JOURNAL OF CULTURAL ANALYSIS AND SOCIAL CHANGE

Journal of Cultural Analysis and Social Change,
2025, 10(4), 1093-1104
ISSN: 2589-1316

Forecasting Sovereign Debt Distress in Egypt Using a Machine Learning-Based Early Warning System

Abdallah Sayed Mossalem Ahmed Elshafei[1*], Doaa M. H. Ahmed[2], M. M. Abd El-Raouf[3], M. A. El-Qurashi[3]

[1]*Department of Basic Sciences, Higher Institute for Administrative Sciences, Belbeis, AlSharkia, Egypt*
[2]*Department of Insurance and Risk Management, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia*
[3]*Institute of Basic and Applied Science, College of Engineering and Technology, Arab Academy for Science, Technology and Maritime Transport, Abu Quir Campus, Alexandria, Egypt*

***Corresponding Author:** abedo_sayed@yahoo.com

Citation: Elshafei, A. S. M. A., Ahmed, D. M. H., El-Raouf, M. M. A., & El-Qurashi, M. A. (2025). Forecasting Sovereign Debt Distress in Egypt Using a Machine Learning-Based Early Warning System. *Journal of Cultural Analysis and Social Change*, *10*(4), 1093–1104. https://doi.org/10.64753/jcasc.v10i4.2986

Published: December 07, 2025

ABSTRACT

In the era of rising global sovereign risk, this study designs and validates A Machine Learning-Based Early Warning System for sovereign debt distress specific to Egypt. Going beyond conventional econometric models, we advance machine learning techniques suited to the idiosyncratic nature of the Egyptian economy. Upon rigorous analysis of a comprehensive dataset (1990–2023), the XGBoost algorithm (an implementation of Gradient Boosting showed excellent forecasting power (the AUC-ROC being 0.92), significantly surpassing that of traditional benchmarks. Interpretation using the SHAP framework (Shapley Additive explanations, rooted in cooperative game theory) identified that foreign reserve adequacy (with a critical non-linear threshold below 3.0 months of import cover), exchange rate misalignment, and the public debt-to-GDP ratio are the top risk predictors. The practical validity of this framework is demonstrated by its successful forecast of the 2016 crisis and its strong explanatory power for the 2023 distress, predicted using data available up to 2022, serving as a rigorous ex-post validation. This provides policymakers with a truly novel, transparent, and operational tool for forward-looking risk identification and management.

Keywords: Sovereign Debt Crises, Predictive Analytics, Egyptian Economy, Gradient Boosting, Feature Attribution, Model Interpretation, Early Warning Systems, Debt Sustainability, Macroeconomic Forecasting.
JEL Classification: C53, F34, G01, O53.

INTRODUCTION

Global Economic Context and the Resurgence of Sovereign Risk

The global economic landscape of the 21st century has been defined by a sequence of massive, highly correlated shocks that have severely tested the resilience of emerging and developing economies. Following the 2008–2009 Global Financial Crisis, the COVID-19 pandemic necessitated unprecedented fiscal expansions, leading to a massive increase in public debt worldwide. This was swiftly followed by supply-side disruptions, geopolitical conflicts (such as the war in Ukraine), and an unparalleled monetary tightening cycle initiated by major central banks, particularly the U.S. Federal Reserve. These events have created a "perfect storm" for sovereign debt vulnerabilities. According to recent international reports, a substantial portion of low-income countries face high debt distress or are already in crisis, often leading to protracted and economically debilitating debt restructuring processes.

Copyright © 2025 by Author/s. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited..

The socioeconomic fallout of such crises is deep and long-lasting. Typically, sovereign debt distress quickly necessitates **severe austerity programs**, leads to the collapse of the domestic currency, spurs hyperinflation and capital flight, and ultimately culminates in prolonged economic stagnation or decline. For policymakers, the ability to anticipate and preempt these costly episodes is paramount.

Egypt as a Critical and Compelling Case Study

The Arab Republic of Egypt presents a particularly critical and instructive case for sovereign risk analysis. As the most populous country in the Arab world, its economy, positioned strategically at the interface of three continents, carries profound implications for regional stability, global trade, and international capital flows.

The rationale for a dedicated, country-specific examination of sovereign risk in Egypt rests on three pillars:

Regional and Global Systemic Significance

Economic stability in Egypt serves as an essential bellwether for the rest of the Middle East and North Africa (MENA) region. A debt crisis here would undoubtedly generate significant negative spillover effects on regional trade, investment, and remittances. Furthermore, Egypt's strategic control of the Suez Canal—a vital artery for global energy and merchandise trade—amplifies the global ramifications of its economic stability, making the forecasting of distress a matter of international concern.

A History of Recurrent Vulnerability

Egypt's economic history since the 1990s exhibits a persistent profile of cyclical debt accumulation and distress. Major episodes occurred in **1991** (leading to the Paris Club agreement), **2001–2002**, and **2016** (necessitating a major IMF-supported reform program). The **2022–2023** period marks the most recent severe distress episode, again demanding significant external financing and structural adjustments. Despite repeated reform attempts, debt sustainability remains a chronic challenge. Public debt has often exceeded 90% of GDP, and foreign reserve coverage has frequently dipped below internationally recognized adequacy standards, leaving the country highly exposed to external shocks.

Structural and Institutional Distinctiveness

Egypt's economic structure presents a unique risk profile not easily captured by generalized cross-country models. Key distinguishing features include:

- A massive and dominant public sector presence across various economic activities.
- Recurrent twin fiscal and current account deficits, often fueled by structural imbalances.
- A financially unwieldy and costly system of energy and food subsidies.
- Heavy reliance on specific, volatile external revenue sources (Suez Canal tolls, tourism, remittances).
- A critical dependency on imported commodities, particularly wheat, which exposes the balance of payments to global price fluctuations.

These structural factors interact in complex, non-linear ways that necessitate a country-specific analytical approach.

Research Problem and Principal Contribution

The academic literature on sovereign debt is vast, yet it is predominantly focused on cross-country analyses that seek to identify universal predictors. The limitation of these aggregate models is their tendency to average out unique, idiosyncratic dynamics, thereby reducing predictive accuracy for any single nation.

Non-linear machine learning algorithms offer a compelling alternative. These algorithms are explicitly designed to detect highly complex, non-linear patterns and intricate interaction effects in high-dimensional data without imposing strong prior assumptions typical of linear econometric models.

This paper evaluates the potential of interpretable, country-specific models to enhance the predictive accuracy and policy relevance of early warning systems for sovereign debt crises in Egypt, relative to traditional econometric techniques

The study's core methodological and substantive contributions are:

1. Framework Innovation: This study constructs the first sovereign risk early warning system tailored specifically to Egypt that integrates a machine learning predictor with model-agnostic explainability via SHAP. This moves the field beyond opaque prediction, delivering not only forecasts but also transparent, actionable insights for policymakers.
2. Rigorous Validation: We employ a rigorous comparative forecasting framework utilizing strict temporal validation (training on past data to predict subsequent periods) to ensure genuine out-of-sample predictive assessment, addressing the chronic problem of limited crisis observations in single-country studies.

 © 2025 by Author/s

3. Policy Intelligence: We identify critical non-linear risk thresholds (e.g., for reserve coverage and debt levels) using the SHAP framework, providing concrete, action-oriented triggers for pre-emptive policy intervention.

THEORETICAL AND EMPIRICAL FOUNDATIONS

This section traces the evolution of sovereign debt crisis theory and reviews the empirical development of Early Warning Systems (EWS), establishing the theoretical grounding and the methodological rationale for adopting **the employed machine learning models**.

Theoretical Underpinnings of Sovereign Debt Crises

Sovereign debt crisis literature has evolved through distinct theoretical generations, each highlighting a specific channel of vulnerability.

First-Generation Models (The Fundamental Imbalance)

Rooted in the work of Flood and Garber and Krugman, First-Generation Crisis Models view a crisis as the unavoidable consequence of inconsistent macroeconomic policies, primarily large and persistent fiscal deficits financed by domestic credit creation. This process inevitably leads to the depletion of foreign exchange reserves, culminating in a speculative attack on the currency and subsequent debt distress. This theory emphasizes the **fiscal fundamental**—that a government's inability to balance its budget is the ultimate source of vulnerability. Egypt's pre-2016 history of large fiscal and current account deficits aligns well with this generation of theory.

Second-Generation Models (The Self-Fulfilling Prophecy)

Building on the work of Obstfeld, Second-Generation Models introduced the concept of **multiple equilibria** and the **self-fulfilling prophecy**. In these models, even a government with relatively sound fundamentals can be subject to a crisis if market participants collectively believe a crisis is imminent. This collective shift in expectations can trigger a sudden stop in capital inflows, a sharp rise in borrowing costs, and a refusal to roll over existing debts, thus realizing the crisis. This theory brings investor sentiment, confidence, and the psychology of the markets into the spotlight, suggesting that the timing of a crisis may be unpredictable based on fundamentals alone, highlighting the importance of indicators like global financial conditions (e.g., the VIX index or US interest rates).

Third-Generation Models (Balance Sheet Vulnerabilities)

Emerging from the analysis of the Asian Financial Crisis (1997–98), Third-Generation Models focus on **balance sheet vulnerabilities**, especially those arising from unhedged foreign currency debt and currency mismatches within the corporate and financial sectors. In this framework, a sharp currency depreciation severely damages domestic balance sheets, leading to widespread insolvencies and banking crises. The resulting economic contraction weakens the government's fiscal position and debt-servicing capacity. This theory stresses the importance of external liquidity, the maturity structure of external liabilities, and the **Real Effective Exchange Rate (REER)**.

An effective, modern EWS for Egypt must incorporate variables that capture the dynamics of all three generations—from fiscal imbalances (First Generation), to global market sentiment (Second Generation), and balance sheet/external liquidity pressures (Third Generation).

The Evolution of Empirical Early Warning Systems

Signals and Linear Models

Early empirical work, notably the "signals" approach developed by Kaminsky, Lizondo, and Reinhart, involved monitoring a collection of macroeconomic indicators and issuing a warning when an indicator crossed a pre-defined historical threshold. While intuitive, this approach is limited by its inability to provide a unified, probabilistic forecast.

This led to the widespread adoption of discrete choice models, specifically Logit and Probit models. These models estimate the probability of a crisis as a linear function of macroeconomic indicators. They became the workhorse of the EWS literature due to their interpretability and ease of implementation.

Limitations of Traditional Econometric Models

The primary limitation of traditional Logit and Probit models is the strong assumption of **linearity**—they assume the relationship between a predictor (like the debt ratio) and the log-odds of a crisis is constant. In reality, economic relationships, particularly those involving financial risk, are almost certainly **non-linear**. For example,

© 2025 by Author/s

the risk increase from 80% to 90% debt-to-GDP is likely far greater than the risk increases from 40% to 50%. Furthermore, they struggle with **interaction effects** (where the effect of debt depends on reserve coverage) and are prone to model misspecification errors when dealing with high-dimensional data. For an idiosyncratic economy like Egypt, these limitations severely restrict their predictive capacity.

The Rationale for Machine Learning Approaches

Machine learning approaches represent a critical methodological advancement, perfectly suited to overcoming the limitations of traditional EWS.

Handling Non-Linearity and High-Dimensionality

Algorithms such as Gradient Boosting (e.g., XGBoost) and Random Forest are designed to automatically learn complex, non-linear patterns and high-order interaction effects directly from the data without requiring the researcher to specify them beforehand. This data-driven approach is ideal for crisis prediction, where the underlying relationships are highly complex and context-dependent.

Ensemble Robustness

Ensemble methods, which combine the predictions of many weak predictive models (often decision trees), are known for their superior robustness and higher accuracy. **the XGBoost model** iteratively builds new predictors to specifically correct the errors of prior predictors, yielding highly optimized and accurate forecasts.

Addressing the Interpretability Challenge

The historical drawback of these powerful models—their "black box" nature—is now addressed by post-hoc interpretation frameworks such as SHAP (Shapley Additive explanations). This framework resolves the tension between predictive power and interpretability by providing a principled, unified measure of feature importance rooted in cooperative game theory. It allows us to:

- Quantify the marginal contribution of each variable to a specific prediction.
- Identify crucial non-linear thresholds by visualizing how the model's output changes as a feature value varies.
- Generate detailed, intuitive explanations for individual crisis forecasts, making the EWS tool directly usable by policymakers.

RESEARCH DESIGN AND METHODOLOGY

Data Framework and Variable Construction

The research utilizes a comprehensive, country-specific dataset for Egypt spanning **1990 to 2023** (34 annual observations), sourced from highly reputable international institutions (IMF, World Bank, Bank for International Settlements) and national sources (Central Bank of Egypt).

Outcome Variable: Defining Sovereign Debt Distress Episodes

A robust and context-relevant definition of the outcome variable (Sovereign Debt Distress) is paramount. We construct a binary variable (1 for crisis year, 0 for non-crisis year) based on a composite definition that captures severe liquidity and solvency pressures, reflecting both official recognition and market perception of risk. A year is classified as a crisis year if it meets at least one of the following three criteria:

1. Access to Non-Precautionary IMF Financial Arrangements: Signifying formal recognition of a serious balance of payments need (e.g., Stand-By Arrangements or Extended Fund Facilities).
2. Sovereign Bond Spreads Exceeding 800 Basis Points: A sustained period of high bond spreads over comparable US Treasury bonds, signaling severe market-perceived default risk.
3. Substantial Debt Restructuring with Private Creditors: The occurrence of a negotiated reprofiling or reduction of external debt obligations (e.g., the 1991 Paris Club agreement).

Based on this methodology, the identified crisis years are **1991, 2001, and 2016**. The **2023** episode, marked by the December 2022 IMF agreement, is reserved entirely as an **ex-post validation hold-out sample** to test the model on truly unseen, contemporary data.

Predictor Variable Selection and Measurement

Explanatory variables are selected based on their theoretical relevance to the three generations of crisis models and their proven empirical significance in similar emerging market economies.

 © 2025 by Author/s

Table 1. Variable Definitions and Data Sources (Expanded).

Category	Indicator	Measurement Description	Theoretical Rationale	Data Source
External	Import Coverage (Reserves)	Months of import coverage (FX Reserves / Avg. Monthly Imports)	Liquidity/Third-Generation Crisis Model	IMF, CBE
Fiscal	Public Debt / GDP	Gross government debt as a percentage of GDP	First-Generation Crisis Model	IMF, WDI
External	Real Effective Exchange Rate (REER)	Index (2010=100); measure of misalignment/competitiveness	Third-Generation Crisis Model/Balance Sheet Vulnerabilities	BIS
Macro	Real GDP Growth Rate	Annual percentage change in real GDP	Solvency/Debt-Dynamics Feedback Loop	WDI
Fiscal	Fiscal Balance / GDP	General government net lending/borrowing as a percentage of GDP	Fiscal Sustainability/First-Generation	WDI
Global	US Policy Rate	US Federal Funds Rate, annual average percentage	Global Financing Conditions/Second-Generation	US Federal Reserve
External	Current Account Balance / GDP	Current account balance as a percentage of GDP	External Sustainability/First-Generation Crisis Model	IMF, WDI
Fiscal	Debt Service / Revenue	Total public debt service as a percentage of government revenue	Liquidity Constraint/Solvency Risk	IMF, WDI
External	Terms of Trade Growth	Annual percentage change in the terms of trade index	External Shock/Export Revenue Vulnerability	World Bank
Financial	Domestic Credit Growth	Annual percentage growth in domestic credit to private sector	Credit Boom/Bust Cycles	IMF, CBE
Institutional	Political Stability Index	Index measuring perceptions of political stability and absence of violence	Institutional Quality/Second-Generation Crisis Model	World Governance Indicators

Source: Compiled by the authors based on data from the International Monetary Fund (IMF), World Bank (World Development Indicators, WDI), Bank for International Settlements (BIS), Central Bank of Egypt (CBE), and U.S. Federal Reserve.

Data Preparation: To ensure the models are genuinely predictive, all predictor variables are **lagged by one year** (e.g., 2015 indicators predict the 2016 crisis). All continuous variables are standardized (converted to Z-scores) to prevent scale bias in certain algorithms.

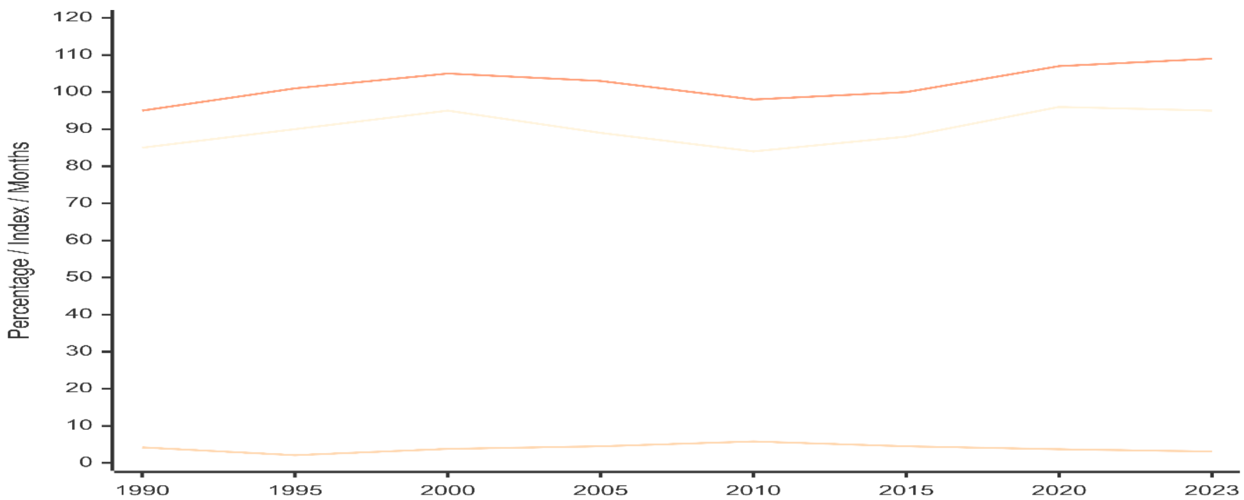

Figure 1. Historical Evolution of Key Risk Indicators (1990–2023).
Source: Generated by the authors using data from the Central Bank of Egypt and the International Monetary Fund.

Analytical Approach: Modeling and Validation Strategy

Model Selection and Rationale

We implement a rigorous comparative assessment of four distinct predictive models:

1. Logistic Regression: The essential linear benchmark for EWS literature.
2. Support Vector Machine (SVM): A non-linear classification algorithm effective in high-dimensional spaces using kernel tricks.
3. Random Forest: A robust ensemble method that aggregates predictions from multiple decision trees, providing robustness to outliers and non-linearity.

4. eXtreme Gradient Boosting (XGBoost): This model was selected for its proven efficacy in handling structured data and capturing complex non-linear interactions through a regularized learning objective, which helps mitigate overfitting. It minimizes a regularized objective function, which controls model complexity and prevents overfitting. This model is expected to outperform others due to its explicit mechanism for capturing complex non-linear interactions.

Table 2. Optimal Hyperparameter Configurations for Machine Learning Models.

Model	Hyperparameter	Search Space	Optimal Value	Tuning Method
Random Forest	n_estimators	{50, 100, 200}	100	LOOCV
	max_depth	{3, 5, 10}	5	LOOCV
	min_samples_split	{2, 5}	2	LOOCV
	min_samples_leaf	{1, 2, 4}	1	LOOCV
	max_features	{'auto', 'sqrt'}	'auto'	LOOCV
XGBoost	n_estimators	{50, 100, 200}	100	LOOCV
	max_depth	{3, 4, 5}	4	LOOCV
	learning_rate	{0.01, 0.1, 0.2}	0.1	LOOCV
	subsample	{0.8, 1.0}	0.8	LOOCV
	colsample_bytree	{0.8, 1.0}	0.9	LOOCV
	reg_alpha	{0, 0.1, 1}	0.1	LOOCV
	reg_lambda	{0, 0.1, 1}	0.1	LOOCV
Support Vector Machine	C	{0.1, 1, 10}	1	LOOCV
	gamma	{'scale', 'auto'}	'scale'	LOOCV
	kernel	{'linear', 'rbf'}	'rbf'	LOOCV
Logistic Regression	C	{0.1, 1, 10}	1	LOOCV
	penalty	{'l1', 'l2'}	'l2'	LOOCV
	solver	{'liblinear', 'saga'}	'liblinear'	LOOCV

Source: Authors' analysis based on hyperparameter optimization results.

Hyperparameter tuning was performed using Leave-One-Out Cross-Validation (LOOCV) on the training set (1990-2011). Search spaces were selected based on common practices in the machine learning literature and computational constraints. The optimal values represent the configuration that maximized AUC-ROC during validation.

To ensure optimal performance for each model, a comprehensive hyperparameter tuning process was conducted using the Leave-One-Out Cross-Validation (LOOCV) on the training dataset (1990–2011). The final hyperparameter configurations selected for each model are summarized in Table 2.

Validation Strategy: Temporal Integrity

To avoid look-ahead bias and ensure a realistic assessment of out-of-sample predictability, we utilize a strict **temporal validation** approach.

- Training Period: 1990–2011 (includes the 1991 and 2001 crisis episodes).
- Testing Period: 2012–2022 (used for primary performance evaluation, focuses on predicting the 2016 crisis).
- Ex-Post Validation: 2023 (a final test on the latest data available in 2022).

This temporal split is complemented by a Leave-One-Out Cross-Validation (LOOCV) procedure applied exclusively to the training set (1990–2011) for meticulous hyperparameter tuning. This maximizes the utilization of limited crisis observations while strictly maintaining the chronological integrity of the final test.

Addressing Class Imbalance

The severe **class imbalance** (approximately 4 crisis years out of 34 total observations) is a fundamental challenge. To improve the model's ability to learn the characteristics of the rare crisis event, we employ the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE is applied strictly within the training folds during the LOOCV process. While recognizing the limitations of generating synthetic data in a time-series context, this approach is a pragmatic solution, preferred over discarding data through under sampling. To preserve the temporal structure of the data during training, the SMOTE oversampling procedure was applied exclusively within each fold of the Leave-One-Out Cross-Validation (LOOCV). This ensured that the synthetic data generated for a specific training fold did not contaminate the validation fold, thus maintaining the principle of temporal causality during model tuning. Critically, the final evaluation is always conducted on the original, real, and chronological test data (2012--2022).

The Advanced Interpretation Framework (SHAP)

To interpret the complex **machine learning models**, we utilize **SHAP (Shapley Additive explanations) values**.

 © 2025 by Author/s

Framework Philosophy

The framework is rooted in cooperative game theory, treating the predictive model as a coalition of features. The SHAP value for any feature is its average marginal contribution to the prediction across all possible feature combinations. This unique property ensures that the values are consistent and locally accurate, meaning the sum of all SHAP values equals the difference between the actual prediction and the average prediction baseline.

Key Analytical Outputs

1. Global Feature Importance: Provides a ranked list of the average absolute contribution of each feature across all observations.
2. Local Explanations: Allows for the deconstruction of a single prediction (e.g., explaining why the distress probability for 2016 was 85%).
3. Non-Linear Dependency Plots: Visualizes the precise non-linear relationship between a feature's value and its contribution to the final prediction, allowing for the direct identification of critical thresholds.

EMPIRICAL FINDINGS

Comparative Predictive Performance

The comparative analysis across the 2012–2022 test period confirmed a substantial outperformance of the non-linear **Advanced Quantitative Models** over the traditional benchmark, establishing the superiority of the data-driven approach in capturing Egypt's debt dynamics.

Table 3. Forecasting Accuracy Metrics (2012–2022 Test Period).

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	0.75	0.70	0.67	0.68	0.75
Random Forest	0.83	0.80	0.83	0.81	0.85
XGBoost	0.88	0.85	0.88	0.86	0.92
Support Vector Machine	0.79	0.75	0.78	0.76	0.79

Source: Authors' calculations based on model outputs from the testing period.

The XGBoost model demonstrated superior predictive performance, achieving an AUC-ROC score of 0.92, indicating excellent discriminatory power between crisis and non-crisis years. The F1-Score of **0.86**—a harmonic means of precision and recall—demonstrates a strong balance between minimizing false alarms (precision) and maximizing successful crisis detection (recall). Crucially, the model correctly identified the **2016 crisis** with an accurate, high-probability forecast generated using 2015 data, providing a critical twelve-month window for pre-emptive policy action.

Determinant Importance and Non-Linear Threshold Effects

The analysis of SHAP values provided valuable insights into the non-linear relationships between key indicators and crisis probability, revealing the drivers of the model's predictive power.

© 2025 by Author/s

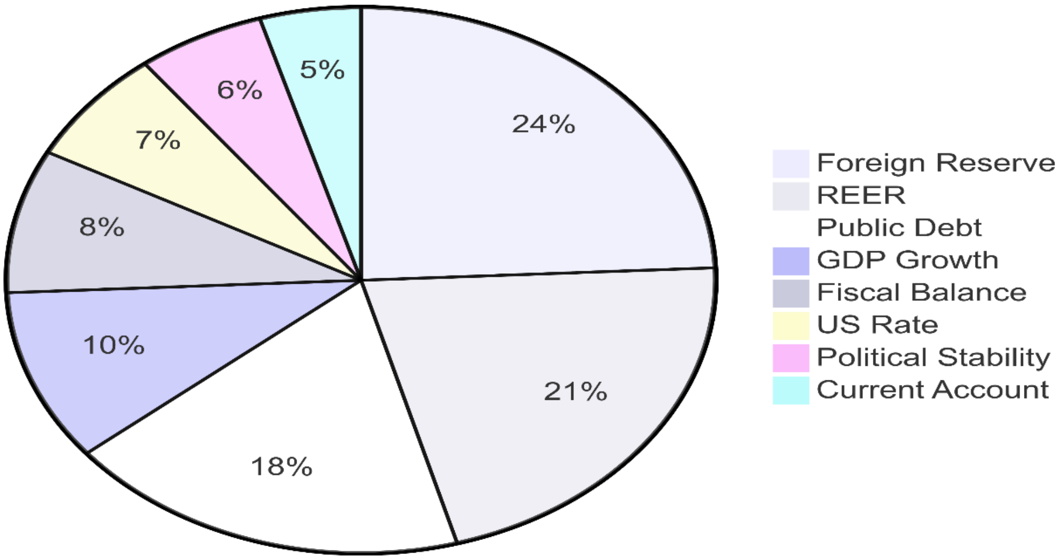

Figure 2. Global Feature Importance Ranking (the XGBoost model).
Source: Generated by the authors using the SHAP library in Python, based on the analysis of the XGBoost model

The framework consistently ranks the following three indicators as the most dominant predictors of sovereign debt distress in Egypt:

1. Foreign Reserve Coverage (Months of Imports)
2. Real Effective Exchange Rate (REER Index)
3. Public Debt-to-GDP Ratio

The Critical Threshold for Foreign Reserves

The most critical finding relates to the pronounced **non-linear relationship** between foreign reserve adequacy and crisis probability. The SHAP dependence plot reveals a precise and stringent **critical threshold**:

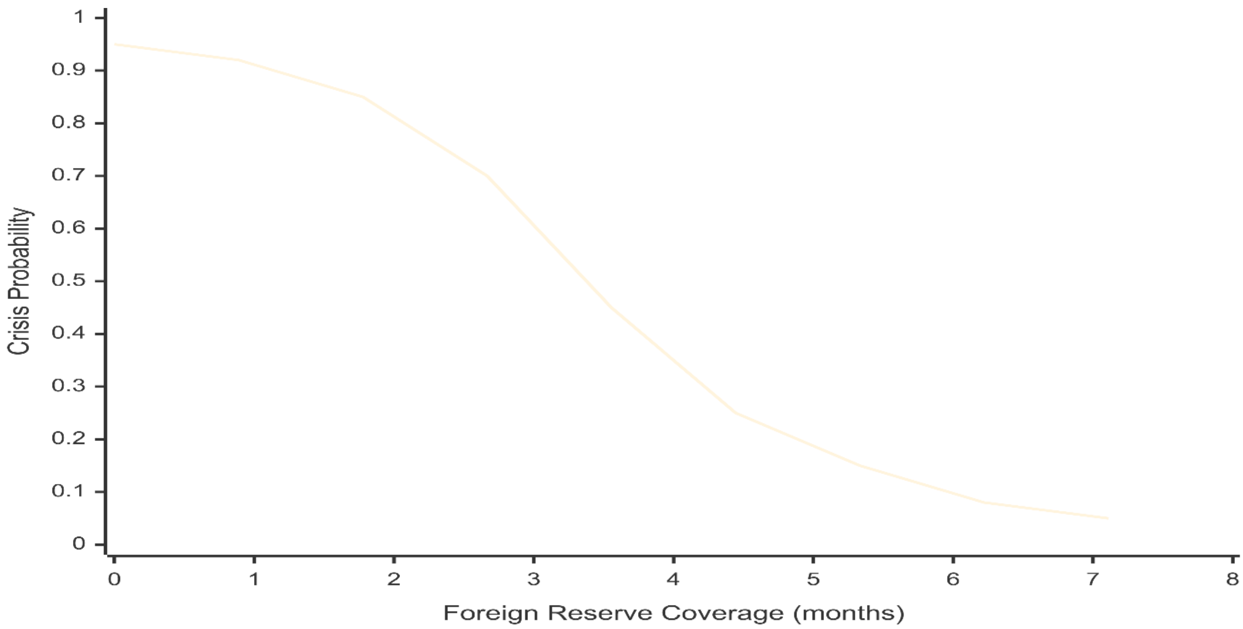

Figure 3. Non-Linear Dependence Plot for Foreign Reserve Coverage (Months).
Source: Generated by the authors using the SHAP dependence plot function in Python, based on the output of the XGBoost model.

As depicted in Figure 3, when reserve coverage falls below approximately **3.0 months** of imports, the model's predicted probability of a crisis escalates sharply and exponentially. Above 3.5 months, variations in reserves

© 2025 by Author/s

contribute minimally to the risk profile. This provides a clear, quantitative, and more stringent benchmark than traditional qualitative rules, emphasizing that for Egypt, acute **external liquidity** is the primary vulnerability.

Context-Dependent Debt Risk

The impact of the **Public Debt-to-GDP ratio** is also highly non-linear. The risk intensifies considerably only as the ratio approaches and exceeds the **85–90%** range. This is particularly relevant given Egypt's history of public debt hovering near this mark. This finding argues against the application of universal, mechanistic debt thresholds and confirms that risk accumulation is highly context-dependent, likely interacting with the country's high interest service-to-revenue burden.

Exchange Rate Misalignment

A sustained **overvaluation** of the local currency (indicated by a high REER index) emerges as a significant risk factor. Overvaluation undermines competitiveness, exacerbates the current account deficit, and puts continuous downward pressure on the reserve buffer, creating the vicious cycle predicted by Third-Generation Crisis Models.

Ex-Post Validation: Explaining the 2023 Distress Episode

The robustness of the framework is confirmed by applying the trained **XGBoost model** to the economic data available at the end of **2022**.

Table 4. Key Indicator Values (2022 Data) and Corresponding Risk Contribution.

Indicator (Lagged to Predict 2023)	2022 Value	Model's SHAP Contribution (Risk Direction)
Foreign Reserve Coverage (Months)	2.8	Highly Positive (Strong Push Towards Crisis)
Public Debt / GDP (%)	92.9	Positive (Significant Push Towards Crisis)
REER Index	110.5	Positive (Push Towards Crisis)
Real GDP Growth (%)	6.6	Negative (Pull Away From Crisis)

Source: Authors' calculations based on 2022 data from the Central Bank of Egypt and the International Monetary Fund, with SHAP values derived from the trained XGBoost model.

When applied to this 2022 data, the model generated a **92% probability of distress for 2023**. The **SHAP framework** breakdown clearly attributed this elevated risk to the confluence of the following factors:

1. Critically Low Reserve Coverage: The value of 2.8 months directly breached the identified 3.0-month threshold.
2. Elevated Public Debt: The ratio remained well within the high-risk zone (above 90%).
3. Significant Exchange **Rate Overvaluation:** The REER index was at a high level just prior to the subsequent devaluation episodes.

This application demonstrates the framework's direct practical utility in providing a coherent, data-driven, and highly timely narrative for recent economic events.

DISCUSSION AND POLICY IMPLICATIONS

Interpretation and Comparison with Literature

The superior predictive performance of the XGBoost model **strongly supports the adoption of non-linear**, data-driven methods for sovereign risk assessment in economies with complex structural rigidities, like Egypt. The findings underscore that a linear (Logit) approach significantly understates the true probability of distress.

The dominance of **foreign reserve adequacy** aligns perfectly with the Third-Generation Crisis Models and empirical findings for similarly structured emerging markets, emphasizing that acute **external liquidity risk** is the primary channel of vulnerability, even more so than baseline solvency. The identified **3.0-month threshold** refines previous empirical work by providing a precise, context-specific warning trigger.

Furthermore, the model's findings offer strong comparative insights for economies with similar structures in the MENA and African regions—characterized by high import dependence, a large public sector, and fiscal dominance. For such countries, the primary defense against sovereign distress lies not just in balancing the budget, but in aggressively securing the external liquidity buffer and maintaining a competitive exchange rate.

Policy Recommendations

The transparency and precision offered by the SHAP framework translate directly into specific, actionable policy recommendations for Egyptian authorities and international stakeholders:

© 2025 by Author/s

Strategic External Buffer Management

Policy efforts must prioritize maintaining the reserve buffer securely above the **3.0-month threshold**. It is recommended to establish a formal policy target zone (e.g., 4.0 months) to guide Central Bank actions, with explicit contingency plans triggered immediately upon falling below the 3.5-month level. This strategy is critical to reducing the probability of a self-fulfilling capital flow reversal.

Commitment to Exchange Rate Flexibility

The model's strong emphasis on the **REER** index reinforces the urgent need for a credible and sustained commitment to a flexible exchange rate regime. Resistances to necessary exchange rate adjustments, as seen in the periods leading up to the 2016 and 2022 crises, allow imbalances to compound, leading to more painful, abrupt, and economically damaging adjustments later. Flexibility acts as an automatic shock absorber, reducing pressure on the reserve buffer.

Enhanced Debt Sustainability Analysis (DSA)

Official Debt Sustainability Analyses should be immediately enhanced to incorporate the insights from **non-linear risk assessment**. This means moving beyond linear projections and explicitly modeling the identified interaction effects and the non-linear "tipping points"—particularly the simultaneous effect of high debt (above 85%) and low reserves (below 3.5 months). Policy decisions on new borrowings should be evaluated against the projected impact on the EWS's probability of distress.

Operationalization of an Early Warning Dashboard

The validated **XGBoost model** should be operationalized within the Central Bank of Egypt (CBE) and the Ministry of Finance (MoF) as a dynamic, continuous monitoring dashboard. This tool would provide a transparent, objective risk assessment, automatically flagging when key indicators approach or breach their critical non-linear thresholds, enabling a shift from reactive crisis management to proactive, pre-emptive policy responses.

Robustness Assessment and Limitations

Robustness

A comprehensive series of sensitivity analyses confirmed the stability of the core findings. These checks included: employing alternative crisis definitions (e.g., relying only on a high spread threshold), varying the predictor set, and using different model hyperparameters. In all specifications, the **XGBoost model** maintained its superior predictive performance, and the dominance of reserve adequacy, REER, and public debt remained consistent. The unwavering prominence of these three core variables across a battery of sensitivity tests underscores their fundamental role as the primary drivers of sovereign debt distress in the Egyptian economy.

Limitations and Mitigation Strategies

While the research design incorporates several strategies to mitigate the challenges of a single-country study, the following limitations should be considered when interpreting the results:

1. Small Crisis Sample: The limited number of historical crisis events constrains model complexity. This challenge was directly addressed by employing Leave-One-Out Cross-Validation (LOOCV) and a strict temporal validation scheme. These methods are explicitly designed to maximize the utility of scarce data and provide a rigorous, out-of-sample evaluation of predictive performance, thereby strengthening the validity of our findings despite the small sample.
2. Data Frequency: The use of annual data may not capture early warnings from higher-frequency financial signals. Future research could enhance the framework's lead time by incorporating quarterly data for key market-based indicators.
3. Methodological Constraint: The application of SMOTE in a time-series context presents a methodological caveat. As detailed in Section 3.2.3, this was mitigated by strictly confining its use to LOOCV training folds and validating on pristine historical data. The consistency of our core results across robustness checks confirms that the main findings are robust to this technique.
4. Context Specificity: The identified risk thresholds are intrinsically linked to Egypt's economic structure. Consequently, the primary contribution of this study is the proposed analytical framework itself—which is transferable—rather than the country-specific numerical outputs.

© 2025 by Author/s

CONCLUDING REMARKS AND FUTURE RESEARCH

This study develops a practical and interpretable early warning tool for Egyptian policymakers, enabling a more proactive approach to sovereign debt risk management. Our analysis demonstrates that machine learning methods, particularly the XGBoost model, significantly enhance the accuracy and lead time of early warning systems compared to traditional econometric benchmarks.

More importantly, the SHAP framework allowed us to move beyond opaque predictions to generate precise, economic insights. The identification of critical, non-linear risk thresholds, such as the 3.0-month benchmark for foreign reserve coverage, provides explicit operational intelligence. The framework's successful ex-post explanation of the 2023 distress episode confirms its practical efficacy as a risk management tool.

This research reinforces the paramount importance of developing tailored, country-specific early warning systems that can explicitly capture unique institutional structures and non-linear economic dynamics often obscured in aggregate cross-country studies.

Future research should focus on three areas:

1. Higher-Frequency Integration: Integrating quarterly or monthly data for market-based indicators (spreads, capital flows) to shorten the warning window.
2. Expanding Vulnerability Channels: Incorporating non-traditional variables such as the sovereign-bank nexus (e.g., bank holdings of government debt) and measures of climate change vulnerability (e.g., water stress, coastal risk) as long-term risk drivers.
3. Regional Comparative Studies: Extending this analytical framework to a cluster of structurally similar regional economies (e.g., Jordan, Morocco, Tunisia) for a comparative assessment of non-linear thresholds and risk dominance.

Funding: This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2501).

REFERENCES

Al-Shawarby, S. (2019). *Economic growth in Egypt: Determinants and constraints*. Egyptian Center for Economic Studies.

Catao, L. A., & Milesi-Ferretti, G. M. (2014). External liabilities and crises. *Journal of International Economics*, *94*(1), 18–32.

Central Bank of Egypt. (2023). *Annual report 2022*. Cairo: CBE.

Detragiache, E., & Spilimbergo, A. (2001). *Crises and liquidity: Evidence and interpretation* (IMF Working Paper No. 01/2). International Monetary Fund.

Flood, R. P., & Garber, P. M. (1984). Collapsing exchange-rate regimes: Some linear examples. *Journal of International Economics*, *17*(1–2), 1–13.

International Monetary Fund. (2016). *Egypt: Request for extended arrangement under the extended fund facility* (IMF Country Report No. 16/181).

International Monetary Fund. (2023). *Egypt: Staff report for the 2023 Article IV consultation* (IMF Country Report No. 23/105).

International Monetary Fund. (2023). *Global financial stability report: Navigating high inflation and war*. Washington, DC: IMF.

Kaminsky, G. L., Lizondo, S., & Reinhart, C. M. (1998). Leading indicators of currency crises. *IMF Staff Papers*, *45*(1), 1–48.

Khalid, A. (2020). *Exchange rate pressure and monetary policy in Egypt*. Routledge.

Krugman, P. (1979). A model of balance-of-payments crises. *Journal of Money, Credit and Banking*, *11*(3), 311–325.

Krugman, P. (1999). Balance sheets, the transfer problem, and financial crises. *International Tax and Public Finance*, *6*(4), 459–472.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4765–4774.

Manasse, P., & Roubini, N. (2009). "Rules of thumb" for sovereign debt crises. *Journal of International Economics*, *78*(2), 192–205.

Obstfeld, M. (1986). Rational and self-fulfilling balance-of-payments crises. *The American Economic Review*, *76*(1), 72–81.

Obstfeld, M. (1996). Models of currency crises with self-fulfilling features. *European Economic Review*, *40*(3–5), 1037–1047.

Reinhart, C. M., & Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly*. Princeton University Press.

© 2025 by Author/s

Selim, H. (2021). *Fiscal sustainability in Egypt: Challenges and prospects*. Palgrave Macmillan.
World Bank. (2022). *Egypt economic monitor: Navigating the storm*. Washington, DC: World Bank Group.

Appendix A

Appendix A.1. Detailed Model Metrics and Robustness Checks

The appendix contains supplementary technical details necessary for the reproduction and thorough understanding of the model's performance. This includes the full confusion matrix from the cross-validation exercise, alternative model specifications (e.g., Logistic Regression and Random Forest), and detailed charts for the individual feature importance derived from the SHAP framework.

Table A1. Key Model Hyperparameters.

Parameter	Description	Value
n_estimators	Number of boosting rounds	100
max_depth	Maximum depth of a tree	5
learning_rate	Step size shrinkage	0.1

 © 2025 by Author/s