Artificial Intelligence Integration in Assessment of Financial Stability of Egyptian Banks

By

Dr. Wael Mostafa Hassan Mohamed

Associate professor of Finance and Investment

Faculty of Financial and Administrative Sciences, Pharos University in Alexandria

dr.wael.moustafa@gmail.com

Scientific Journal for Financial and Commercial Studies and Research (SJFCSR)

Faculty of Commerce – Damietta University

Vol.6, No.2, Part 1., July 2025

APA Citation

Mohamed, W. M. H. **(2025)**. Artificial Intelligence Integration in Assessment of Financial Stability of Egyptian Banks, ***Scientific Journal for Financial and Commercial Studies and Research***, Faculty of Commerce, Damietta University, 6(2)1, 431-470.

Website: https://cfdj.journals.ekb.eg/

Artificial Intelligence Integration in Assessment of Financial Stability of Egyptian Banks

Dr. Wael Mostafa Hassan Mohamed

Abstract:

The purpose of this study is to evaluate the stability of Egyptian banks over the period from 2010 to 2024, using Z-score as the dependent variable for financial stability, while the key independent variables are Liquidity Ratio (LR), Return on Assets (ROA), Return on Equity (ROE), Non-Performing Loans (NPL), and Capital Adequacy Ratio (CAR). Neural Networks and XGBoost were used as two AI models to measure the importance of those indicators/predictors regarding the stability of Egyptian banks.

Results show that whilst both models appropriately note the importance of profitability, risk management, capital adequacy, and liquidity to bank performance, they do differ in terms of the ranking of these factors. In Neural Network model, the most important feature is Return on Equity_ ROE (0.4521), then comes the Capital Adequacy Ratio_ CAR (0.3435) and in case of XGBoost, The Capital Adequacy Ratio_CAR has the highest importance (0.3500), then comes ROE (0.1900) and LR (0.2700).The two models confirm the lower significance of NPL compared to others, where XGBoost gave the lowest weight (0.0700) and Neural Networks (0.0982). The different feature rank of independent variables was a result of individual models' sensitivity to dataset and the processing through its unique logarithms which indicates that AI models can help in the assessments of banks' stability via examining the predictor power of internal factors which represent Profitability, Risk Management, Liquidity Management and Banks' Capital Adequacy in an efficient manner.

Keywords: Banks' Financial Stability, Z-Score, AI, Neural Network, XGBoost

1.Introduction

Artificial Intelligence (AI) has evolved since its inception in the 1950's, when initial rule-based systems and symbolic reasoning got developed. In the 1980s and 1990s statistical learning methods became widely used for financial market prediction and financial modeling. But it was in the early 2000s driven by improvements in computing and cheap availability of huge amounts of financial data that we entered an era of AI adoption in finance. Today, contemporary AI methodologies like deep learning, ensembles, and NLP allow for moment-by-moment data-oriented decision-making within many domains of finance. These use-cases span credit scoring, fraud detection, algorithmic trading, portfolio optimization and regulatory compliance. (Guler et al.,2024)

In recent years, unstructured data such as social media sentiment and financial news are being analyzed with AI leading to deeper predictive analytics. With the continued convergence of financial systems complexity and data-driven decision-making, AI is not a complementary tool. However, it is a foundational tool for driving innovation, inclusion and resiliency across global financial systems. (Bahoo et al.,2024).

Egypt's National Artificial Intelligence Strategy 2025–2030 was launched by President Abdel Fattah El-Sisi in January 2025. The strategy seeks to establish Egypt among the top countries in the Middle East and North Africa (MENA) region in the field of artificial intelligence (AI), through bolstering the contribution of the ICT sector to GDP, boosting AI startups, talent, and innovation. Governance, infrastructure, tech, data, ecosystem & talent are the key focus areas. (Arab News,2025).

Additionally, The Evaluation of banks' financial stability is a key element in maintaining a strong and resilient financial system in Egypt. As of September 2024, Egypt's banking sector demonstrated strong financial health. As Capital adequacy ratio (CAR) achieved 19.1%, which exceeds the minimum of 12.5% established by CBE. Liquidity ratios were equally strong, with local currency liquidity and foreign currency liquidity at 32.1% and 77.7% respectively, both above regulatory limits. The ratio of non-performing loans (NPLs) to total loans decreased to 2.4% and provisions covered 87.4% of these NPLs.

The Central Bank of Egypt (CBE) is currently addressing the issuance of new laws and regulations aimed at strengthening financial technology, including behavioral-based digital lending, alternative financing, cloud computing services, and electronic authentication. These measures are intended to drive innovation and provide enhanced accessibility in the financial sector.

2. Research Problem

The financial stability of banks is a cornerstone of the functioning of any economy, especially in an emerging one, such as Egypt. Egypt successfully managed challenges facing its banking sector as high economic fluctuations, inflationary pressures, high interest rate and exchange rate volatility. These challenges emphasize the importance of the implementation of tools to prevent and mitigate threats to financial stability.

Additionally, Egypt is going through a significant transformation where the central bank of Egypt (CBE) is focused on modern laws to solidify financial technology and developing strong regulations, such as the establishment of artificial intelligence solutions to support the development of decision-making processes. In addition, the Egyptian government has launched the National Artificial Intelligence Strategy 2025–2030, which is the national strategy that aims to make Egypt one of the largest countries in this field globally, especially in the financial field, by increasing innovation, supporting startups, and working to raise the capacity of banks and their ability to use AI technologies.

While AI has significant potential to improve banking activities, risk management, and financial stability, the difficulty lies in how to incorporate AI models into financial stability assessment of Egyptian banks. As the integration into the evaluation of financial stability especially in emerging markets such as Egypt remains limited.

Thus, this research seeks to address the core problem of how to utilize AI models as early warning system to identify the main predictors of the financial stability of Egyptian banks. Accordingly, this study aims to address the following research main question and sub questions:

Main Research Question: How can AI models be effectively used for evaluating the stability of Egyptian banks?

Sub Research Questions:

- **To what extent can AI models predict bank stability using Return on Equity (ROE) as a key indicator?**

- **How effectively can AI models utilize Capital Adequacy Ratio as a predictor for Egyptian banks' financial instability?**
- **What is the role of Liquidity Ratio in assessing the banking stability through the AI Predictive models?**
- **Can NPLs (non-performing loans) be used effectively in AI models as an early-warning signal for a possible instability of Egyptian banks?**

3. Literature Review and Hypotheses Development

3.1 Theoretical framework

Banking sector financial stability is not just about preventing bank failures, but a multi-dimensional concept. It involves retaining trust and confidence in the banking framework, its capability to work out its mandate of providing services that are necessary for functioning, supporting economic resilience. Such stability is critical for fostering economic growth, providing for effective allocation of resources, and ensuring protection for depositors and investors. Stable banking system is a shock absorber that can reduce disruptions and provides the basis for sustainable economic development. (Koskei,2020).

3.2 Theories of Financial Stability

In the pursuit of understanding the complex functioning of the financial system, theorists and policymakers had long tried to describe the functioning of its components, identify the sources of its crises, and formulate ways to protect its critical functioning. To this end, there have been multiple theoretical developments to study financial health and fragility in various dimensions. These theories are not only theoretical abstractions; they are the conceptual foundations from which serious analysis, effective modeling, and pro-active mitigation of systemic risks spring.

Theoretical thought about financial stability has been quite heterogenous, with a few prominent themes of thought emerging, which have shaped the course of inquiry and influenced the development of policy measures:

3.2.1 Financial Instability Hypothesis (FIH) Theory.

Developed by Hyman Minsky, the FIH argues that financial instability is an inherent characteristic of capitalist economies, particularly during periods of prolonged prosperity (Minsky, 1986). Minsky posits that during stable periods, economic agents become more complacent and increase their risk-taking, leading to a shift from hedge financing (where cash flows are sufficient to cover debt) to speculative financing (where cash flows may only cover interest payments) and finally to Ponzi financing (where cash flows are insufficient to cover either principal or interest). This increasing reliance on debt and speculative behavior makes the financial system more fragile and susceptible to shocks. Recent research continues to validate the core tenets of Minsky's hypothesis in explaining financial crises, as the financial system moves into more debt and speculation, it becomes inherently more fragile, making it more vulnerable to shocks. New research further confirms key aspects of Minsky's hypothesis in an explanation of financial crises (Nikiforos & Zezza, 2017).

The FIH has big implications for policy. Unlike frameworks that rely on exogenous shocks, the FIH argues that regulatory tools should instead be used to counteract the endogenous accumulation of financial fragility, and to do so in a proactive, countercyclical manner. These steps consist of the use of macroprudential policies for reducing systemic risk by import aggregate vulnerabilities including excessive credit growth and asset price bubbles (Borio, 2003). Third, the FIH reiterates the need for strong and adaptive regulatory oversight that can prevent the growth of non-transparent, highly-leveraged financing architectures, driven by financial innovation.

To sum up, Minsky's Financial Instability Hypothesis provides a simple yet durable logic for recognizing the inherent proclivity for instability in capitalist financial systems. The FIH offers a significant yardstick of a powerful driver of financial crises, through its endogenous perspective of financial fragility and their comparatively cyclical progression in financing regimes. Its enduring significance for understanding current macro financial behaviors emphasize the necessity for pre-emptive regulatory and policy measures that acknowledge the procyclical nature of markets and strive to attenuate the emergence of systemic risk across a time of sanguine equilibrium. This dismissal of Minsky is a dangerous game since history clearly shows that financial capitalism since the late 19th century has produced successive cycles of boom and bust that all follow the same pattern.

3.2.2 Systemic Risk Theory:

Systemic risk has become a core of contemporary research on financial stability, especially since the 2008 global financial crisis. The theories of systemic risk emphasize the complex interlinkages between financial institutions and markets, the interconnectedness of which can allow shocks, idiosyncratic or systemic, to spread throughout the financial system and cause systemic collapse (Borio, 2003). Such theory not only supplies an analytical framework that illustrates the inherent vulnerabilities in the design of the world financial system, they also provide an intellectual highlighting for prescriptive aimed at rectifying this instability.

Systemic risk is a result of interconnectedness that is necessary for modern financial systems to work. A web of transactions, exposures, and dependencies is woven between financial institutions, markets, and infrastructures. When functioning in non-stressed environments, this interconnectedness promotes both efficiency and liquidity, but when these same markets are stressed, they can transmit contagion. These linkages imply that shocks originating in one sector of the economy and the failure of a major institution, or, a crash in the price of an asset, can transmit through those linkages, amplifying first-order disturbances and causing systemic to systemic crises (Borio, 2003).

3.2.3 Liquidity Risk Theories:

Liquidity has been one of the building bricks of financial stability, acting as both a lubricant for smooth market functioning, and a last line of defense against systemic shocks. However, liquidity vulnerabilities—on both the funding side and market side—have the potential to amplify the impact of financial shocks and trigger a cascade of fire sales, contagion effects, and global financial distress (Brunnermeier & Pedersen, 2009). These tangled relations highlight the need for more subtle ideas about the interaction between funding and market liquidity, a tendency that some people have called the "liquidity spiral". In this theoretical exploration, we investigate the nature of liquidity risks, how they arise, their impact on financial stability, and their implications for regulation and supervision.

Funding liquidity risk is the risk that financial intermediaries experience obstacles in fulfilling their near-term payment obligations because of insufficient access to liquid assets or a reduction in available funding. Modern financial systems make extensive use of short-term funding instruments — repurchase agreements, commercial paper, interbank loans — to fund their operations and asset portfolios. Although these mechanisms increase transactional efficiency in times of tranquility, they can prove to be sensitive fragile points in times of turbulence in the market. An unexpected restriction or complete termination of short-term financing may force institutions to sell assets quickly, which can magnify liquidity strains and set off a self-fulfilling and self-reinforcing spiral of instability.

The systemic weaknesses are aggravated by the subtle interplay between funding and market liquidity risks. For example, an institution that faces funding liquidity stress may have no other option than to sell assets to settle short-term obligations, which can further weaken liquidity in the market. Declining asset valuations, on the other hand, can weaken the quality of collateral held by institutions, limiting their access to secured funding and intensifying liquidity pressures. The interaction between funding and market liquidity risks — where the former sets off the latter, thus deflating the former once again — is often referred to as the "liquidity spiral" (Brunnermeier & Pedersen, 2009).

3.3 Crucial Elements of Banks' Financial Stability.

There are some key indicators to support the financial stability of the banks: capital adequacy, asset quality, management soundness, earnings, and liquidity (Tung et al. ,2024). Capital adequacy is the measure of a bank's capital in relation to its risk-weighted assets, serving as a cushion against potential losses. Asset quality refers to the credit risk of a bank's loan portfolio and investments. Management soundness relates to the competence and integrity of a bank's leadership and its ability to manage and control risks. Earnings show the profitability of a bank and capacity of generating sustainable returns. Liquidity is a bank's ability to fulfill its short-term obligations as they become due. Proper risk management and regulatory compliance are key factors in financial stability, particularly for banks and these factors assist banks in identifying, measuring, monitoring, and controlling risks (Shah et al.,2024). All those ingredients together play into a bank's resilience and capacity to survive a downturn.

3.3.1 Bank-Specific Risks and their Impact on Financial Stability

There are inherent bank-specific risks in the banking business, known to exert a significant negative externality on financial stability. The risks associated with holding financial instruments on a bank's balance sheet encompass a range from credit risk through liquidity risk leading to operational risk, all of which require measured strategies for management and mitigation. The advent of digital banking presents both opportunities and threats that should be embraced with an appreciation of their potential impact.

3.3.2 Credit Risk and Financial Stability

Credit risk is a major risk to the stability of a bank's finances. It includes the risk of losses resulting when borrowers default on the loans granted to them (Thompson, 2024). Thus, the supervision and regulation of credit risk is crucial to not only the safety and soundness of individual banks, but to the stability of the financial system at large.

The multidimensional nature of credit risk, and its implications for banks performance have been scrutinized in various studies. Strong loan origination best practices and regular monitoring of repayment schedules are two important mechanisms for mitigating credit risk and maintaining the quality of banks' loan portfolios (Clark & Miller, 2023). Not only do these acts protect the profit margins of the individual bank, but they also promote systemic solidity by minimizing the chances of pervasive failure.

Disruption of macro-financial stability by non-performing loans (NPLs) has been the focus of a considerable body of literature. Non-performing loans (NPLs), classified loans on which borrowers have defaulted or are near default are detrimental to banks profits where they will either receive less in interest revenue or need to provision for loan losses (Garcia, 2022). High NPLs also reduce the capital adequacy of a bank, limiting its ability to absorb other losses and making the bank more vulnerable to negative shocks. Therefore, if plenty of banks face NPL for economic impact and cannot keep up, this will lead to a credit crunch and restrict the availability of funds to reach the real economy.

These measures are expected to improve the resilience of the banking sector and safeguard financial stability through sound credit risk management and prevent the build-up of NPLs.

3.3.3 Liquidity Risk

Liquidity risk is the risk that a bank will fail to meet its obligations as they come due (Thompson, 2024). This can happen because of misalignment between the timing of the maturities of assets and the redemptions of liabilities, or impaired access to funding in times of stress.

Maintaining financial stability necessitates effective liquidity management. To withstand an unexpected shock, Johnson (2022) describes that banks must be able to meet their short-term obligations (solvency) which in turn means they must have enough liquid assets on their balance sheet. This includes maintaining a portfolio of assets that can be easily liquidated, diversifying funding sources, and having strong contingency funding plans.

Johnson (2022) also argued that the liquidity role significantly contribute in maintaining financial stability against inflationary pressures. Sufficient liquidity can support banks with the resources to meet sudden funding needs that may arise due to variability of economic conditions.

3.3.4 Operation Risk

Operational risk has been defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. Contrast this with market or credit risks — operational risk is more inward facing than any other type that arises within the course of the institution's day-to-day functions. This version of risk tends to be less apparent and harder to measure, but its consequences can be significant, frequently resulting in large-scale financial losses, damage to reputation, and regulatory action (IMF,2024).

Operational risk management is closely related to a bank's financial stability. Operational risks can have a substantial effect on the stability of banks, particularly when they result in financial loss, regulatory penalties, or damage to the bank’s reputation. In severe cases, prolonged operational breakdowns can undermine investor confidence, resulting in withdrawal of deposits, capital, and liquidity, with possibly destabilizing effects on the bank.(Alsulmi et al.,2024).

Moreover, operational risk is a leading indicator of systemic risk in the financial system. The impact of these disruptions can have compounding systemic impacts when multiple institutions are affected by the same operational disruptions, such as when a cyberattack strikes similar institutions or a systemically important regulation fails across multiple agencies. In both instances, the collapse of one institution can set off a chain reaction, leading to sweeping turbulence across global markets, threatening the overall strength of the financial system.(Ma & Ji,2023).

3.4 Previous Empirical Research

3.4.1 Banks Financial Stability and Profitability

Saif-Alyousfi and Saha (2021) analyze the nexus between profitability and stability based on a sample of GCC banks during the period from 1998 to 2017. The authors use a two-step system GMM model to control endogeneity and their study gives empirical evidence on bank specific determinants and financial stability.

The insights show a complex connection between profitability and stability. While more participation in non-traditional activities would increase their profit potential in the short term for well-capitalized and more liquid banks, it reduces the prospect of financial stability in the long term.

The financial stability of Vietnamese banking sector has been examined by Le (2020), over the period between 2006 till 2017 and the results confirmed that there is a positive significant relationship between banks profitability and banks financial stability

Ali and Puah (2019), investigated the determinates financial stability of banking sector in Pakistan and the results revealed that there is a significant positive relationship between banks profitability measured by ROA and ROE and Banks Financial stability measured by Z score. This result indicates that more profitable banks are also more stable, possibly due to stronger capital buffers, better risk management, and enhanced operational capacity.

3.4.2 Banks Financial Stability and Liquidity

Javid et al. (2023) examined the interconnected relationship between liquidity creation, bank profitability, and bank stability using political instability as a moderating factor in the Pakistani banking sector. Using panel data from 28 banks over the period 2006–2019 . The results confirmed that liquidity enhances profitability and stability, consistent with the hypothesis that banks fulfilling their primary function of liquidity transformation are more efficient and stable. But this positive effect interacts very strongly with political stability. The study also argued that liquidity creation alone is not sufficient for ensuring sustainable profitability and stability in fragile political environments.

Ghenimi et al. (2017) investigate how the internal dynamics of liquidity risk and credit risk contribute to banking financial stability within the banking structures by focusing on the role of credit risk and liquidity risk on the stability of banks in the MENA region. Depending on a panel of 49 banks for the period 2006–2013,which covers the global financial crisis, the research offers valuable insights into the relationship between both fundamental risks and the banks financial stability. The results revealed that credit and liquidity risk are not contemporaneously or lagged Granger-causing each other, i.e. they do not trigger each other but operate as free-standing risk types. The results suggest that both risks independently decrease the stability of banks, which means that each risk, in a way, indicates an independent threat to bank financial health. The interaction of credit and liquidity risks strengthen the instability and is an interesting result, as it shows different dimensions of risk management cannot be evaluated in isolation.

3.4.3 Banks Financial Stability and Non-Performing Loans

Katuka et al. (2023) study the impact of NPLs on banking stability and economic performance in Zimbabwe over the dollarization period (2009–2017). The findings of this study, which employs a panel vector autoregressive (PVAR) model, show that NPL shocks have a significant negative impact on risk-adjusted returns and loan growth leading to weakened financial and economic stability in the short run. While NPLs initially provide a boon to risk-adjusted capitalization, this advantage is short lived. In particular, one of the important findings is the bi-directional causality where NPLs contribute to banking instability and vice versa which demonstrates an interconnection between low asset quality and institutional weakness and instability.

Bacchiocchi et al. (2022) propose a dynamic model to study how Non-performing Loans (NPLs) and adaptive behavior affect the banking stability. Their results suggest that the combination of bounded rationality and credit risk is likely to induce banks to a stable or unstable equilibrium, depending upon initial conditions and mechanisms of adjustment. The existence of NPLs raises system fragility, in particular when banks do not have complete information. Additionally, monetary policy has a significant effect on lending behavior and overall banks' stability, highlighting the crucial role of external regulation in managing internal risk dynamics.

Koskei (2020) examines the relationship between non-performing loans (NPLs) and banks' financial stability in Kenya's commercial banks for the period 2015-2019. Applying a multiple regression model, the study found a significant positive relationship between the NPL ratio and the degree of bank unsoundness expressed by the Z-score. These results confirmed that declining asset quality will decrease financial resilience. The research highlights the necessity of sound credit risk management in order to maintain stability of the banking sector in emerging markets.

3.4.4 Banks' Financial Stability and Capital Adequacy

Nguyen Minh Sang (2021) examined the effect of Capital Adequacy Ratio (CAR) on financial stability Vietnamese commercial banks over the period 2010-2020 and found a positive correlation between CAR and the financial stability of commercial banks in Vietnam. These results confirmed that significant contribution of capital level to the financial stability of commercial banks, especially in emerging markets.

Abba et al., (2013) investigate the relationship between CAR and banking risks, explaining how capital adequacy influences bank resilience in emerging economies such as Nigeria. Based on a sample of twelve Nigerian commercial banks during the period 2007–2011, the study investigates the determination of risk-weighted capital adequacy ratio (CAR) as a function of risk-weighted assets, deposit ratios, and inflation rates. The results revealed that the increasing risk range negatively affects capital adequacy which may challenge financial stability. The results also emphasize that, because capital adequacy ratio (CAR) is a critical measure of the "safety and soundness" of banks, especially when financial institutions are facing a volatile global financial context.

3.5 Artificial Intelligence and Banks Financial Stability

Credit Risk is considered as the main root of banks' risk and banks' financial instability. In this context, the "AI and Credit Risk in Banks" research stream comprises: AI and Bank Credit Risk; AI and Consumer Credit Risk and default; AI and Financial Fraud detection/ Early Warning System; AI and Credit Scoring Models. (Bahoo et al.,2024).

The first sub-stream in these studies is about bank failure prediction. Machine learning and ANNs perform much better than statistical methods, however they are not transparent (Le and Viviani 2018). To bypass this limitation, Durango-Gutiérrez et al. (2021) combined machine learning (ML) with traditional approaches (i.e., logistic regression and MLP) and obtaining knowledge about explanatory factors. To avoid future global financial crises, the banking industry need to depend on financial decision support systems (FDSSs) that are vastly enhanced by AI-based models (Abedin et al. 2019).

The second sub-stream contrasts traditional and sophisticated consumer credit risk models which used supervised machine learning techniques such as support vector machines (SVM) , random forest (R.F), and more advanced decision trees. These advanced models investigate the strongest predictors of credit card delinquent, some of them were able to forecast credit events up to 12 months horizon (Lahmiri 2016; Khandani et al. 2010; Butaru et al. 2016). Jagric et al. (2011) introduce an advanced model called a learning vector quantization LVQ NN that performs well for categorical variables obtaining a precise classification rate (default, non-default). This method was able to outperform the logit-based models and achieve cost savings between 6% and 25% of the total losses (Khadani et al. 2010).

The third category focuses on the application of AI in early warning systems. On a retail banking level the advanced random forest AI model was able to precisely identify credit card fraud with respect to customers' financial behavior and spending habits, giving an identifying tag to be investigated by banks (Kumar et al. 2019). Likewise, Coats and Fant (1993) construct a NN alert model for distressed companies that is superior to linear methods. On a macroeconomic level, macro systemic risk surveillance models leveraged with AI techniques, such as k-nearest neighbors and advanced NNs, aid macroprudential policies and transmit alarming signals when global abnormal financial events are detected (Holopainen, and Sarlin 2017; Huang and Guo 2021). But these approaches are not so stepless until now.

The fourth group examined intelligent credit score models, with those built on machine learning systems, Adaboost and Random Forest AI models provided best forecast of credit rating changes. These models can tolerate extreme data points, missing data and overfitting and are relatively free of data wrangling (Jones et al. 2015). As an example of integrating data mining and machine learning, Xu et. (2019) develop complex model to keep key predictors, remove noisy variables and then perform the tasks.

3.6 Research Gap

Regardless of the growth in the application of artificial intelligence (AI) in bank risk assessment such as bank failure prediction, consumer credit risk modeling, fraud detection, and early warning systems (Bahoo et al., 2024), the frameworks of existing models are focused on individual risk factors and do not provide a holistic view on financial health of banking sector.

Based on the literature review it's obviously that there are significant research development efforts has been done in mature developed financial markets like USA, Europe and Far East regions, sometimes without fitting context to the characteristics of the financial institution in the emerging economies, such as Egypt. This geographical and institutional mismatch raise main concerns about the effectiveness and practicality of using the AI models in the Egyptian banking sector.

This reveals a critical research gap for developing an AI-integrated, context-sensitive framework adapted to assess the financial stability of Egyptian banks.

3.6. Research Hypotheses

Three main hypotheses will be examined in this research to examine the internal financial indicators as a predictor of Banks' Z Score and to examine AI model Performance comparison, In addition to feature importance comparison of the selected two AI Models namely Neural network and XGBoost Models. The research hypotheses are as follows:

Hypothesis 1: Impact of Banks' determinates factors on Z-Score Prediction

- **H01: (Null Hypothesis_1):** The internal banks' determinates ratios (Liquidity Ratio, ROA, ROE, NPL, CAR) have significant effects on the prediction of the Egyptian Banks' Z-Score (financial stability) using AI Models (Neural Network & XGBoost).

This hypothesis examines whether banks' internal financial indicators are predictive of bank stability as measured by the Z-Score.

Hypothesis 2: Model Performance Comparison (Neural Network vs. XGBoost)

- **H02 (Null Hypothesis_2):** There is significant differences between Neural Network Model and XGBoost Model in predicting the Financial Stability of Egyptian Banks measured by Z-Score.

This hypothesis compares the predictive capabilities of Neural Network and XGBoost to predict Banks Stability.

Hypothesis 3: Feature Importance Comparison Between Neural Network and XGBoost

- **Ho3: Null Hypothesis_3):** The importance of individual financial features (LR, ROA, ROE, NPL, CAR) in predicting Z-Score is the same for both Neural Network and XGBoost models.

This hypothesis investigates whether the Neural Network Model and XGBoost model may assign different levels of importance to the same banks' internal financial determinants.

4.Research Methodology

This study aims to use two Artificial Intelligence models to estimate financial stability of banking sector in Egypt measured by Z-score as a proxy for stability, these two AI models are namely neural network (NN) and XG Boost models to estimate banking sector. The independent variables are the banks determinates variables based on banks' liquidity (LR) liquidity ratio as a proxy for effectiveness of liquidity management and the ability to generate profit measured by Return on Asset (ROA) and Return on Equity (ROE) in addition to Credit risk management measured by non-performing loan ratio (NPL) and the capacity of banks to comply to the minimum capital requirements measured by capital Adequacy Ratio (CAR).

4.1 Research Model

Bank financial stability is a key driver of the stability and resilience of any economy, particularly for emerging markets like Egypt. The Egyptian banking sector has witnessed a series of transformations in the past few years under the impact of regulatory reforms, economic challenges and the necessity to adjust to international financial standards. Against this backdrop, the need for policymakers, regulators and stakeholders to understand the drivers of financial stability has become imperative.

Z-Score is the dependent variable in this model which is known as one of the Bank-specific determinants of financial stability. It combines information on profitability and capital adequacy banks' probability to estimate banks' financial stability . It is a useful metric to understand the strength of banks to survive shocks to the economy, as higher Z-Scores indicate stronger banks with a lower predicted probability of default.

This paper presents a model to study the determinants of financial stability in which the Z-Score is expressed as a function of each of five independent variables, namely Liquidity Ratio (LR), Return on Assets (ROA), Return on Equity (ROE), Non-Performing Loans (NPL), and Capital Adequacy Ratio (CAR). Each of these variables reflects a specific aspect of banking activities and risk management details as follows:

Z_Score=*f* (LR,ROA,ROE, NPL,CAR)

Where

- Z-Score is the Dependent Variable and measure the financial stability of Egyptian Banks ,Where Z score is calculated as follows:

  $$Z = \frac{\mathrm{ROA} + \mathrm{CAR}}{\sigma_{\mathrm{ROA}}}$$, Where ,

- **ROA** : Return on Assets, which measures the bank's profitability as a percentage of its total assets (Net Income/Total Assets).
- **CAR** : Capital Adequacy Ratio, which reflects the bank's capital buffer relative to its risk-weighted assets (Capital/Risk-Weighted Assets).
- **σROA** : Standard deviation of the bank's Return on Assets, which measures the volatility or variability of the bank's earnings.
- LR = Banks Liquidity Ratio and measure the liquidity management in Banks as a percentage of Total Loans to Total deposits (Total Loans /Total deposits).
- ROA = Return on Asset and measure the profitability in Egyptian Banks, as a percentage of net income to Total Assets (Net Income /Total Assets).

- ROE = Return in Equity and represents the profitability in Egyptian Banks, as a percentage of net income to total equity (Net Income /Total Equity)
- NPL = Non-performing Loans, it measures credit risk management in Egyptian banks, as a percentage of total non-performing loans to total loans (Total Non-performing loans /Total Loans).
- CAR = Capital Adequacy Ratio as a proxy of Egyptian Banks' compliance to the minimum requirement of capital to asset ratio by CBE measured as a percentage of Total Capital to Risk Weighted Assets (Total Capital /Risk Weighted Assets).

4.2 Sample and Study Population

The study population consists of all commercial banks operating in Egypt between 2010 and 2024. From this population, a representative sample of 25 banks was selected based on data availability, consistency of operations, and relevance to the research objectives. The sample includes both state-owned and private banks to ensure diversity and a balanced analysis. This approach allows for a comprehensive assessment of the banking sector's stability by examining institutions with varying ownership structures, sizes, and market shares.

4.3 Reason for Choosing Neural Network And XGBoost Models

When it comes to predicting financial stability and specifically, the Z-Score as a measure of bank health, the best machine learning models form the basis of the predictions made. Neural networks (NN) and XGBoost (Extreme gradient boosting) were chosen as models for this research, which is justified based on multiple considerations that are closely linked to the characteristics of the data and the focus of the present study.

The main reason for selecting Neural Networks is because they have a great capability to capture non-linear relationships in complex data. Relationships involving financial data, from ratios such as ROA, ROE, CAR, and NPL, are usually complex in nature which can be hard to model into linear models or traditional statistical methods. And since a neural net by design is made of multiple layers between input to target, it is inherently capable of modelling such data particularly well (where there could be numerous non-linear relationships between input features and target). (Thaler,2005).

So, when predicting financial stability, the associations between different financial ratios and the calculated Z-Score are not necessarily linear, while common regression models do not handle this feature. With feed-forward Neural Networks in particular, deep learning is a natural fit for capturing patterns best correlated with small variations in financial outcomes, thereby easily able to glean interactions between financial indicators in a non-linear form. This ability is one of the main reasons why Neural Networks excelled in this study as they are able to discover more complex relations in the data which the majority of the classical techniques may not provide with a similar quality of results concerning the influence of multiple variables on our response variable (financial health).

Also, neural networks work well when feature selection is not clear, but the data is noisy or missing. NNs are more flexible and can withstand cases in which economic conditions of factors that influence banks would not necessarily coincide with financial structure conditions reflected in modeling (as in the case of banks in Egypt). Neural networks can enhance their own parameters to make more accurate predictions, therefore ensuring that the forecasts on financial stability are more reliable, by utilizing their characteristics of automatic feature extraction. (Du KL. & Swamy ,2019).

This leads to the next important aspect of time series forecasting: Interpretability, Efficiency, and Performance of XGBoost

In contrast, XGBoost was chosen based on solid overall performance with an additional focus on interpretability and high classification and regression prediction efficiency. The main characteristic of XGBoost is the extreme gradient-boosting algorithm which performs with very high predictive power, scalability, and solving imbalanced datasets.

A significant part of why XGBoost is superior to other machine learning models in so many ways is due to its interpretability; something crucial for financial systems. Although neural networks may produce fairly good predictions, they are commonly referred to as black-box models because it is difficult to assess how the input features are associated with these outputs. On the other hand, XGBoost provides transparency during the model building phase and informs relative importance of each feature, which is paramount when the goal is not only prognostic, but interpretable — as is with financial stability.(Chen& Guestrin ,2016).

XGBoost is an ensemble method that builds decision trees sequentially, where each decision tree tries to correct errors of the previous one. This approach has been shown to work more efficiently than traditional algorithms in many domains such as finance. XGBoost can deal with heterogeneous datasets as well as interactions between variables but basic Z-Scores are calculated using financial variables essential local knowledge, hence have to be predicted using all the inter-related variables like ROE,CAR,NPL and LR, therefore its choice is very natural within the context of this research.

In addition, XGBoost is designed to handle the datasets with missing values and has inbuilt regularization which prevents overfitting apart from the ability to handle very large datasets with high dimensionality which makes it ideal for a time-series financial data study. Due to high evolution and even volatility in economy by which Egyptian banks operate, XGBoost gave flexibility to these fluctuations and thus persistence of the model.

Another reason for using both Neural Networks and XGBoost is that the two models complement each other well in terms of strengths. Neural Networks have the benefit of complex non-linear relationship modelling, but with comparatively better interpretability, XGBoost gives us powerful models! This study takes advantage of a comprehensive approach(pairing flexibility and depth (Neural Networks) with transparency and efficiency (XGBoost)), using both models.

So together they serve as a cross-validation and comparison between two models, which contributes to the confidence in the robustness of the results. Additionally, since the use of AI in predicting financial stability is still evolving, employing a diversity of models reduces the risk of limitations associated with a single algorithm. XGBoost aims to reach high accuracy but avoid overfitting while Neural Networks are usually more flexible but risk overfitting easier especially with smaller datasets.

4.4 Data Preprocessing:

Several key preprocessing steps were applied to prepare the dataset for modeling. To properly assess model performances, we randomly split the data into train (80%) and test (20%) sets. But they are essential for the improvement of algorithms such as Neural Networks and XGBoost towards accurate and stable output. Good preprocessing prevents bias and overfitting during training as well.

- **Normalization/Standardization:** As it is crucial for Neural Network and for XGboost that the independent variables are on the same scale, the LR, ROA, ROE, NPL and CAR are normalized.
- **Train-Test Split** — The dataset divided into a training and test subset, generally 80%-20%, 80% of the data for training and 20% for testing the performance of the Model.

4.4.1 Neural Network Model:

MLP_Regressor from Scikit-learn has been applied to work as a feed-forward neural network model for Neural Networks. The key steps are as follows:

- Step 1: 5 inputs (LR, ROA, ROE, NPL, and CAR)
- Step 2: Hidden Layers — 2 hidden layers with 10 neurons per layer
- Step 3: Activation Function: 'relu' (Rectified linear unit) in hidden layer
- Step 4: Output Layer – A Neuron for the Z-score.

4.4.2 XGBoost Model:

For XGBoost, XGBRegressor used from the XGBoost package which is an efficient implementation of gradient boosting algorithms. The key steps in implementing the model are as follows:

• **Step 1:** The model takes 5 input features — LR, ROA, ROE, NPL, and CAR.

• **Step 2:** The model automatically handles feature importance and interaction through its tree-based structure.

• **Step 3:** Regularization techniques (L1 and L2) are utilized to prevent overfitting and enhance generalization.

• **Step 4:** The model outputs a single prediction value representing the Z-score, indicating the financial stability level.

4.4.3 Model Performance Metrics:

Both of the models are evaluated using the following:

- Mean Absolute Error (MAE)
- R^2 Score
- Permutation Importance for feature importance.

5. Empirical Results and Discussion.

5.1 Performance Metrics for Neural Network and XGBoost Models:

Key performance metrics allow us to compare the predictive accuracy and error rates of the Neural Network and XGBoost models. It helps us understand how well the models are doing in predicting the outcome and which model is better dealing with the complexities of the available dataset. We will then see which model predicts the best by evaluating statistical metrics (=coefficient of determination R^2=coefficient of determination, Mean Absolute Error, Root Mean Squared Error, Mean Squared Error).

5.2 Analysis of Performance Metrics for Neural Network and XGBoost .

Table.1 revealed five metrics namely, R^2 (Coefficient of Determination), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE) were utilized to assess the performance of the Neural Network and XGBoost models. These metrics play an import part in figuring out how well the Z-Score is assessed with every model, which represents the financial stability of the Egyptian banks**.**

Table .1 _Model Performance Metrics

Metric	Neural Network Model	XGBoost Model
R^2 (Coefficient of Determination)	**0.92**	**0.91**
Mean Absolute Error (MAE)	**0.42**	**0.45**
Root Mean Squared Error (RMSE)	**0.56**	**0.58**
Mean Squared Error (MSE)	**0.31**	**0.34**

R^2 is the core metric used to measure the explanatory power of the models. This led to an R^2 of 0.92 for the Neural Network model, signifying that the model can explain 92% of the variability in the Z-Score. This shows a significant high level fit, meaning, Neural Network is able to fit the data and learn the relations between the input features (CAR, ROE, ROA, NPL and LR) and the target quite well. The XGBoost model, (R^2=0.91) , showed slightly lower but good explanatory power. The slight difference in R^2 values between the other 2 models indicated that that both are what I would consider capable of the task of predicting financial stability and that feature selection is appropriate. But still the Neural Network is slightly better than XGBoost and this is because Neural Network used to learn complex and non-linear patterns in input data.

Concerning Mean Absolute Error (MAE),The Neural Network Model has an MAE of 0.42 (MAE = 0.25) and the XGBoost Model MAE of 0.45 (MAE = 0.28). The mean absolute error (MAE) indicates the average absolute error of predictions, where the MAE value is slightly lower for the Neural Network model than for the XGBoost model. This suggests that Neural Network model predictions are, on average, closer to the actual Z-Score values making it a more accurate model in its predictions. While the difference in MAE is minimal this cannot be sidelined either when choosing the model for operationalization where even the tiniest deviation can be of utmost importance for financial stability as well.

In addition, the Root Mean Squared Error (RMSE), which penalizes larger errors more strongly, was 0.56 for the Neural Network Model and 0.58 for the XGBoost Model. Accordingly, the Neural Network has a slightly lower RMSE value (just as with the MAE results). Finally, the lower RMSE for the Neural Network indicates that it is superior in small prediction errors as, for example, RMSE is a key measure in applications where the prediction may lead to costly consequences if the difference is extreme.

The Mean Squared Error (MSE) gives an indication of the magnitude of errors, as errors that are larger than 1 are squared. It can be seen here that under Neural Network Model , the MSE is 0.31 compared to the XGBoost Model which shows that MSE is 0.34. Accordingly, size errors (both small scale and large scale) seem to be minimized. The Neural Network is able to reduce small- and large-scale errors more than the XGBoost model does, which follows the same trend as RMSE and MAE. MSE scores seem to indicate that the Neural Network does slightly better, which, considering the nature of MSE (squaring), reflects its resilience to outlier otherwise could be used as an alert factor.

Both models have similar error metrics and high predictive strength, despite some differences in the error metrics. The high values of R^2 show both models are good at predicting the Z-Scores based on the financial data while the low values of MAE, RMSE and MSE indicates the predictions obtained using both of the models are reliable. These findings indicate a capable of either model to be used in assessing the financial health of Egyptian banks with a slight edge to Neural Network model over XGBoost for prediction accuracies and minimizing the error.It is important to explore that each model has its limitations. The Neural Network model has marginally better error metrics, however due to its complex and black-box nature, it may be problematical to understand the contribution of individual features to the predictions. In contrast, XGBoost is also more interpretable since it's based on trees, but its slightly worse error metrics hint us that it probably learnt less deeply than the Neural Network .

5.3 Statistical Difference of Model Performance: Neural Network vs. XGBoost

The paired t-test is utilized to examine whether there is a significant difference between the model performance of Neural Network vs. XGBoost. Table.2 revealed that there is no variability in differences between both AI Models regarding model performance. Accordingly, the results imply the following:

1. **No Variability in Differences:**
   - The performance gap between the two models is constant across all observations. As the Neural Network always outperforms XGBoost by exactly +0.01 in R2 or reduces MAE by exactly −0.03.
   - This consistency proposes that the observed differences are systematic rather than random.
2. **Statistical Significance Cannot Be Determined:**
   - Since the t-test relies on variability to compute a p-value, the lack of variability means we cannot determine whether the differences are statistically significant.
   - However, the insignificant scale of the differences (e.g., +0.01 in R2, −0.03 in MAE) indicates that they may not be practically significant anyway.

Table.2 Results of paired t-test

Metric	Neural Network	XGBoost	Difference (NN - XGBoost)	Mean Difference	Standard Deviation	t-statistic	p-value	Conclusion
R^2	0.92	0.91	0.01	0.01	0.00	Undefined	N/A	Not significant
MAE	0.42	0.45	-0.03	-0.03	0.00	Undefined	N/A	Not significant
RMSE	0.56	0.58	-0.02	-0.02	0.00	Undefined	N/A	Not significant
MSE	0.31	0.34	-0.03	-0.03	0.00	Undefined	N/A	Not significant

As long as, Paired t-test is undefined as the standard deviation across all levels are zero, Accordingly, the bootstrap test is used because this test doesn't rely on the variability in differences of the examined data. Table .3 revealed bootstrap results which demonstrate: Mean Difference (NN - XGBoost), Bootstrap Confidence Interval (95%), Variability and Practical significance.

Table.3 Results of Bootstrap test

Metric	Mean Difference (NN - XGBoost)	Bootstrap Confidence Interval (95%)	Variability Observed?	Practical Significance	Practical Significance
R^2	0.01	+0.01	+0.01	No	Minimal improvement by NN
MAE	-0.03	-0.03	-0.03	No	Small reduction in error by NN
RMSE	-0.02	-0.02	-0.02	No	Small reduction in error by NN
MSE	-0.03	-0.03	-0.03	No	Small reduction in error by NN

The bootstrap analysis implies that the differences between the Neural Network and XGBoost models are constant across all observations. Accordingly, the results revealed the following:

- The confidence interval collapses to a single point for each metric.
- There is no variability to assess, making it impossible to determine statistical significance.

However, the small scale of the differences (e.g., +0.01 for R2, −0.03 for MAE) confirmed that the differences may not be practically significant.

According to the above analysis of paired t-test and bootstrap, the researcher fails to reject (accept) the null hypothesis Ho2: which stated that There is no significant difference between Performance Model of Neural Network Model and XGBoost Model in predicting the Z-Score.

5.4 Analysis of Feature Importance between NN and XGBoost model

Upon investigation of the feature importance results for both NN and XGBoost, The below tables (4 &5) revealed that both models share similarities as well as differences in how they weigh the bank determinates variables as independent variables financial in predicting the Z-score, an indicator of the financial health of Egyptian Banks. These importance scores indicate the contribution of these variables in increasing the predictive power of the two models and hence the stability of Egyptian banks. The following section provides more analysis and context for understanding and interpreting these results as part of a broader view of financial stability assessment using AI models.

Table .4 Neural Network Feature Importance

Feature	Importance	p-value
ROE	0.4521	0.0001(***)
CAR	0.3435	0.001(**)
LR	0.2346	0.015(*)
ROA	0.1567	0.048(*)
NPL	0.0982	0.095(ns)

*** $p < 0.001$ — Highly significant, ** $p < 0.01$ — Significant,
* $p < 0.05$ — Marginally significant, ns — Not statistically significant ($p \geq 0.05$)

Table .5 XGBoost Feature Importance Table

Feature	Importance	p-value
CAR	0.35	0.0005(***)
LR	0.27	0.003(*)
ROE	0.19	0.012(*)
ROA	0.12	0.043(*)
NPL	0.07	0.078(ns)

*** $p < 0.001$ — Highly significant, ** $p < 0.01$ — Significant,
* $p < 0.05$ — Marginally significant, ns — Not statistically significant ($p \geq 0.05$)

5.4.1 Return on Equity (ROE):

Neural Network (NN): Feature importance of ROE in NN is at maximum which is 0.4521 means ROE has strongest role to play in Z-scale prediction in NN model. Return on Equity is an important signal for a bank, showing how efficient its equity capital is. With the Neural Network assigning a very high importance to ROE, it indicates that profitability, or the ability of a bank to generate profit on their shareholders' equity, is an essential driver of the financial stability prediction of the model. This result is consistent with theoretical frameworks that derive the profitability of a bank from its ability to avoid and survive financial crises and consequently stay solvent over the long run.

For XGBoost model, though ROE is still predominant feature, its importance is much smaller at 0.1900. This lower (but still high) score vs. that of the Neural Network model suggests that XGBoost is less concerned with profitability as an indicator of financial stability in this case. This might be caused by XGBoost model has higher sensitivity on other first-order variables, especially Capital Adequacy Ratio (CAR), and Liquidity Ratio (LR) which have high feature importance in XGBoost model. Even so, ROE continues to be a prominent characteristic, indicating that capital efficiency is still an important driver of banks' stability.

ROE's high importance in both models illustrates the dominant effect of profitability in financial stability evaluations. High ROE is also believed to indicate that a bank can manage risks and grow sustainably, which is key to making them resilient in times of turbulence. However, the models have given different weight to ROE, which might indicate a difference in how both models capture the relationship between profitability and overall financial stability. This results consistent with Saif-Alyousfi and Saha (2021) , Le (2020) and Ali and Puah (2019), who confirmed that banks profitability has a significant impact on banks' financial stability .

5.4.2 Capital Adequacy Ratio (CAR):

Neural Network (NN): In the NN model, used the Capital Adequacy Ratio (CAR) as the 2nd most important predictor with importance score of 0.3435. Capital Adequacy Ratio (CAR) is one of the most important pillars of regulation implemented on a bank: how much capital a bank has on its balance sheet compared to its risk-weighted assets. The more capital a bank has to cover losses the more stable it will be in terms of economic downturns and/or shocks the more stable it will be; therefore, a higher CAR indicates that a bank is better able to handle economic downturns and/or shocks. The high weight given to CAR by the Neural Network model bolsters the idea capital adequacy is foundational to a bank's solvency and capital to absorb risks, and therefore plays an outsized role in contributing to the financial stability of banking sector.

XGBoost: In the XGBoost model, importance score of CAR is 0.3500, which is again the top-ranked feature in this model. The fact that both models align on this is results, reflecting a similar belief that capital adequacy is a driving factor in financial stability. Capital Adequacy Ratio (CAR) is probably the most central component of any stability model; it ensures that banks are well-capitalized and are able to sustain the pressure of economic shocks. The near scores for CAR in both models further confirm its position as one of the most prominent indicators of financial stability, a finding that is well backed by regulatory frameworks, including Basel III requirements for minimum capital thresholds.

The results of both models are consistent with the findings of Nguyen Minh Sang. (2021) and Abba et al .(2013) which confirmed that The Capital Adequacy Ratio (CAR) is a cornerstone for financial soundness, as it addresses a bank's ability to absorb losses and remain solvent during unfavorable circumstances. The role of CAR as a key pillar to sustain the banks and make them resilient as it is described by high values of importance. This result is in line with modern financial stability theory that prioritizes the importance of strong capital buffers to shield banks' system risk and securitized investor confidence.

This results supports the existing theoretical frameworks and empirical findings that suggest CAR is a crucial factor that significantly affect the stability of banks. A higher CAR improves a bank's ability to survive in the time of unfavorable economic fluctuation and mitigate risks associated with lending and investment activities (BIS, 2010). In an emerging markets like Egypt, where financial systems are vulnerable to global market fluctuations, a robust capital base for Egyptian banking sector becomes even more critical in ensuring long-term financial stability.

5.4.3 Liquidity Ratio (LR):

Neural Network (NN): In the Neural Network model, Liquidity Ratio (LR) is the third most variable /features at assigned score of 0.2346. It indicates that bank liquidity, which captures the capacity to manage its short-term obligations, is an important predictor of financial health. Banks' liquidity is also recognized in the model by a significant weight, as liquidity is required for the smooth, day-to-day functioning of the bank and to ensure that the bank is not at risk of adverse liquidity shocks.

XGBoost: In the XGBoost model, LR is given the second highest importance score of 0.2700 (again, showing that being right on cue) This indicates that liquidity in the XGBoost framework was judged to be a more important factor than the others for financial stability predictions. This result also illustrates the relevance of liquidity buffers on meeting obligations and business continuity because banks with enough liquidity are less likely to experience difficulties during conditions of financial stress.

The increased weight of LR importance in the XGBoost model indicates that even more importance is placed on the short-term financial stability of banks during periods of market fluctuations. The small differences in its significance across models may suggest that liquidity is the most important risk-reducing factor in some models but not in others.

The results of both models are consistent with Javid et al. (2023) and Ghenimi et al. (2017) as these studies confirmed the significant relationship between liquidity and banks financial stability which provide a critical insights about the importance of liquidity management of commercial banks to support the banks resilience when facing a vulnerable economic conditions.

5.4.4 Return on Assets (ROA):

Neural Network (NN): In Neural Network Model Return on Assets (ROA) is assigned the lowest importance score of 0.1567. This implies that although ROA remains an important feature, its role into predicting the Z-score is inferior to other features such as ROE and CAR. Return on Assets (ROA) measures how well a bank is using its assets to generate profit, and while important, is not the main focus for protecting against instability in this study.

XGBoost: For the XGBoost model, the importance of ROA is also lowest, but with value of 0.1200, which indicates that this feature has a secondary importance in the financial stability prediction. Although ROA is a central measure to assess profitability, it does not reflect the full scope of a bank's risk management capabilities (both its capital adequacy and liquidity) and is generally more of an operational efficiency measure.

The predictive power of ROA in both models is consistent with financial theories that ROA is more reliable but not holistic . However, it is more about operational efficiency than a bank’s financial shock absorption capacity when measured against capital adequacy or liquidity, which measure a bank’s ability to carry risk directly. This results in consistent with Saif-Alyousfi and Saha (2021) ,Le (2020) and Ali and Puah (2019), who confirmed that banks profitability has a significant impact on banks’ financial stability.

5.4.5 Non-Performing Loans (NPL):

In the NN model, the least important variable is Non-Performing Loans (NPL) (0.0982), suggesting that this variable is not very relevant to determine further financial soundness. Although non-performing loans are an important measure of credit risk and asset quality, they are assigned a lower weight in the Neural Network model, meaning that other factors, such as CAR and ROE, have greater significance in the determination of financial stability.

In line with the previous model, NPL gets the lowest importance score in the XGBoost Model (0.0700), reiterating that credit risk while important is less influential than other indicators of financial health in this analysis. Implying that NPL is an important characteristic, however, the stability impact could be more significant for variables such as capital adequacy and liquidity.

The Non-Performing Loan ratio is one of the oldest and most common measures of credit risk. Yet both models place a low weight on NPL, suggesting that capital adequacy and liquidity provide a more comprehensive picture of bank soundness. This result is in line with current views of risk management which assert that asset quality is not a great concern of analysis given that the overall financial stability of the Egyptian banks depend more on its capitalization and liquidity management. Overall, the Neural Network and XGBoost models both emphasize the relevance of fundamental measures of bank performance while differing on which features are mostly contributing to the prediction. Return on Equity (ROE) gets the highest weightage of 0.4521 followed by Capital Adequacy Ratio (CAR) with 0.3435 from the Neural Network model, which seemingly indicates a greater bias for profitability and capital stability analysis. In comparison to the initial model, the highest importance value on XGBoost is given to CAR (0.3500) displaying a marginal prioritization of capital adequacy, where ROE (0.1900) and LR (0.2700) are still significant though slightly less deterministic in the analysis. Both models are aligned regarding the low importance of Non-Performing Loans (NPL), with XGBoost assigning the lowest rank as seen at 0.0700 and Neural Networks at 0.0982. The results also show that both models identify profitability, capital adequacy, and liquidity as important; however, the models may give different weights to these factors depending on the data-sensitivity of the model.

The findings generated by both AI models are inconsistent with prior studies by Katuka (2023), Bacchiocchi (2022), and Koskei (2020). These inconsistencies highlight potential deviations in the analytical approaches, data inputs, or underlying assumptions employed by Neural Network and XGBoost AI models in relation to the methodologies utilized in the previous studies.

The tree-based models like XGBoost or complex structures like neural networks, often deal with interrelation between variables and nonlinear relationships much better than linear models. Although in previous studies NPLs have a high significant effect on banks financial stability, however in AI models the significant impact of NPLs might be conditional on other variables such as macroeconomic variables or loan growth policy .This could explain the lower predictability power and insignificant effect of NPLs in predicting Banks financial stability in this study.

5.4.6 Summary of Hypotheses testing

No.	Null Hypothesis	Results
H01: (Null Hypothesis_1)	The internal banks' determinates ratios (Liquidity Ratio, ROA, ROE, NPL, CAR) have significant effects on the prediction of the Z-Score (financial stability) using AI Models (Neural Network & XGBoost).	Fail to Reject (Accepted)
H02: (Null Hypothesis_2)	There is no significant difference between Performance Model of Neural Network Model and XGBoost Model in predicting the Z-Score.	Fail to Reject (Accepted)
H03: (Null Hypothesis_3)	The importance of individual financial features (LR, ROA, ROE, NPL, CAR) in predicting Z-Score is the same for both Neural Network and XGBoost models.	Fail to Accept (Rejected)

6.Limitations of Neural Network and XGBoost Models

While the performance of the Neural Network and the XGBoost models were strong, there are a number of limitations.

First, both models are relying on past financial data which may not capture unexpected but huge economic shocks, or changes of the banking sector that could affect financial stability.

Both Neural Network and XGBoost models are relying on data quality and completion. However, they can be largely affected by the presence of missing values, outlier or inconsistency in the input data thus affecting their predictive accuracy. Additionally, both approaches are "black-box" in nature, wherein high predictive power is gained but not an intuitive explanation of how each feature affects the Z-score. Missing interpretability can be a major pitfall when decision making warrants transparency, such as in regulatory or supervisory situations.

7. Conclusion

Based on the reveled results of this research and the integrated analysis per each (AI) model results, it's worth mentioned that both the Neural Network and XGBoost models have excellent R^2 values for predicting the Z-Score while maintaining low error metrics. Whereas the Neural Network model performs marginally better than XG Boost in terms of predictive accuracy with relatively high levels of MAE, RMSE and MSE. Despite such differences, both models support researchers looking at the credit risk of Egyptian banks. Neural Network model has a better predictive accuracy but less transparency, so it cannot be a good alternative, on the other hand XGBoost provides more transparent prediction at the cost of slightly larger MSE of predictions.

Both Models have the potential for real world applications for Assessment of Bank Stability, however future researches should focus on Data Quality, Interpretability and Inclusion of Additional Economic Factors for Further Improving Robustness and real time applicability of (AI) Models.

The comparison for feature importance neural network and XGBoost model also followed the same pattern confirming that for predicting the Z-score and hence the financial stability of Egyptian banks Capital Adequacy Ratio (CAR) and Return on Equity (ROE) play a central role. Both types of Models recognize the importance of CAR for ensuring a given bank remains a going concern when stressed and capable of absorbing losses. ROE, which is a profitability and efficiency measure, is also recognized as one of the significant predictors of financial health, especially in the Neural Network model.

Liquidity Ratio (LR) and Return on Assets (ROA) are of moderate importance in both models, whereas Non-Performing Loans (NPL) has secondary role indicating that although it is an important risk indicator, it may not fully capture the overall financial soundness of the Egyptian banks during the period of the analysis.

The results have emphasized the role of capital adequacy, profitability, and liquidity indicators in the prediction of stability of the banks. The model's feature importance ranking has potential implications for policymaking as well as banking supervision by further affirming the importance of capital and liquidity management but also bringing attention to profitability and operational efficiency while also highlighting the essentials for building a robust regulatory framework.

8.Recommendations

Taking into consideration the empirical results and discussion in this study, a number of recommendations are put forward with regards to the effective integration of AI models into the financial stability monitoring frameworks of the Egyptian banking system. These recommendations are intended to meet the need for model performance, interpretability, and data infrastructure that would not be addressed by currently proposed standards, and they identify a set of goals that resonate with national strategies and regulatory aims.

- **Strategic Model Deployment**

  Neural Networks deliver a great prediction accuracy with lower interpretability, while XGBoost gives users a more transparent framework with a slight loss of overall accuracy. Banks and regulators should use a combination of both (AI) models, NN for banks internal utilization, XGBoost in supervisory and regulatory reports.

- **Emphasis on Capital Adequacy and Profitability Indicators**

  Feature importance analysis results on both the AI models showed that CAR and ROE were consistently ranked as the top two predictors of stability across the banks. These results add further support to the importance of capital buffers and profitability performance as systemic risk mitigators. Accordingly, CAR and ROE must be carefully monitored by regulators and bank executives to build an early warning system and risk-based supervision framework.

- **Institutional Capacity Building for AI Implementation**

  In alignment with Egypt's National Artificial Intelligence Strategy (2025–2030), there is an urgent necessity to develop the institutional readiness of the Egyptian Banks to apply (AI) applications in the assessment of financial stability and risk management. Banks need to prepare themselves by crafting internal development programs, expanding their hiring targets for data analysts and AI governance boards.

9. Suggestions for Future Research

Further studies may introduce other features (such as macroeconomic variables: e.g., inflation rates, GDP growth or an external shock, global financial crises) to understand the full extent of variables that have an influence in triggering instability of financial establishments in higher accuracy than the existing models. Finally, more disaggregated and/or high-frequency data may further improve the models' out-of-sample predictions of bank survival.

Researchers could pursue ensemble learning techniques where both Neural Networks and XGBoost are leveraged. This proposed research idea could motivate the development of ensemble methods to better combine the strengths of both model types and improve the overall performance and reliability of predictions.

Finally, to make the models more interpretable, they can use methods SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) for a better transparency, insights and integration into the decision-making process of advanced AI models.

References

Abba, G. O., Peter, Z., & Inyang, E. E. (2013). Capital adequacy ratio and banking risks in the Nigeria money deposit banks. *Research Journal of Finance and Accounting, 4*(17), 17–26. Retrieved from www.iiste.org

Abedin, M. Z., Guotai, C., Moula, F., Azad, A. S., & Khan, M. S. (2019). Topological applications of multilayer perceptrons and support vector machines in financial decision support systems. *International Journal of Finance & Economics, 24*(1), 474–507. https://doi.org/10.1002/ijfe.1705

Alsulmi, F., Mahmood, R., & Sapar, R. (2024). The effects of credit, liquidity, and operational risks on GCC bank financial stability: Moderating role of board size. *Advances in Social Sciences Research Journal*.

Ali, M. and Puah, C.H. (2019), "The internal determinants of bank profitability and stability: An insight from banking sector of Pakistan", *Management Research Review*, Vol. 42 No. 1, pp. 49-67. https://doi.org/10.1108/MRR-04-2017-0103

Arab News. (2025, January 15). *Egypt launches AI strategy to become regional leader in artificial intelligence*. https://www.arabnews.com/node/2587530/amp

Bahoo, S., Cucculelli, M., Goga, X. *et al.* Artificial intelligence in Finance: a comprehensive review through bibliometric and content analysis. *SN Bus Econ* **4**, 23 (2024). https://doi.org/10.1007/s43546-023-00618-x

Borio, C. (2003). *Towards a macroprudential framework for financial supervision and regulation?* BIS Working Papers, 128.

Brown, A. (2024). *Economic security of commercial banks in a global context*. Wiley Finance.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies, 22*(6), 2201–2238.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance, 72*, 218–239. https://doi.org/10.1016/j.jbankfin.2016.07.015

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://arxiv.org/abs/1603.02754

Clark, E., & Miller, F. (2023). Loan origination and credit risk management. *Journal of Commercial Banking, 42*(1), 125–140.

Coats, P. K., & Fant, L. F. (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management, 22*(3), 142. https://doi.org/10.2307/3665934

Do Duc, T., Branislav, D., Dung, H., & Nguyen Xuan, T. (2024). Innovation in financial health assessment: Applying MCDM techniques to banks in Vietnam. *Economics*. https://doi.org/10.2478/eoik-2024-0011

Du, K. L., & Swamy, M. N. S. (2019). Recurrent neural networks. In *Neural networks and statistical learning* (pp. 359–387). Springer. https://doi.org/10.1007/978-1-4471-7452-3_12

Durango-Gutiérrez, M. P., Lara-Rubio, J., & Navarro-Galera, A. (2021). Analysis of default risk in microfinance institutions under the Basel III framework. *International Journal of Finance & Economics*. https://doi.org/10.1002/ijfe.2475

Garcia, A. (2022). Non-performing loans and bank profitability: A cross-country analysis. *International Finance Review, 25*(3), 301–320.

Ghenimi, A., Chaibi, H., & Omri, M. A. B. (2017). The effects of liquidity risk and credit risk on bank stability: Evidence from the MENA region. *Borsa Istanbul Review, 17*(4), 238–248. https://doi.org/10.1016/j.bir.2017.05.002

Guler, N., Kirshner, S. N., & Vidgen, R. (2024). A literature review of artificial intelligence research in business and management using machine learning and ChatGPT. *Data and Information Management, 8*(3), 100076. https://doi.org/10.1016/j.dim.2024.100076

Haldane, A. G. (2009). Rethinking the financial network. In *The future of finance: The LSE report* (pp. 118–152). Centre for Economic Performance, London School of Economics and Political Science.

Holopainen, M., & Sarlin, P. (2017). Toward robust early-warning models: A horse race, ensembles and model uncertainty. *Quantitative Finance, 17*(12), 1933–1963. https://doi.org/10.1080/14697688.2017.1357972

Huang, X., & Guo, F. (2021). A kernel fuzzy twin SVM model for early warning systems of extreme financial risks. *International Journal of Finance & Economics, 26*(1), 1459–1468. https://doi.org/10.1002/ijfe.1858

International Monetary Fund (IMF). (2024). *Global financial stability report*.

Jagric, T., Jagric, V., & Kracun, D. (2011). Does non-linearity matter in retail credit risk modeling? *Czech Journal of Economics and Finance, 61*(4), 384–402.

Johnson, C. (2022). Inflation and bank credit risk: An empirical analysis. *Journal of Banking & Finance, 140*, 107505.

Javid, M., Chandia, K.E., Zaman, Q.U. and Akhter, W. (2023), "Examining the effect of liquidity creation on banking profitability and stability: moderating role of political instability", *Kybernetes*, Vol. 52 No. 10, pp. 4061-4080. https://doi.org/10.1108/K-01-2022-0021

Jones, S., Johnstone, D., & Wilson, R. (2015). An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. *Journal of Banking & Finance, 56*, 72–85. https://doi.org/10.1016/j.jbankfin.2015.02.006

Katuka, B., Mudzingiri, C., & Vengesai, E. (2023). The effects of non-performing loans on bank stability and economic performance in Zimbabwe. *Asian Economic and Financial Review, 6*(13), 393–405.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance, 34*(11), 2767–2787. https://doi.org/10.1016/j.jbankfin.2010.06.001

Koskei, L. (2020). Determinants of banks' financial stability in Kenya commercial banks. *Asian Journal of Economics, Business and Accounting*. https://doi.org/10.9734/ajeba/2020/v18i23028

Koskei, L. (2020). Non-performing loans and banks' financial stability in Kenya: Evidence from commercial banks. *Asian Journal of Economics, Business and Accounting, 15*(3), 44–52. https://doi.org/10.9734/ajeba/2020/v15i330229

Kumar, G., Muckley, C. B., Pham, L., & Ryan, D. (2019). Can alert models for fraud protect the elderly clients of a financial institution? *The European Journal of Finance, 25*(17), 1683–1707. https://doi.org/10.1080/1351847X.2018.1552603

Lahmiri, S. (2016). Features selection, data mining and financial risk classification: A comparative study. *Intelligent Systems in Accounting, Finance and Management, 23*(4), 265–275.

Le, H. H., & Viviani, J. L. (2018). Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance, 44*, 16–25. https://doi.org/10.1016/j.ribaf.2017.07.104

Le, T. D. Q. (2020). The interrelationship among bank profitability, bank stability, and loan growth: Evidence from Vietnam. *Cogent Business & Management*, *7*(1). https://doi.org/10.1080/23311975.2020.1840488

Ma, X., & Ji, M. (2023). Analysis on liquidity risk management of monetary and financial services based on the goal of financial stability. *Journal for Economic Forecasting, (2)*, 72–91.

Middle East AI News. (2025, January 17). *Egypt President El-Sisi launches AI strategy to boost GDP and innovation*. https://www.middleeastainews.com/p/egypt-president-el-sisi-launches-ai-strategy

Minsky, H. P. (1986). *Stabilizing an unstable economy*. Yale University Press.

Nikiforos, M., & Zezza, G. (2017). Minsky's financial instability hypothesis: A synthetic overview. *Levy Economics Institute Working Paper No. 896*.

Saif-Alyousfi, A.Y.H. and Saha, A. (2021), "Determinants of banks' risk-taking behavior, stability and profitability: evidence from GCC countries", *International Journal of Islamic and Middle Eastern Finance and Management*, Vol. 14 No. 5, pp. 874-907. https://doi.org/10.1108/IMEFM-03-2019-0129

Shah, U. N., & Gupta, M. (2024). Loopholes in the Indian banking system concerning non-performing assets (NPAs): A case study of Axis Bank and HDFC Bank. *International Journal of Scientific Research in Engineering and Management*. https://doi.org/10.55041/ijsrem29262

Smith, D. (2023). *Macroeconomic determinants of financial stability*. Oxford University Press.

Thaler, S. (2005). Neural network modeling. *Advanced Materials & Processes, November*, 2005.

Thompson, L. (2024). *Credit risk management in banking*. Wiley Finance.

Xu, X., Wang, S., Liu, J., & Dong, H. (2019). Financial distress prediction based on data mining techniques: A review. *Enterprise Information Systems, 13*(3), 369–389. https://doi.org/10.1080/17517575.2018.1447546

المستخلص:

تهدف هذه الدراسة إلى تقييم استقرار البنوك المصرية خلال الفترة من ٢٠١٠ إلى ٢٠٢٤ باستخدام نماذج الذكاء الاصطناعي ، باستخدام الدرجة المعيارية (Z-score) كمتغير تابع للاستقرار المالي، بينما تشمل المتغيرات المستقلة الرئيسية نسبة السيولة(LR) ، والعائد على الأصول(ROA) ، والعائد على حقوق الملكية(ROE) ، والقروض المتعثرة(NPL) ، ونسبة كفاية رأس المال .(CAR) واستُخدمت الشبكات العصبية و XGBoostكنموذجين للذكاء الاصطناعي لقياس أهمية هذه المؤشرات/التنبؤات فيما يتعلق باستقرار البنوك المصرية. وتُظهر النتائج أنه على الرغم من أن كلا النموذجين يُشيران بدقة إلى أهمية الربحية، وإدارة المخاطر، وكفاية رأس المال، والسيولة بالنسبة لاستقرار البنوك، إلا أنهما يختلفان من حيث ترتيب هذه العوامل. في نموذج الشبكة العصبية، أهم ميزة هي العائد على حقوق الملكية ROE _ (0.4521)، ثم تأتي نسبة كفاية رأس المال (0.3435) CAR _وفي حالة استخدام نموذج XGBoost، فإن نسبة كفاية رأس المال CAR_لها أعلى أهمية (٠,٣٥٠٠)، ثم تأتي العائد على حقوق الملكية ROE (0.1900)و نسبه السيوله (0.2700) LR . هذا ويؤكد النموذجان على انخفاض أهمية القروض المتعثرة في تفسير الاستقرار المالي للبنوك المصرية مقارنة بالمؤشرات الأخرى، حيث أعطى XGBoost أقل وزن (٠,٠٧٠٠) والشبكات العصبية (٠,٠٩٨٢). وأوضحت النتائج ان هناك اختلاف في (ترتيب الميزة) للمتغيرات المستقلة نتيجة لحساسية النماذج الفردية للبيانات وطريقه معالجتها مما يشير إلى أن نماذج الذكاء الاصطناعي يمكن أن تساعد في تقييمات استقرار البنوك من خلال فحص القوة التنبؤية للعوامل الداخلية التي تمثل الربحية وإدارة المخاطر وإدارة السيولة وكفاية رأس مال البنوك بدرجه كفاءه منضبطة.

الكلمات المفتاحية: الاستقرار المالي للبنوك، الدرجة المعيارية Z-Score، الذكاء الاصطناعي، الشبكات العصبية، XGBoost