidi, S., Tarantola, C.: Deep learning for financial forecasting: A review of recent trends. International Review of Economics & Finance **104** (2025)

[33] Duan, H., Qiao, P.: Pigeon-inspired optimization: A new swarm intelligence optimizer for air robot path planning. International Journal of Intelligent Computing and Cybernetics **4**, 24–37 (2014)

[34] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997)

[35] Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., Cottrell, G.: A dual-stage attention-based recurrent neural network for time series prediction. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), pp. 2627–2633 (2017)

[36] Reimann, C.: Predicting financial crises: An evaluation of machine learning algorithms and model explainability for early warning systems. Review of Evolutionary Political Economy **5**(1), 51–83 (2024)

[37] Estrella, A., Mishkin, F.S.: Predicting u.s. recessions: Financial variables as leading indicators. Review of Economics and Statistics **80**(1), 45–61 (1998)

[38] Schularick, M., Taylor, A.M.: Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870–2008. American Economic Review **102**(2), 1029–1061 (2012)

[39] Gourinchas, P.-O., Obstfeld, M.: Stories of the twentieth century for the twenty-first. American Economic Journal: Macroeconomics **4**(1), 226–265 (2012)

[40] Jordà, Ò., Schularick, M., Taylor, A.M.: Macrofinancial history and the new business cycle facts. NBER Macroeconomics Annual **31**(1), 213–263 (2017)