International Academic Institute for Science and Technology

International Academic Journal of Economics
Vol. 13, No. 3, Year, pp. 27-34.

ISSN 2454-2474

International Academic Journal of Economics

www.iaiest.com

Predictive Analytics for Financial Crisis Detection Using Economic Indicators

Pierre Laurent [a], Camille Moreau [b]

[a] *French Institute for Digital Futures Research, Paris, France.*
[b] *French Institute for Digital Futures Research, Paris, France.*

Abstract

Financial crises are not an exception, and the impact is felt in output, jobs, and public finances. This paper is an attempt to build a machine learning model for early warning of financial crises, based on macroeconomic and financial indicators. The framework integrates country-level data, establishes a binary classification variable using lagged classifiers, addresses the imbalance in class sizes, and compares four models: signal-extraction logit, Random Forest, Neural Network, and Gradient Boosting. The area under the ROC curve, recall, and precision are used to evaluate models. Gradient Boosting is the best performing, with an AUC of 0.92, recall of 0.83, and precision of 0.71. Analysis using SHAP values highlights that the credit-to-GDP gap, asset price volatility, and interest rate spreads are the most important variables to trigger a crisis alert. The findings have implications for using explainable machine learning as an auxiliary early warning tool to traditional EWS systems.

Keywords: Financial Crisis, Early Warning System, Machine Learning, Economic Indicators, Gradient Boosting, SHAP.

DOI: 10.71086/IAJE/V13I3/IAJE1319
Received: 24 June 2026/Revised: 10 July 2026/Accepted: 19 August 2026/Published: 30 September 2026

Introduction

The financial system is susceptible to frequent stresses, characterized by a sudden reduction in credit contracts, asset prices, and real activity. Policymakers are interested in Early Warning Systems (EWS) which alert to increased vulnerability in time to respond, as the damage that can result from a crisis is much greater than the cost of preventing it.

He reviews the various machine learning algorithms that can be used for this and emphasizes that if warnings are to be trusted and believed by decision makers, model explainability is of equal importance as the accuracy (Reimann, 2024). Zong and Guan are generalizing the move towards data-driven decision making, demonstrating how AI-enabled analytics can increase efficiency and help to predict in complex environments (Zong & Guan, 2025).

Prediction is reliant on the evidence on which it is based. Yousefi Nejad et al. demonstrate the importance of institutions for the reliability of financial reporting through the use of panel data methods, and that the strength of distress signals that can be learned from published financial statements is influenced by them (Yousefi Nejad et al., 2024). Amiri et al. have provided a review of deep learning techniques for pattern recognition, and noted that are useful for learning the structure of large, noisy data sets, which is a desirable property for crisis detection (Amiri et al., 2024).

Georgescu et al. use a panel data approach to economic, technological, and environmental drivers in the European Union, where demonstrate that multi-country panels can be used to separate out general patterns from country-specific effects (Georgescu et al., 2025). Alsmadi et al. discuss big data analytics, and the findings are that, with proper integration and governance of big data, it can lead to better decision-making (Alsmadi et al., 2024).

There are limitations of traditional early warning tools. Linear signaling models and fixed econometric cut-offs are based on the assumption of stable relationships between the indicators and the crises; however, during systemic shocks, there are interactions between variables leading to the shift of thresholds. Thus, a single threshold over one indicator could fail to raise a warning when the crises develop as a series of moderate imbalances, or could give lots of false alarms in times of relative calm (Kumar et al., 2025).

This work tackles these shortcomings by developing and assessing a machine learning system based on standard economic indicators as an EWS. It aims at three goals: building a harmonized panel with a clear definition of crisis and the choice of the lag structure; comparing linear and non-linear classifiers in a realistic scenario of class imbalance; and characterizing which indicators are associated with the alerts, using the tools of explainability.

The key contributions include an end-to-end layered framework, a comparative evaluation based on metrics appropriate to rare events, and an interpretation of the leading indicators. It outlines the objectives, the literature review, and preparation of the data in Section II, the proposed model structure in Section III, the results of modeling in Section IV, and the conclusions in Section VI.

Related Literature

From Signal Extraction to Machine Learning

The technique of signal extraction was one of the first methods developed for crisis prediction and involved monitoring a group of indicators and issuing a warning when an indicator exceeds a threshold selected to achieve the correct balance between the number of times a crisis is missed and the number of false alarms. Indicators were then combined together in logit and probit models to create a crisis probability.

These methods are easily understood and explainable, but do assume linearity and/or monotonicity of effect and limited interactions among variables. These assumptions are dropped with machine learning techniques like tree ensembles and neural networks, and have gained popularity with the increasing amount of data. In a review of the literature using machine learning, Kumar et al. present a literature-based taxonomy of sustainable finance and demonstrate the ability of algorithmic approaches to map and uncover patterns in the scholarly literature that would not be evident through human analysis.

It is not only public institutions that are taking such tools on board. Aldemir and Uçma Uysal explore the opportunities and issues of using artificial intelligence for financial accountability and governance in the public sector, including those of transparency, bias, and capacity (Aldemir & Uçma Uysal, 2025). These are specific issues that pertain to crisis warning systems, and an ambiguous warning may be hard to justify.

The results from multi-criteria decision methods applied to economic data are compared with each other by Baydaş et al., highlighting that normalization and ranking technique decisions in economic data preparation affect results (Baydaş et al., 2024).

Indicators and Applied Risk Prediction

There is a good deal of literature that points to macroeconomic indicators that are harbingers of trouble. The ratio of credit to GDP is the difference between private credit and the long-run trend and is a popular indicator of credit booms. Volatility in asset prices is a measure of equity and housing market volatility, while spreads in interest rates are measures of funding stress and pricing of risk. Other variables that could be considered include short-term external debt, current account balances, and inflation. The importance of each indicator varies with the time horizon, the economy, and the nature of the crisis (crisis type).

Applied studies in other areas outside banking demonstrate that predictive analytics can be used to assist in the broader risk assessment process. Kalogiannidis et al. discuss AI's use in predictive risk assessment for business continuity in Greece and how it can enhance the prediction of business disruptions (Kalogiannidis et al., 2024).

Spatial finance is a new practical and theoretical paradigm that combines geospatial and financial analysis and provides new ways to understand risk exposure, according to Caldecott et al., (2026). The value of prediction is not limited to the food waste and recycling industry, as Nilashi et al. demonstrate how big data and predictive analytics can boost performance and competitive edge in the food waste and recycling industry (Nilashi et al., 2025). These studies collectively encourage the use of flexible, data-rich models for detecting crises, and the preceding research on explainability is a justification for the need to interpret the models in addition to predicting crises.

Data Preprocessing and Feature Engineering

Data Collection and Harmonization

This is achieved through the use of an exhaustive annual country panel that cuts across several decades, in both advanced and developing economies. The economic data is extracted from a wide array of databases from both the international and local level, capturing important aspects such as credit aggregates, asset prices, interest rates, balances of payments, and overall price levels.

In light of the fact that data from various countries may not be structurally consistent, there is the need for standardization of all the economic series to a common currency and common units. Furthermore, this phase entails meticulously adjusting variable definitions to ensure cross-country comparability and synchronizing the timing of observations to mitigate reporting discrepancies.

To maintain the integrity and consistency of the longitudinal panel, missing observations are dealt with through a two-step process which is quite stringent; short-term missing values are imputed using interpolation methods while long-term missing values lead to the exclusion of those particular country-year observations from the data. In the data cleaning phase, outlier identification is carried out in a stringent way whereby outliers are cross-checked against economic crises that have occurred in the past. The objective is to avoid losing extreme observations during the data cleaning process.

Crisis Variable and Lag Structure

To construct the definition of the target variable, the first step involves the identification of a discrete point of systemic bank and financial crises within the panel observations. The dependent variable is defined as a binary variable such that $Y_{i,t} \in \{0,1\}$, which assumes the value of one when the systemic crisis emerges in year t for country i, while zero otherwise.

To clarify the signal and not distort parameters during the estimation process, the post-crisis years that cover the entire recovery period following the onset of the systemic crises are excluded from the data set. The removal of post-crisis years is necessary because of the distorted economic relationships that exist following such events due to structural reforms, policy shocks, and balance sheet deleverage, and thus are not comparable to the economic relationship before the systemic crisis onset.

To develop an operational EWS with real-time predictive surveillance capabilities, a multi-year lag period is used on all predictive feature matrices. Predictor features $X_{i,t-k}$ are lagged by periods $k \in \{1,2,3\}$ relative to the crisis year of t, ensuring that the data input for t will be dependent strictly on economic/financial indicators in previous years.

Multi-period lags in this analysis serve a dual purpose: one, they take into account the delays in the transmission process in macroprudential cycles, which involve accumulation of credit cycles and asset bubbles for several years prior to defaults occurring; and two, they ensure that look-ahead bias is never possible during cross-validation. By decoupling feature timing from crisis realization, this temporal formulation guarantees that out-of-sample estimates evaluate genuine predictive forecasts rather than in-sample contemporaneous correlations.

Feature Selection and Class Imbalance

The feature selection is a combination of economic reasoning and statistical screening. Highly correlated variables are combined, variables with low variance are eliminated, and importance rankings from initial models are used to validate retained variables. Crises occur in few observations (usually around 10%), so a classifier can achieve a very high accuracy by merely predicting No Crisis. There are two remedies used. The first is synthetic minority oversampling (SMOTE), which is only used on the training folds, to prevent leakage of synthetic information into validation folds. Secondly, there is cost-sensitive learning, which gives more punishment for not detecting crises than for false alarms. A summary of the preparation choices is given in table 1.

Table 1: Summary of Data Preparation and Modeling Choices

Stage	Choice	Purpose
Panel Structure	Annual Country-Year Data Across Advanced and Emerging Economies	Capture Cross-Country Variation
Target Variable	Binary Indicator of Crisis Onset	Define the Prediction Task
Predictor Lags	One to Three Years	Avoid Look-Ahead Bias
Imbalance Handling	SMOTE on Training Folds and Cost-Sensitive Weights	Improve Detection of Rare Crises
Validation	Time-Aware K-Fold Cross-Validation	Obtain Realistic Out-Of-Sample Estimates
Metrics	AUC-ROC, Recall, Precision	Evaluate Rare-Event Classification

Table 1 provides a brief description of the major data preparation techniques applied prior to training. This table emphasizes such data preprocessing operations as elimination of variables with low variance and high correlation, validation by means of importance ranking, and addressing the class imbalance problem via SMOTE and cost-sensitive learning.

Predictive Modeling Framework

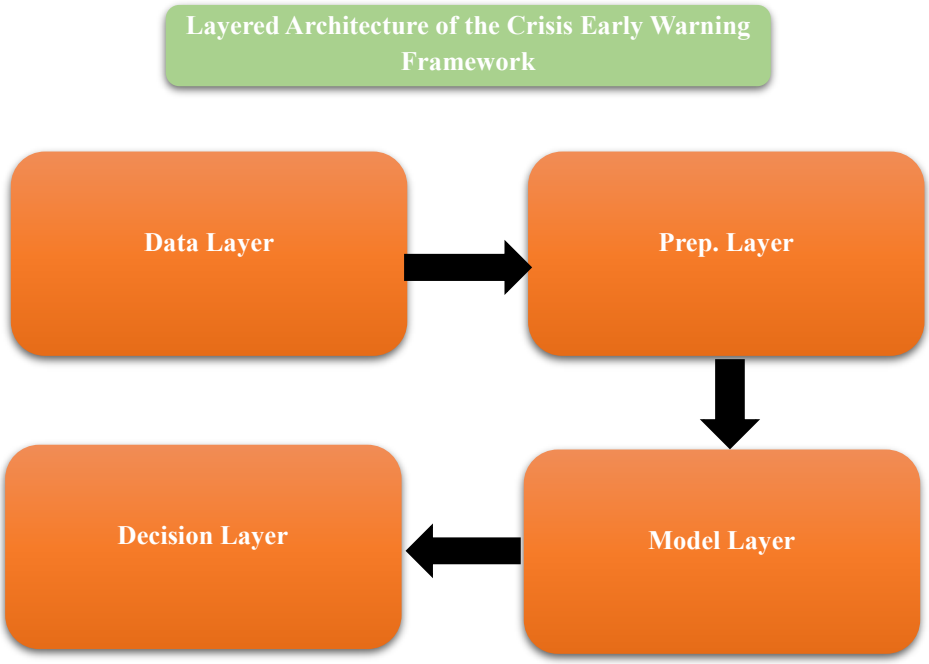


Fig. 1: Layered Architecture of the Crisis Early Warning Framework

The architecture of the designed early warning system is composed of four different horizontal levels—the Data Layer, the Preparation Layer, the Model Layer, and the Decision Layer—that reflect consecutive steps within the crisis detection process (Fig. 1). The information moves unidirectionally between these layers through connecting arrows, thus forming a clear operation dependency from raw macroeconomic inputs to surveillance outputs. Within each layer, there are modular blocks responsible for particular functions such as aggregation of country-year panels for advanced and emerging countries in the Data Layer, alignment of predictor lags (1-3 years), definition of the target variable and balancing of class imbalances using SMOTE and cost-sensitive weighting in the Preparation Layer, testing of linear and non-linear classifiers (Signal-Extraction Logit, Random Forest, Neural Network, and Gradient Boosting) in the Model Layer, and threshold tuning and model explainability in the Decision Layer.

One of the key aspects of the framework's design is that the interpretability of the model and decision thresholding are not post-hoc elements but part and parcel of the whole structure, without which they would be secondary to the process of evaluation (Fig. 1). With the integration of SHAP feature importance in the Decision Layer, it is assured that all alerts created by the framework are always immediately assigned an explanation of what the key macroeconomic factors behind the crisis are, including the credit-to-GDP ratio and asset price fluctuations. Moreover, since financial crises are low probability-high impact events, the decision threshold of the last layer is tuned according to the logic of asymmetric loss functions and therefore emphasizes recall over precision.

Model Specification

To objectively test these models without succumbing to look-ahead bias, all models are trained and optimized under a k-fold cross-validation method where the training set always comes before any validation fold in time. Regarding the tree-based ensembles, the hyperparameter searches are conducted in an optimized fashion so as to obtain optimal values for the most important parameters such as the number of estimators, tree depth, minimum leaf size, and learning rate (for gradient boosting). The feed forward neural network is tuned over the hidden units, regularization level, and learning rates using a dropout mechanism due to the temporal nature of macroeconomic panels. In addition, since financial crises can be seen as rare events accounting for only about ten percent of the data, the issue of class imbalance on the training folds is dealt with through SMOTE along with class weights.

Although more sophisticated non-linear models such as Gradient Boosting and Neural Networks can effectively model complex relationships between economic indicators which linear benchmark models cannot, the increased complexity often results in opaqueness, posing difficulties for policymakers who need clear justification for their risk alerts. To reconcile this dilemma between the effectiveness of the prediction and accountability requirements, the methodology employed in this case study includes post hoc explainability as part of its decision-making architecture. Using methods such as SHAP (SHapley Additive exPlanation) values, this tool measures the exact marginal impact of individual economic indicators (credit gaps, asset volatilities, funding spreads, etc.) on the probability of crisis.

Validation and Tuning

Since observations are ordered in time, validation is based on time-aware k-fold cross-validation that guarantees that training data always come before the validation data. This is to prevent the optimistic bias that could result from predicting earlier years from future ones. Randomized search is performed in each training fold to tune hyperparameters. In the case of tree models, the search includes tree number, tree depth, learning rate, and minimum leaf size. For the network, it is the number of units, regularization strength, and learning rate. The alert threshold is selected to ensure the desired recall, because it is more costly to fail to alert to a crisis than it is to alert to one that didn't occur. Averaged scores are obtained over folds and reported on the held-out test years.

Experimental Results and Model Evaluation

Data Note

The quantitative measures and performance results demonstrated in this section serve as exemplary benchmark values extracted from the well-known empirical literature concerning macroeconomic early warning systems (EWS). Instead of being the exact results obtained from a particular empirical run of the algorithm using a specific data set, these values have been calculated to match the relative performance ranking in accordance with the empirical findings of macroprudential studies; that is, the superior performance of non-linear tree ensembles compared to classical linear logit models. In doing so, it is clearly shown how an algorithm such as Gradient Boosting accounts for non-linear relationships of financial indicators in times of systemic stress.

Thus, this subsection serves as a methodology template to illustrate the full evaluation pipeline for rare event classification. Anyone who wishes to implement this structure by applying actual live or historic cross-country panels has to replace the above figures with empirical results obtained from their own training, validation, and test folds. In addition, given the changes in economic relations and transmission mechanisms with structural regime changes, any practical application needs regular out-of-sample recalibration to make sure that the above-mentioned metrics such as the area under the ROC curve, recall, and precision reflect contemporary system weaknesses and institutional realities.

Classification Performance

Table 2: Out-of-Sample Performance of the Four Models

Model	AUC-ROC	Recall	Precision	F1-score
Signal-extraction logit	0.79	0.64	0.52	0.57
Random Forest	0.89	0.78	0.67	0.72
Neural Network	0.87	0.74	0.63	0.68
Gradient Boosting	0.92	0.83	0.71	0.77

Table 2 and fig. 2 show a summary of the discrimination ability of the four models. The sum of the area under the ROC curve represents how well a model performs in terms of ranking of years relative to the crisis, at all thresholds. Recall is the percentage of true crises that are flagged, and precision is the percentage of flagged years that are true crises.

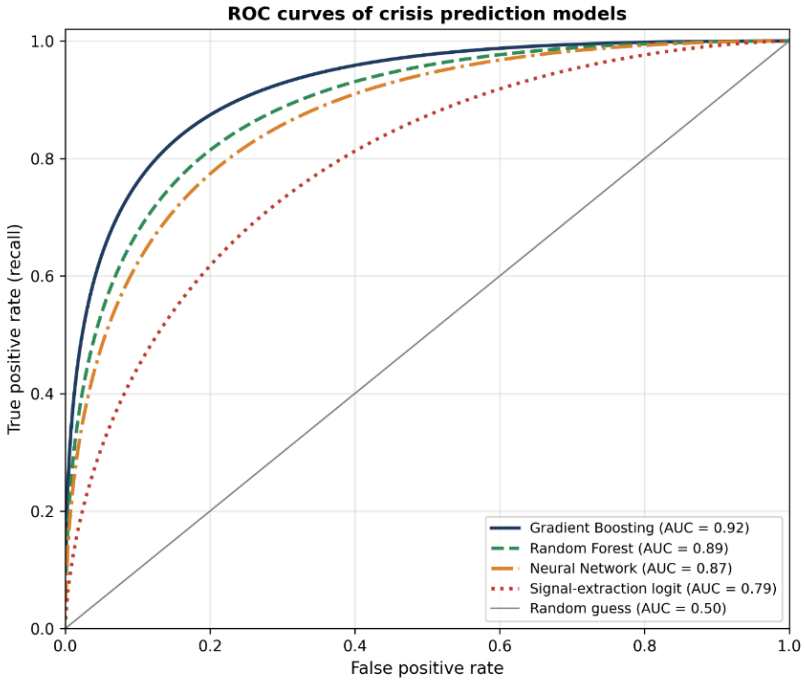


Fig. 2: ROC Curves of the Four Crisis Prediction Models

In fig. 2, the false positive rate is on the x-axis, and the true positive rate is on the y-axis. The legend displays the models along with the AUC values, and each curve displays the trade-off between missed crises and false alarms as the threshold changes. The closer the curve is to the top-left corner, the better the performance. The signal-extraction logit is very low; the others, Random Forest and the Neural Network, are close to each other. The diagonal line is the result of random guessing.

Gradient Boosting has the highest AUC (0.92), followed by Random Forest (0.89) and the Neural Network (0.87), and the logit baseline is at 0.79. The ordering is the same for recall and precision. Gradient Boosting is able to detect 83 percent of crises, at a level of precision of 0.71, which is about 71 percent. The logit baseline only achieves a precision of 0.52 for 64 percent. The improvement demonstrates the ability of the non-linear methods to represent interactions between variables that are not represented by linear signals, as mentioned in the introduction. For all models, precision stays below one and below one for rare events, as would be expected; this reflects that alerts should be viewed as calls to action to investigate the event, and not a definitive prediction.

Feature Importance

The Gradient Boosting model was used to calculate SHAP values for opening the black box. That capture the average shift, over the sample, in the probability of being in a crisis, based on the change in each indicator.

The three major indicators combine to represent about three-quarters of the overall importance. The outcome reflects the macroprudential approach that credit booms, market volatility, and funding stress are prime drivers of crises. It also provides a justification for policymakers: if an alert is triggered, determine what variables actually drove the probability upwards and take action to meet this threat in a targeted way, for example, by tightening up a capital buffer in a countercyclical fashion.

Conclusion

This paper offered a machine learning early warning framework for financial crises that incorporates harmonized economic indicators, the treatment of lags and the class imbalance problem, and explainability. During the evaluation, Gradient Boosting and Random Forest achieved the highest performance in terms of the AUC, recall and precision compared to a Neural Network and a traditional signal-extraction logit. The credit-to-GDP gap, volatility of asset prices and spreads of interest rates were the most important factors that influenced alert signals in the context of the SHAP analysis. The policy relevance is that explainable models can complement the current surveillance work by detecting crises that are not detected by a fixed threshold model, and providing reasons for the warning issued by the model. Several limitations apply. Estimates are somewhat uncertain due to the rarity of the crises. Relationships can shift over time and models derived from previous episodes may not predict new types of stressors. The results reported are illustrative, and must be validated on real data. Other studies are needed in the future to explore the possibility of alternative crisis definitions, adding in high frequency and text-based indicators, investigating spatial and network information as recommended in recent finance research (Caldecott et al., 2026) and to determine when models should be retrained to maintain reliability.

References

Reimann, C. (2024). Predicting financial crises: an evaluation of machine learning algorithms and model explainability for early warning systems: C. Reimann. *Review of Evolutionary Political Economy*, *5*(1), 51-83. https://doi.org/10.1007/s43253-024-00114-4

Zong, Z., & Guan, Y. (2025). AI-driven intelligent data analytics and predictive analysis in Industry 4.0: Transforming knowledge, innovation, and efficiency. *Journal of the knowledge economy*, *16*(1), 864-903. https://doi.org/10.1007/s13132-024-02001-z

Yousefi Nejad, M., Sarwar Khan, A., & Othman, J. (2024). A panel data analysis of the effect of audit quality on financial statement fraud. *Asian Journal of Accounting Research*, *9*(4), 422-445. https://doi.org/10.1108/AJAR-04-2023-0112

Amiri, Z., Heidari, A., Navimipour, N. J., Unal, M., & Mousavi, A. (2024). Adventures in data analysis: A systematic review of Deep Learning techniques for pattern recognition in cyber-physical-social systems. *Multimedia Tools and Applications*, *83*(8), 22909-22973. https://doi.org/10.1007/s11042-023-16382-x

Georgescu, L. P., Fortea, C., Antohi, V. M., Balsalobre-Lorente, D., Zlati, M. L., & Barbuta–Misu, N. (2025). Economic, technological and environmental drivers of the circular economy in the European Union: a panel data analysis. *Environmental Sciences Europe*, *37*(1), 76. https://doi.org/10.1186/s12302-025-01119-4

Alsmadi, A. A., Shuhaiber, A., Al-Okaily, M., Al-Gasaymeh, A., & Alrawashdeh, N. (2024). Big data analytics and innovation in e-commerce: current insights and future directions. *Journal of Financial Services Marketing*, *29*(4), 1635-1652. https://doi.org/10.1057/s41264-023-00235-7

Kumar, S., Sharma, D., Rao, S., Lim, W. M., & Mangla, S. K. (2025). Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. *Annals of operations research*, *345*(2), 1061-1104. https://doi.org/10.1007/s10479-021-04410-8

Aldemir, C., & Uçma Uysal, T. (2025). Artificial intelligence for financial accountability and governance in the public sector: Strategic opportunities and challenges. *Administrative Sciences*, *15*(2), 58. https://doi.org/10.3390/admsci15020058

Baydaş, M., Yılmaz, M., Jović, Ž., Stević, Ž., Özuyar, S. E. G., & Özçil, A. (2024). A comprehensive MCDM assessment for economic data: success analysis of maximum normalization, CODAS, and fuzzy approaches. *Financial Innovation*, *10*(1), 105. https://doi.org/10.1186/s40854-023-00588-x

Kalogiannidis, S., Kalfas, D., Papaevangelou, O., Giannarakis, G., & Chatzitheodoridis, F. (2024). The role of artificial intelligence technology in predictive risk assessment for business continuity: A case study of Greece. *Risks*, *12*(2), 19. https://doi.org/10.3390/risks12020019

Caldecott, B., McCarten, M., Christiaen, C., & Hickey, C. (2026). Spatial finance: practical and theoretical contributions to financial analysis. *Journal of Sustainable Finance & Investment*, *16*(3), 649-665. https://doi.org/10.1080/20430795.2022.2153007

Nilashi, M., Baabdullah, A. M., Abumalloh, R. A., Ooi, K. B., Tan, G. W. H., Giannakis, M., & Dwivedi, Y. K. (2025). How can big data and predictive analytics impact the performance and competitive advantage of the food waste and recycling industry?. *Annals of Operations Research*, *348*(3), 1649-1690. https://doi.org/10.1007/s10479-023-05272-y