AMERICAN UNIVERSITY OF BEIRUT

PREDICTING BANKING CRISES USING MACHINE LEARNING: THE CASE OF LEBANON

by
LEA NIZAR EL HALABI

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science in Business Analytics
of the Suliman S. Olayan School of Business
at the American University of Beirut

Beirut, Lebanon
April 2023

AMERICAN UNIVERSITY OF BEIRUT

PREDICTING BANKING CRISES USING MACHINE LEARNING: THE CASE OF LEBANON

by

LEA NIZAR EL HALABI

Approved by:

Dr. Victor Araman, Professor — Advisor
BIDS Track, Suliman S. Olayan School of Business

Dr. Ibrahim Jamali, Associate Professor — Advisor
FAME Track, Suliman S. Olayan School of Business

Date of thesis defense: April 28, 2023

AMERICAN UNIVERSITY OF BEIRUT

THESIS RELEASE FORM

Student Name: El Halabi Lea Nizar
Last First Middle

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my thesis; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

☒ As of the date of submission

☐ One year from the date of submission of my thesis.

☐ Two years from the date of submission of my thesis.

☐ Three years from the date of submission of my thesis.

lea el halabi May 4, 2023

Signature Date

ACKNOWLEDGEMENTS

First, I would like to thank my family and friends for believing in me and supporting me through this journey, in all its ups and downs.

Second, special thanks are due to my advisors: Dr. Jamali and Dr. Araman.
Their constant support, guidance, expertise, and knowledge were of immense value to this work. Also, appreciation to Dr. El-Fakhani is due, for the insightful comments.

My gratitude is finally addressed to Dr. Gregor von Schweinitz for inspiring this work and for the valuable contribution.

Without you, this work wouldn't have been possible.

Thank you.

ABSTRACT
OF THE THESIS OF

Lea Nizar El Halabi for Master of Science Business Analytics
Major: Finance

Title: Predicting Banking Crises Using Machine Learning: The Case of Lebanon

This thesis revisits the existing empirical evidence on the predictability of systemic banking crises using machine learning methods with a particular emphasis on Lebanon's crisis of 2019.

More specifically, the dataset of Laeven and Valencia (2020) is extended by appending to it Lebanon's systemic banking crisis and the predictive ability of machine learning techniques such as Logit, KNNs, SVMs, Trees and XGBoost is assessed.

Evaluating the methods using the F-1 score and the ROC AUC suggests that the best-performing models over the testing period are the KNNs and XGBoost.

TABLE OF CONTENTS

ILLUSTRATIONS

Figure

TABLES

Table

CHAPTER I

INTRODUCTION

Banking and financial crises have been a recurrent phenomenon over time, across the world. In fact, 268 banking crises have been accounted for over a period of 208 years (Reinhart and Rogoff, 2009). Additionally, the frequency of banking crises seems to have increased following financial liberalization (Bordo et al., 2001). Such crises emerge unexpectedly and often sound the end of a period of growth. Many theories discuss the causes of banking and financial crises. Laeven (2011) notes that some of these theories support the claim that banking crises are a result of unsustainable macroeconomic policies whereas other theories state that bank runs, and depositor fear, are the primary drivers of banking crises. As banks play a central role in the distribution and allocation of capital (i.e., as financial intermediaries), the depth of banking crises and their effects on the economy can be detrimental and even catastrophic.

A sizable literature examines the predictability of systemic banking crises. In the age of data, Machine Learning (ML) techniques have also been applied for the same purpose. Beutel et al. (2019) apply machine learning algorithms and models on data retrieved from developed countries and evaluate the predictive performance of each model. They find that the Logit model appears to have the best predictive performance.

A. Research Objectives

As Lebanon's financial, banking and economic crisis unfolds, we revisit the evidence on the predictability of banking crises by enlarging the existing data set with the addition of Lebanon's 2019 banking crisis to Laeven and Valencia (2020)'s dataset. The research question of interest is the following: how accurately do the existing predictors of banking crises signal the onset of the 2019 Lebanese banking crisis? How does Lebanon's addition change our empirical understanding of the predictors of financial crises? Is there a commonality in the predictors of banking crises?

The goal of this study is to examine whether it is possible to predict banking crises using ML techniques while specifically focusing on the case of Lebanon. Also, the research aims to investigate whether studying the case of Lebanon would generate new insights regarding banking crises' predictors. This research aims to further shed light on the different benefits that such models can bring to the banking and financial industry, especially in a country with heavy political inertia, significant association between policy makers and bank owners and a hindered independence of the financial regulator. Eventually, we aim to contribute to Laeven & Valencia's Database (latest updated: 2020) by adding the case of Lebanon. The findings have important policymaking implications in that they can would benefit the regulators and Central banks in monitoring signals that that might indicate potential occurrence of crises and implement the necessary measures and take the required action to protect depositors' interests.

B. Research Methodology

1. Literature Review

This part aims to explore scholarly research and publications related to banking crises and machine learning applications in predicting banking crises.

2. Data Collection

Data is retrieved from different sources: the main source for Banking Crises will be the Laeven & Valencia Database (latest update: 2020). Macroeconomic data is retrieved from the databases of the IMF's International Financial Statistics, the St. Louis Fed (FRED), the Bank of International Settlements (BIS), the OECD, countries' central banks, Blom Bank's BRITE, and the European System of Central Banks and the European Systemic Risk Board.

3. Data Processing

We identify variables that might represent significant early warning indicators of banking crises. From the review of related work that has previously been conducted, the following main predictors: Credit expansion:

Change in Debt to Private Sector (in real terms), Non-Performing Loans (to Gross Loans), Change in Debt Service (Debt Service Ratio as a percent of Exports), Change in Real Narrow Money (M1), Change in Real Broad Money (M3), Change Public Debt, Change in Inflation, Real and Nominal effective Exchange Rates and Oil prices. We preprocess variables into features to better identify relevant signals of potential banking crises. Preprocessing procedures include but are not limited to applying feature normalization or scaling and encoding of events or categorical variables). In our data, the occurrence of a crisis takes a value of 1 and a 0 otherwise. As for the remainder of the data is numerical and expressed as period-to-period percentage change.

4. Modeling & Testing

Since this research is inspired by the work of Beutel et al., similar models are used in our study.

Building explanatory models: Logistic Regression, XGBoost, Decision Trees, Random Forests, Support Vector Machines and Neural Networks, with Logistic Regression being the benchmark. Evaluating and Comparing the Predictive Performance

of the different models using the metrics employed by Beutel et al., such as the Relative Usefulness, Area Under the ROC Curve (AUC), and the F-1 Score.

5. Analysis & Recommendation

Based on the results of the evaluations of the different models, we determine which Machine Learning Model is best at predicting banking crises in Lebanon when applied on unseen data. Analysis of results is supported by the necessary visuals.

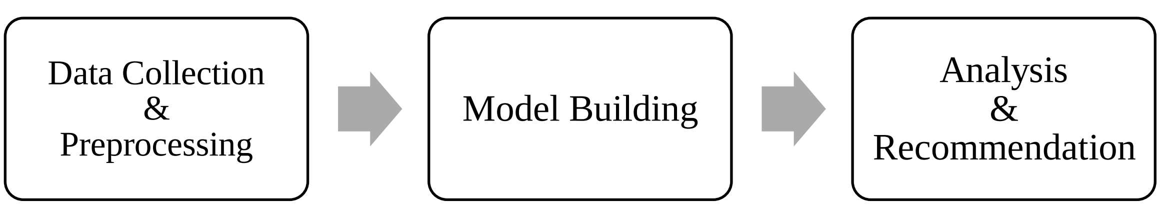


Figure 1: Research Methodology

The remainder of this paper is organized as follows. Chapter Two summarizes the previous literature on defining and identifying banking and financial crises, determining the drivers of crises along with their causes and consequences. Also, a summary on the literature discussing the application of Machine Learning in predicting crises is presented. Chapter Three discusses the data collection process, as well as the limitations in data availability. Moreover, the third chapter presents the processing applied on the data to ensure consistency and robustness. Chapter Four discusses several machine algorithms, such as the logistic regression, decision trees, random forests, K-nearest-neighbors, support vector machines and neural networks. Models are trained on in-sample data and evaluated on test data. The models are then compared based on the different evaluations metrics and the model which outperforms others will be selected. Chapter Five reports the most significant predictors of banking crises and discusses the importance of close monitoring of said variables. Finally, Chapter Six presents the recommendations and suggestions for further researchers and recommendations for at the industry level.

CHAPTER II

LITERATURE REVIEW

Banking and financial crises have been a recurrent phenomenon over time. Additionally, the frequency of banking crises seems to have increased following financial liberalization (Bordo et al., 2001). Such crises emerge unexpectedly and often sound the end of a period of growth. Many theories discuss the causes of banking and financial crises. As banks play a central role in the distribution and allocation of capital, the depth of banking crises and their effects on the economy can be detrimental and even catastrophic. Much research has been conducted with the aim of predicting these banking crises, whether through developing prediction models (Beutel et al., 2019) or early warning indicators (Babecký et al., 2014 and Drehmann et al., 2014). In the age of data, advanced statistical, econometric and machine learning techniques have been applied with the aim to predict banking and financial crises. Our interest to explore this topic further was elicited by the simultaneous crises that have been ongoing in Lebanon since October of 2019. As a first step in this research, an in-depth literature review is conducted to understand what previous research has been done in this scope and identify gaps for the purpose of defining our contribution.

This section starts by presenting some research that has been conducted in the pursuit of a better understanding of banking crises. It proceeds to discuss the previous application of statistical, econometric and machine learning techniques in predicting banking crises.

Finally, the simultaneous crises that Lebanon has witnessed and continues to witness, along with the effects they had on the economy are portrayed.

A. Overview of Banking & Financial Crises

Banking and financial crises emerge unexpectedly and often sound the end of a period of growth. In fact, banking crises ruin the economy by restricting credit and triggering massive liquidations (Banking Crises: A Review, 2015). In 2007, what seemed to be a US-based financial crisis shook the world when its effects revibrated in international financial markets, requiring governmental rescue actions across the world. Laeven (2011) examines theoretical and empirical literature on causes and consequences of banking crises, summarizes lessons to be learned in relation to resolving banking crises.

Despite having different origins, banking crises tend to display similar patterns and seem to have similar causes: unsustainable macroeconomic policies, market failures, inefficiency in regulation, and government interference. According to Laeven, when crises are not addressed or resolved swiftly, their costs can be large. Laeven then ends his article with proposals to enhance financial stability. It has been observed that there are many common factors between current and previous crises. Banking crises are usually preceded by periods of credit boom and overpriced assets (bubbles) and are typically followed by governmental action to save the financial system and hence the economy. Although the exact timing of the crisis is uncertain, it remains inevitable (Laeven, 2011) and although some experts have warned against excessive credit expansion (Rajan, 2005), such warnings were ignored until it was too late, especially by governments and regulators.

Although banking crises are often accompanied by other financial crises, like bursts in asset prices, currency devaluations, and sovereign debt defaults (Flood and Marion, 1999, Krugman 2000, and Dooley & Frankel 2003, Sturzenegger and Zettelmeyer, 2007, and Kaminsky and Reinhart, 1999), the focus usually falls on banks since these institutions are usually at the center of financial intermediation. Furthermore, and since banks play an essential role in the allocation of funds in the economy, banking crises have shown to have detrimental effects on the economy by paralyzing production, significantly reducing output, and causing high increases in unemployment levels (Laeven, 2011).

In the paper "Financial Crises: A Survey," Amir Sufi and Alan M. Taylor provide an overview of the history and causes of financial crises. The authors argue that financial crises are a recurring feature of modern capitalist economies, and that understanding their causes and consequences is crucial for policymakers. The authors begin by defining financial crises and discussing the various types of crises that have occurred throughout history. They note that financial crises can have a range of causes, including asset price bubbles, banking panics, and sovereign debt crises, among others. Sufi and Taylor then examine the various theories of financial crises, including the "Minsky moment" theory, which argues that crises occur when financial institutions and investors become overly confident and take on excessive risk. They also discuss the role of macroeconomic imbalances, such as trade imbalances and fiscal deficits, in causing crises. The authors also explore the consequences of financial crises, including the impact on economic growth, inequality, and political stability. They argue that financial crises can have long-lasting effects on the economy and society, and that mitigating these effects requires careful policy interventions.

Finally, Sufi and Taylor discuss the policy responses to financial crises, including monetary and fiscal policy measures, as well as financial sector reforms. They note that policymakers face a trade-off between providing liquidity to the financial system and moral hazard, and that finding the right balance is crucial for avoiding future crises (Sufi and Taylor, 2018).

B. Explaining Crises

The book titled "Financial Crises: Causes, Consequences, and Policy Responses" consolidates the efforts of the world's best experts in the field and presents the different types of financial crises, the characteristics and causes, the different response and management methods, the short- and long-term effects, the preemptive ways to avoid them, and finally, how to restructure overhanging debts.

The first chapter of the book: "Financial Crises: Explanations, Types and Implications" is written by Claessens and Kose. It reviews a selected sample of literature on financial crises and attempts to answer three questions: What are the key factors that explain financial crises? What are the main types of financial crises? What are the repercussions of financial crises on the real and financial sectors? Answering the first question: explaining financial crises, the authors highlight the importance of severe movements in asset and credit markets.

Typically, financial crises are often associated with one or more of the following: significant increases in asset prices and credit volume, severe disruptions in external financing supply to players in the economy and in financial intermediation, major balance sheet problems (for individuals, businesses, and the sovereign), and a large-scale government support (through liquidity support and recapitalization). The drivers of crises have been identified in the literature but determining the exact causes remains a challenge. Although fundamental phenomena are regularly observed, crises sometimes appear to be driven by "irrational" factors such as runs on banks, contagion and spillovers, limited arbitrage in times of stress, asset markets busts, credit crunches and fire sales. As the authors note, the notion of "animal spirits" has long been of interest in attempting to explain crises and the earliest mentioned expert is Keynes in 1930. The authors define financial crises as asset and credit booms that eventually turn into busts, then proceed to raise questions that are challenging to answer: why haven't financial market players and/or policy makers been able to anticipate the risks accompanying these asset and credit booms and why haven't there been attempts to control credit expansion and asset price increases?

C. Types of Crises

Claessens and Kose (2014) offer a taxonomy the different types of financial crises in the first chapter of the book. In fact, the authors refer to Reinhart and Rogoff's (2009b) classification of crises into two types: crises that can be identified quantitatively such as Currency and Sudden Stop Crises, and crises that are identified qualitatively or using judgement such as Debt and Banking Crises.

The authors acknowledge that other classifications are plausible, but the different crises will likely overlap. Their identification, dating and frequency of crises are determined analytically and empirically.

a. <u>Currency Crises</u>

A currency crisis can be defined as a devaluation or severe depreciation of the currency which often necessitates a large expenditure of international reserves, imposing capital controls, and / or sharply increasing interest rates (Claessens and Kose, 2014). Theories examining currency crises have progressed and changed alongside the change in the nature of these crises. In fact, research shifted from the focus on fundamental causes of crises to stressing the scope for multiple equilibria and the role of financial components (specifically changes in balance sheets) in triggering currency crises. In this section of the chapter, the authors discuss three generation of models: The first-generation models were mainly developed following the collapse of gold prices in 1970s which impacted the value of currencies in Latin America & other developing countries. These models, also known as "KFG" models – in reference to Krugman (1979), Flood, and Garber (1984) – show that a sudden attack on a fixed or pegged currency can be caused by rational actions and behavior of investors who accurately detect that a government has been operating in significant deficits financed through credit from the central bank. Investors hold the currency until they expect the peg to end. When investors dump the currency, the central bank forcibly loses liquid assets and foreign currency supporting the exchange rate which then results in the collapse of said currency (Claessens and Kose, 2014).

Flood and Marion (1997) explain that, in first-generation models, policies enacted before the attack can translate into a crisis. The second-generation models highlight the importance of multiple equilibria: concerns about a government's willingness to maintain its currency's exchange rate peg could cause multiple equilibria and currency crises (Obstfeld, 1986). These models do not dismiss the possibility of self-fulfilling prophecies: investors target a certain currency simply because they anticipate that other investors will attack it. Flood and Marion (1997) contrast policy actions in first- and second-generation models: changes in policies in anticipation of a possible attack can lead to an attack and hence represent the trigger of crises in second generation models. Second-generation models are partially driven by events like the European Exchange Rate Mechanism (ERM, 1992), which aimed to reduce variability between currencies in Europe, before moving towards a unified currency: at that time, the UK withdrew from the treaty and the British Pound (GBP) ended up losing around 5% of its value (Claessens and Kose, 2014). The third-generation models investigate how sharp deteriorations in balance sheets, alongside variability in prices of assets (of which currencies and their exchange rates) lead to currency crises, resembling what occurred in the Asian crises in the 1990s.

In these countries, macroeconomic imbalances appeared to be manageable (surplus in fiscal positions and controllable current account deficit), but weaknesses in the financial and corporate industries were significant. The work of Chang and Velasco (2000) is mentioned as an example: the researchers show that large debt that is denominated in foreign currency and carried by local banks may lead to twin crises: a banking crisis and a currency crisis.

Furthermore, the third-generation models also consider the role of banks in the "self-fulfilling nature of crises": on one level, it is discussed that vulnerabilities resulting from overborrowing by banks can trigger currency crises (McKinnon and Pill, 1996, Krugman, 1999, and Corsetti, Pesenti and Roubini, 1998). On another level, Eichenbaum and Rebelo (2001, 2004) discuss that crises become self-fulfilling due to fiscal concerns and volatile real exchange rate fluctuations. Finally, Radelet and Sachs (1998) present self-fulfilling crises in a more general manner: they argue that when financial institutions face panics, they are forced to liquidate assets. This eventually becomes a confirmation of the panic and leads to a currency crisis. While the empirical research has not yielded conclusive evidence as to which of these models best characterizes currency crises, subsequent empirical work moved towards applying censored dependent variable models such as the logit model (Eichengreen et al., 1995, Frankel and Rose, 1996, and Kumar et al., 2003). Other methods employed by researchers like Kaminsky, Lizondo and Reinhart (1998) and Kaminsky and Reinhart (1999), are Signaling Models which evaluate the predictive power of certain variables in signaling the occurrence of a crises. Findings of these works show an association between certain indicators and crises, but the methods failed to accurately predict the timing of crises' occurrence.

b. Sudden Stop (Capital Account | Balance of Payments) Crises

Claessens & Kose define a sudden stop crisis as being a large drop in capital inflows to a country, accompanied with a significant increase in credit spreads. In fact, Claessens and Kose claim that models dealing with sudden stops resemble the third-generation models of currency crises in that they highlight balance sheet disparities arising from currency as well as from maturity, in financial and corporate sectors (Calvo et al., 2004). In addition, they assign greater importance to the role of international factors in instigating sudden stops in capital flows. Sudden stops typically occur in countries which have relatively little tradable sectors and a large base of foreign exchange liabilities. Such crises affect countries with dissimilar levels of GDP per capita, levels of financial development and exchange rate regimes and countries with different levels of reserve coverage. As per Calvo et al. (2008), sudden stops usually share two common elements: a small supply of tradable goods relative to domestic absorption and a domestic banking system with a large liabilities' base, denominated in foreign currency. In developing markets, the model incorporates current account reversals as well as the depreciation of the real exchange rate which typically occurs in times of crisis. In an attempt to better explain instances and match data more accurately, various friction variables were factored into more recent sudden stop models such as Fisherian channels, financial accelerator mechanisms, frictions in the labor market, in order to generate a drop in output during a sudden stop while maintaining the ability to account for movements of other variables. Literature on models with financial frictions reveals that such models reflect the dynamics of output and productivity more accurately in sudden stops.

Claessens and Kose provide an example of friction: when businesses are required to borrow in advance to settle working capital expenses, a drop in credit (sudden stop along with increasing external financing costs) reduces aggregate demand and consequently causes a reduction in output (Calvo and Reinhart, 2000). Similarly, in sudden stop crises, setting constraints on collateral in lending can cause debt-deflation, accelerated decline in credit, sharp drops in prices and volume of collateralized assets, resulting as such to a decline in output. Financial distress and bankruptcies compel banks to become more risk averse, reducing as such new lending which in turn causes a further decline in credit and consequently contributing to a recession (Calvo, 2000). Claessens and Kose explain that relatively small shocks can impose collateral constraints on debt, especially when borrowing is high compared to the value of assets. Sudden stops can be caused by Fisher's debt-deflation mechanisms through an amplified decline in asset prices and holdings of collateral assets (Fisher, 1933). Empirical research shows that many sudden stop crises have been linked with global shocks. Usually, a period during which large capital flows into economies triggers global shocks which eventually lead to a reversal of capital flows as sudden stops are more likely to occur when large cross border financial linkages are present (Claessens and Kose, 2014).

c. Debt Crises

Under the umbrella of Debt Crises, Claessens and Kose define Foreign Debt Crises and Domestic Debt Crises. The authors explain that a foreign debt crisis occurs when a country is unable or not willing to service its foreign debt, sovereign debt, private debt or all. As for the domestic public debt crisis occurs when a country does not respect its domestic fiscal obligations, either by explicitly defaulting, by deliberately inflating its currency (deliberate devaluation), or through other means of financial repressions. Theories on foreign debt crises are highly associated with theories on sovereign lending. In the absence and avoidance of military action, lenders are unable to seize collateral from another country when that borrowing country does not honor its liabilities. As such, economic reasons are indispensable to understand reasons behind international lending. Claessens and Kose present two notions that can impact a country's willingness to repay its debt: intertemporal and intratemporal sanctions. According to Eaton and Gersovitz (1980), intertemporal sanctions arise when a country faces the threat of no longer having access to credit if it defaults on its current obligations. In this type of models, the country defaults when the opportunity cost of not being able to borrow in the future is low which is likely the case when trade terms are favorable and expected to remain favorable (Kletzer and Wright, 2000). As for the intratemporal sanctions, Bulow and Rogoff (1999) explain that these arise when a country no longer has access to foreign currency because of sanctions imposed by their trading partners or make that country's access to international markets difficult. In such models, costs of trade cutoffs are low when trade terms are unfavorable (Aguiar and Gopinath, 2006). Nevertheless, models are unable to identify the reasons behind sovereign defaults and the reasons behind the significant amounts of lending.

Also, models seem to underestimate the probabilities of default. Unlike what most models predict, defaults do not occur during bad times. Furthermore, models underestimate the exposures that lenders and investors are willing to have to countries with high default risk. Panizza et al. (2009) find that changes in institutional environment does not seem to impact the relationship between economic and political variables and the probability of defaulting on debt which confirms that models are still unable to accurately capture fully the features that are necessary in understanding and explaining defaults. Theories and literature on domestic debt crises have been limited until recent times. In fact, Claessens and Kose (2014) explain that economic theory and models assume that governments will always respect their domestic debt engagements and so the role of domestic debt crises has been insignificant in economic theory. Yet, Reinhart & Rogoff (2009b) show that a mere few countries were able to avoid domestic debt default at the expense of severe unfavorable economic repercussions. The authors explain that defaulting locally often occurs when inflation is unusually high due to currency issuance. It is often found that debt defaults in supernormal inflation times is usually followed by currency crises. Reinhart et al. (2011) find another way of defaulting: through financial repression. Actually, following a devaluation or "debasing" of a local currency, it often takes a long time to reinstate the trust of the public in that currency and therefore fiscal costs of stabilizing inflation increase considerably, leading to significant negative real effects of the abnormal inflation and accompanying currency crisis. In developing countries, debt intolerance appears to be associated with duress that would otherwise be controllable in advanced economies. Normally, safe debt thresholds depend on factors specific to the country, such as its default history and inflation levels.

Reinhart and Rogoff (2009a) find that when the external debt level of a country exceeds 30-35% of its Gross National Product (GNP), the probability of occurrence of an external debt crisis increases significantly. These authors also find that, when a country defaults recurrently, its debt intolerance increases and its access to global capital markets becomes a challenge. Claessens and Kose discuss the many challenges that are still present regarding modeling countries' abilities to sustain and service domestic and global debt obligations. One important challenge presented is that the form of financing used by countries is endogenous. Jeanne (2003) claims that short-term debt denominated in foreign currency can be advantageous to a country in driving the implementation of good macroeconomic policies. Another claim is made by Diamond and Rajan (2001) whereby the experts hypothesize that financial institutions in emerging markets often operate in lower quality institutional environments and so they become required to hold short term debt to finance illiquid projects. Finally, Eichengreen and Hausmann (1999) assume that there is an "original sin" in the motivation behind borrowing. They explain that countries with unfavorable conditions have little to no choice but to depend on short term debt denominated in foreign currency as a main source of capital, in the absence of foreign direct investments and no access to equity. In general, the deeper causes of debt crises are not to be separated from the immediate or proximate causes.

In fact, much of the weaknesses and vulnerabilities that increase the riskiness of a debt crisis is often rooted in financial integration, political economics and institutional environments. McKinnon and Pill (1996, 1998) find that countries with weak morals and poor supervision tend to be more vulnerable to banking and currency crises when capital inflows are unrestricted.

Furthermore, the authors highlight that debt crises are often accompanied by sudden stop, currency, and banking crises (or different combinations of the different crises) and so empirical studies in identifying the deeper causes of crises face the challenges of endogeneity and simultaneity.

d. Banking Crises

The global financial crisis that occurred in the last decade inspired research on the significance of a stable financial system for macroeconomic stability (Beutel et al., 2019). Banking crises, although being relatively common, are the least comprehended. Some experts attribute the occurrence of banking crises to bank runs while others argue that systemic banking crises can be a consequence of micro and macro level policies (Laeven, 2011). On one hand Claessens and Kose (2014) explain that systemic banking crises occur when actual or impending bank failures and runs prompt banks to suspend convertibility of liabilities (limitations on deposit withdrawals) and require government intervention to extend banks with liquidity and provide them with capital assistance on a large scale. In fact, Claessens and Kose explain that banks are intrinsically fragile and subject to runs by depositors. Furthermore, a delinquency in one bank can rapidly spread across the entire system.

In addition, a weak institutional environment can increase crises' risks as financial institutions rely heavily on the legal and judicial environments and on informational systems in their investment decision making and debt collection processes. Claessens and Kose discuss that bank runs have taken place in many countries across the world and throughout history.

Runs were common during the Great Depression (1900s) and are common in emerging and developing markets. Runs can also occur in non-bank financial markets (like the mutual funds case of the US in fall of 2008). As previously mentioned, banks are inherently fragile and highly leveraged entities with an important role to transform maturities and create liquidity. A significant challenge caused by fragility in financial markets is coordination, or the lack of it. Problems rooted in coordination arise when investors or entities proceed to withdrawing capital or liquidity out of fear that others will proceed to do the same. Therefore, small shocks can transform into market turmoil and even into a financial crisis. A bank run is an example of a coordination problem: banks typically borrow over the short term (deposits have short term maturities) and invest / lend over the long term. This makes banks vulnerable and weak to sudden increased demand for liquidity by a large number of customers wishing to withdraw their deposits out of the fear that the bank will become insolvent. As a bank run unfolds, it gains momentum and turns into a self-fulfilling prophecy: the more people withdraw deposits, the more probabilities of defaults increase, leading to further withdrawals. This cycle can hinder a bank's stability to the extent that it faces bankruptcy. On the other hand, and according to Laeven (2011), vulnerabilities in a banking system can be the result of policies at both micro and macro levels.

Dewatripont and Tirole (1994) explain that policymakers, markets, and institutions have identified the banking system's weaknesses and fragilities and have developed means of response to them. Good market practice recommends limiting vulnerabilities whereas at the financial institutions' level, firms have adopted risk-management strategies to hedge against these fragilities. In addition, micro-prudential regulation with rule-enforcing supervision aims to reduce the risk behavior of certain individual financial institutions and maintain stability.

Furthermore, the lender-of-last-resort: the Central Bank can assist banks by providing them with short term liquidity during periods of high stress while policy intervention by the public sector can mitigate systemic risk when financial turmoil arises. Although regulation and support measures can help, their poor design and inadequate implementation could actually increase the likelihood of a banking crisis occurrence.

Often, regulation and supervision seem to constantly lag behind innovation, leading to a certain degree of poor design and implementation of safety net and regulation measures. Finally, moral hazard can lead banks to assume excessive leverage: when an institution believes it is too big to fail, it can take too many risks and consequently creating some systemic vulnerabilities. In the second chapter of the IMF's book – written by Laeven and Valencia – the authors investigate systemic banking crises and their initial conditions. They also discuss crises containment policies as well as crises resolution policies. To investigate the initial conditions, the authors study certain macroeconomic variables as well as information on the banking system.

The researchers collected data on the following macroeconomic variables: Fiscal Balance (government balance) / GDP, Public Debt / GDP, Inflation, Central Banks' Net Foreign Assets, Central Banks' Net Foreign Assets / Broad Money (M2), Total Deposits at Deposits receiving institutions / GDP, GDP Growth and Current Account / GDP. The researchers also retrieved bank specific data: Peak NPL: Peak Ratio of Non-Performing Loans to Total Loans (%) during a 5-year period (t to t+5), Government Ownership: percentage of banking system assets that are owned by the government at t-1, Significant Bank Runs: whether the country's banking sector experienced a bank run by depositors, estimated at being a one-month percentage drop in total outstanding deposits exceeding 5% during t to t+1, and Credit Boom: defined as three-year pre-crisis average growth in credit to the private sector, exceeding 10% per year, computed throughout t-4 to t-1.

D. Causes of Banking Crises

The drivers of banking crises have long been an issue of debate. Some theories attribute banking crises to depositors' fears causing run-on-banks that can be illustrated by unjustified withdrawals. These run-on-banks exert excessive pressure on banks' liquidity (Friedman and Schwartz, 1963), rendering them consequently insolvent. If no action is immediately taken and policy swiftly implemented, bank failures can become systemic. Bank runs are not necessarily directly related to changes in the economy. Some bank runs are deemed as self-fulfilling prophecies mainly when long-term investments are illiquid (Bryant, 1980 and Diamond and Dybvig, 1983). Bank runs may also be caused by withdrawal of funds when economic declines are expected.

The researchers Ben Bernanke, Douglas Diamond and Philip Dybvig have been awarded the 2022 Nobel Prize in Economics for their contributions to the field of econometrics, specifically their work on the importance of preventing widespread bank collapses. The laureates in Economic Sciences, Ben Bernanke, Douglas Diamond, and Philip Dybvig, have demonstrated the importance of preventing widespread bank collapses to manage financial crises.

They have explained the central role of banks in society, why they are vulnerable to collapse, and how society can lessen this vulnerability through modern bank regulation. Bernanke's research shows how failing banks played a decisive role in the Great Depression, and the importance of well-functioning bank regulation in preventing catastrophic consequences during the 2008-2009 financial crisis. The article discusses how Ben Bernanke's work on the Great Depression of the 1930s changed the conventional wisdom about its causes. Prior to Bernanke's research, experts believed that printing more money could have prevented the depression. However, Bernanke showed that the main cause of the depression was the decline in the banking system's ability to channel savings into productive investments. This decline was caused by a banking crisis that began in 1930, leading to bank runs and ultimately bankruptcies. Bernanke established that bank collapses were decisive for the recession developing into a deep and prolonged depression. The relationship between a bank and its borrowers contains knowledge capital that is necessary for efficient lending, and building up such knowledge capital takes a long time. Repairing a failed banking system can therefore take many years, during which time the economy functions very poorly. Bernanke demonstrated that the economy did not start to recover until the state finally implemented powerful measures to prevent additional bank panics. Banks are important because they receive money from deposits and channel it to borrowers, which creates financial intermediation. Banks solve conflicts between the needs of savers and investors. They create liquidity for savers while borrowers can access long-term financing. Diamond and Dybvig's theoretical model explains how banks create liquidity and why they work, but it also highlights that the banking system is inherently vulnerable, leading to a need for regulation.

Banks create money through maturity transformation, which is when the bank transforms assets with long maturity into bank accounts with short maturity. However, banks are vulnerable to rumors that cause bank runs, and deposit insurance from the government is a solution to this problem. Banks also monitor borrowers to ensure that they honor their commitments.

a. Bank Runs

Runs reduce the value of bank assets and consequently increase the probability of banks failing to service their debt towards depositors (Jacklin and Bhattacharya, 1988). Such crises are more likely when information about bank distress is asymmetric (Laeven, 2011). In 2005, Diamond and Rajan build on their 2001 model to show bank runs can happen even if depositors are not running because of their fear that other depositors will run as well. Depositor panics are most detrimental when they cause liquidity pressures to spread across the banking system.

So, failure of a certain bank creates a domino effect on the entire system. But depositors' panic-based theories have become less frequent since deposits have become insured.

Figure 2: Illustration of a run on a bank (source: Nobel Prize Org.)

b. <u>Systemic Vulnerabilities</u>

Other theories attribute banking crises to heavy losses banks' asset-side on the balance sheets, consequently rendering banks insolvent. Losses generally stem from adverse macroeconomic distresses, market failures, government interference, or fraudulent acts. Many of these theories are centered around changes in economic fundamentals and regard banking crises as a natural effect of business cycles (Minsky 1982, Gorton 1988). Banking crises that are caused by macroeconomic fundamentals are basically caused by unsustainable macroeconomic policies, misalignments in exchange rates and globalization of financial conditions. Some macroeconomic policies include monetary and fiscal policies that induce growth and cause increasing credit volumes, accumulation of debt, and concentration of investments in certain asset types.

These unsustainable macroeconomic policies ultimately cause a deterioration of bank assets, asset price bubbles to burst and eventually leads to banking crises (Reinhart & Rogoff, 2009). Laeven (2011) discusses that macroeconomic shocks can cause exceptionally severe bank distress in emerging markets. In fact, banks in developing countries tend to borrow from foreign markets. So, credit risk derived from currency risk or maturity gaps in balance sheets can swiftly turn into direct losses for banks in these emerging markets: the depreciation of a certain currency or increases in global interest rates can have detrimental effects on banks. For example, for exporters, currency depreciation may cause significant changes in trade terms which will consequently hinder these companies' ability to pay back their debt towards banks.

c. Unsustainable Macroeconomic Policies

Unsustainable macroeconomic policies can have significant effects on the banking and financial sectors, often leading to crises. Policymakers need to be aware of the risks associated with these policies and take steps to mitigate them to ensure economic stability and prevent future crises. In fact, policies, such as fiscal and monetary policy, can impact the economy in a variety of ways, including through their effects on inflation, exchange rates, and asset prices. When these policies are unsustainable, they can lead to imbalances and distortions in the economy, which can eventually result in a crisis. One common effect of unsustainable macroeconomic policies is the buildup of debt in the economy. When policymakers engage in deficit spending or other expansionary fiscal policies, they can stimulate economic growth in the short term but can also lead to an increase in government debt levels. This, in turn, can lead to higher interest rates and reduced investment, as investors become more hesitant to lend to the government. In addition, unsustainable monetary policy can lead to inflation and exchange rate instability, which can have significant impacts on the financial sector. For example, high inflation can erode the value of banks' assets and reduce their ability to lend, while exchange rate volatility can increase the risks faced by banks and other financial institutions. Another effect of unsustainable macroeconomic policies is the buildup of asset price bubbles. When interest rates are kept artificially low or when policymakers engage in other policies that encourage excessive risk-taking, asset prices can become inflated beyond their true values. This can lead to a bubble, which, when it bursts, can result in significant losses for investors and destabilize the financial sector.

In the article "Predictable Financial Crises," Robin Greenwood and co-authors argue that financial crises can be predicted based on certain economic indicators. Specifically, the article focuses on the relationship between credit growth, asset prices, and economic fundamentals. Greenwood et al. (2022) suggest that financial crises occur when credit growth outpaces economic fundamentals, leading to overvaluation of assets and an eventual market correction. He cites several historical examples, such as the housing market bubble in the United States in the mid-2000s, to support this argument. The authors also discuss the role of financial intermediaries, such as banks and other financial institutions, in exacerbating these imbalances. He notes that these intermediaries often engage in "financial innovation" that allows them to increase leverage and take on more risk, which can further amplify the effects of a financial crisis.

To mitigate the risk of financial crises, the authors suggest that policymakers should monitor credit growth and asset prices closely and intervene when necessary to prevent imbalances from becoming too severe. He also advocates for greater regulation of financial intermediaries to limit their ability to take on excessive risk.

Overall, the article argues that financial crises are predictable and can be prevented with the right policies in place. (Greenwood et al., 2022).

d. Credit Growth

It has been observed that credit growth may lead to financial crises and eventually severe macroeconomic consequences. When the economy is witnessing growth, the overall sentiment about the future is optimistic. As a result, credit volumes grow as lending requirements become more relaxed. But when the economic environment returns to its initial state, the sub-par lending quality causes a burst of the credit bubble (Dymski, 2011). Investigating further, Sufi and Taylor (2021) assess whether the source of credit is a determinant of banking crises and whether it is relevant if credit is financed domestically or from foreign creditors (Sufi and Taylor, 2021). Jorda, Schularick, and Taylor (2011) explore the current account deficit of 14 developed economies over the period extending from 1870 to 2008 and study whether a deficit in the current account can help predict financial crises. Findings of their research suggest that credit growth is the best predictor of systemic banking crises. However, they also find that the predictive power of credit growth improves when external variables are accounted for. Jorda, Schularick, and Taylor (2011) find that borrowing from abroad increases the likelihood of the occurrence of a financial crisis vs. borrowing from local sources. Mian, Sufi, and Verner (2017) investigate this question in more recent times:

they study a sample of 30 developed economies over the period extending from 1960 to 2012. They focus on factors that impact GDP growth and find that an accumulated current account deficit – as a standalone variable – does not justify lower economic growth as a consequence.

Conversely, an increase in household debt to GDP has a significant power in predicting lower future growth. Yet, the authors find that the interaction between increasing household debt and the accumulated current account deficit is a powerful predictor. An important finding is related to the source of debt: rising household debt that is financed by foreign sources forecasts lower GDP growth as compared to debt that is financed locally.

In the paper entitled "Finance and Business Cycles: The Credit-Driven Household Demand Channel," Atif Mian and Amir Sufi examine the role of household debt in driving business cycles. Specifically, they focus on the relationship between household debt and consumption, and how this relationship can impact the broader economy. The authors argue that changes in household debt can have a significant impact on consumption, as households often use debt to finance their spending. They note that during periods of easy credit, households tend to take on more debt and increase their consumption, which can lead to a boom in the economy. Conversely, during periods of tight credit, households tend to reduce their debt and consumption, which can lead to a contraction in the economy. Mian and Sufi (2018) also examine the impact of household debt on the housing market. They note that during periods of easy credit, households tend to take on more debt to purchase homes, which can lead to a housing market bubble. They argue that when this bubble bursts, it can have significant negative effects on the broader economy, as was seen during the 2008 financial crisis.

To mitigate the risks associated with household debt, Mian and Sufi suggest several policy solutions. These include greater regulation of the financial sector to prevent excessive lending, as well as policies to help households manage their debt burdens and reduce their reliance on debt to finance consumption.

Overall, the article highlights the importance of household debt in driving business cycles and the need for policymakers to carefully manage this relationship to ensure economic stability (Mian and Sufi, 2018).

So, banking crises may be caused by macroeconomic factors, systemic problems in the industry, market sentiment (optimism & lending, pessimism & run-on-banks), currency and a reactive policy-making system rather than a proactive one. As financial crises in emerging markets represent a great problem for public policy, intervention and speed of action are crucial. The gravity of uncertainty in emerging markets, and more specifically from an economic and political perspective, makes hedging against crises and proactive policy-making more difficult (Dooley, 2003).

E. Examining Crises Quantitatively

Numerous studies have been conducted around investigating banking and financial crises quantitatively. In fact, Demirgüç-Kunt and Detragiache (2005) employ two econometric approaches to identify the determinants of banking crises: the signals approach and the multivariate logit approach. They study 77 crises in 94 countries over the period extending from 1980 and until 2002.

Findings yielded several significant determinants of banking crises such as individual bank measures of fragility, financial liberalization, international shocks, exchange rate regime, bank ownership and structure, and finally the role of institutions and political systems (Demirgüç-Kunt and Detragiache, 2005).

Another study, conducted by Sufi and Taylor (2021), surveys existing literature on financial crises and studies how crises are measured. Sufi and Taylor (2021) also investigate the association between financial crises and economic downturns and question whether crises are predictable. The first part of the study is dedicated to defining a crisis by measuring how GDP is affected “h” periods after a crisis occurs. They find that GDP falls following a crisis. Then, they define a financial crisis as being a crisis in the banking sector by combining data and narrative criteria (Reinhart and Rogoff, 2009b) and Schularick and Taylor (2012). Schularick and Taylor (2012) categorize recessions as normal and financial types and classifies them based on the proximity of the recession to the start of a financial crisis (Jorda, Schularick, and Taylor, 2013). Sufi and Taylor (2021) also define crises through influential binary classification. They present two different sets of data developed by subject matter experts: the first includes episodes dating back to the 19th century (referred to as the Long Dataset) whereas the other includes data from the year 1970 onwards (referred to as the Recent Dataset) in advanced and developing economies. To detect consistency in the classification of financial crisis events using the most widely used methods to date, they only focus on instances where the two datasets overlap, although the overlap is limited. On one hand, it is found that, in the Long Dataset, differences in classifications in developed economies are not significant. Yet, the subjective element of classifications is uncovered by the data. On the other hand, findings of the Recent Dataset show that significant progress has been made in the systematic classification of financial crisis occurrences.

The databases developed include many countries and widely cover the recent decades. Evidently, experts sometimes disagree on some specific historical events given the subjectivity involved, but generally agreement on the same criteria results in substantial.

Another approach to classify crises uses data driven criteria (Sufi and Taylor, 2021). This approach allows detection of stress in the financial system based on tangible financial data. In theory, the results of this approach could be completely dissociated from the findings of the qualitative or narrative information. However, it is considered as valuable information when qualitative data displays bias.

Sufi and Taylor's (2021) survey mainly provides a comprehensive overview of financial crises, with a focus on the impact of the pandemic. The authors highlight the ongoing challenges facing policymakers and call for continued vigilance to prevent future crises. The authors' 2021 survey builds on their earlier work on financial crises, providing an updated overview of the causes and consequences of financial crises. The authors begin by defining financial crises and discussing the various types of crises that have occurred throughout history. They then examine the various theories of financial crises, including the "Minsky moment" theory, which argues that crises occur when financial institutions and investors become overly confident and take on excessive risk. They also discuss the role of macroeconomic imbalances, such as trade imbalances and fiscal deficits, in causing crises.

The authors also explore the consequences of financial crises, including the impact on economic growth, inequality, and political stability. They argue that financial crises can have long-lasting effects on the economy and society, and that mitigating these effects requires careful policy interventions.

In their 2021 survey, Sufi and Taylor focus on the impact of the COVID-19 pandemic on the global economy and financial system. They note that the pandemic has led to a sharp increase in government debt levels and a rise in nonperforming loans, among other challenges. The authors also discuss the policy responses to the pandemic, including the massive fiscal and monetary policy interventions that have been implemented around the world. They note that these policies have helped to stabilize the economy and financial system in the short term but warn of the longer-term risks associated with rising debt levels (Sufi and Taylor, 2021).

Drehmann and Juselius (2014) and Drehmann, Juselius, and Korinek (2018) explore data containing information about new debt and debt service payments for a panel of 16 countries over a period of 35 years (from 1980 to 2015). They concentrate on the debt service ratio (DSR = Ratio of (interest + principal payments made by borrowers) scaled by the disposable personal income) and find that the increase in private credit is a strong predictor of a financial crisis. However, they also consider the rise in the Debt Service Ratio as a factor that could explain – on the short-term – how the crisis is actually triggered. This literature finds that worsening financial conditions are strong predictors of the unfavorable economic growth. Also, it is evident that, despite favorable financial conditions on the short run, negative consequences are reached on longer runs, especially if credit growth increases (Adrian, Grinberg, Liang, and Malik, 2018). Therefore, the general result found is that the critical tipping point leading to financial crises is an adverse shock to the financial sector, occurring following a period of financial growth. Finally, Babecký et al. (2014) explore 40 crises episodes in developed countries that occurred between the years 1970 and 2010.

They begin by presenting facts on banking, debt and currency crises and they find that banking and credit crises are correlated and usually precede currency crises but not the other way around. Findings also reveal that banking crises are the costliest and recovery happens over six years, on average. Then, Babecký et al. (2014) try to identify early warning indicators of crises in developed economies and for that they employ Bayesian model averaging. Results of their investigation shows that the most reliable signal across various specifications and time horizons is the substantial increase in domestic credit that occurs before banking crises. Other robust signals found are increasing money market rates and international corporate spreads and require close and continuous supervision. To examine currency crises, Babecký et al. (2014) also investigate the role of increasing domestic credit and money market rates in detecting local currency overvaluation. They also find that indicators are most effective over time horizons that are shorter than two years. It was difficult for the authors to determine early warning indicators of debt crises because of the low occurrence of such episodes in their dataset. Lastly, they employ a signals approach to derive a threshold value for the best performing indicator (which was determined to be domestic private credit) and develop a composite early warning index that increases the effectiveness of the model even further.

F. Quality of Institutions and Creditor Rights

As previously mentioned, in the second chapter of the IMF's book – written by Laeven and Valencia, the authors discuss crises containment policies as well as crises resolution policies. Regarding crisis containment policies, the authors discuss several mechanisms. In fact, a government's initial policy options are limited to policies that do not require complexity. Timely policy responses include suspension of convertibility of deposits (limiting withdrawals of deposits), regulatory capital forbearance: allowing banks to continue to operate while their capital has been fully depleted (a grace period is granted to banks to regularize their situation by the regulator), emergency liquidity support to banks and government guarantee of depositors. According to the authors, the appropriate policy to be implemented depends on the trigger of the crisis: loss of depositor confidence, bank insolvency, spillover, currency and / or macroeconomic pressures. Once response measures have been implemented, the government must face the long-term challenges of crisis resolution on the level of the economy, financial institutions, and private sector (individual and corporations). The main policy approaches used in the resolution phase are the following (Laeven and Valencia, 2014): conditional government-subsidized but decentralized workouts of distressed loans, debt forgiveness, establishment of government-owned asset management company to buy and resolve distressed loans and government-assisted recapitalization of the financial institutions through capital injection.

To assess institutional development, Laeven and Valencia examine Creditors' Rights using an index of protection of creditors' rights from Djankov, McLiesh and Shleifer (2007).

In their paper, Djankov et al (2007) examine the role of institutions in the resolution of systemic banking crises for a sample of over 40 countries. In fact, they explain that banking crises are costly on a fiscal level. According to the authors, an occurrence of a systemic banking crisis involves intricate coordination issues. The many players include, but are not limited to, corporations, financial institutions, the general economic outlook, financial systems, and governments. The latter often blazes a trail in the restructuring of the banking system. During the restructuring process, governments incur large fiscal costs. The researchers discuss that institutional weaknesses typically aggravates a crisis, which in turns becomes a complicating factor: disclosures and accounting rules for financial institutions and the private sector may be fragile, equity and creditor rights may be inadequately defined or poorly enforced, and the judiciary system is often found to be inefficient. As a consequence, the government itself may face credibility problems as it may be held accountable for the crisis, faces consistency problems, and corruption may be significantly large. The work's findings show that better institutions, represented by reduced corruption, improved laws & regulations, and an efficient legal system and bureaucracy support the accelerated resolution of systemic banking crises.

The authors find that these results are robust to estimation techniques, and that these results suggest that countries should enforce strict policies to resolve crises and even take advantage of the crisis' occurrence to implement structural reforms on the medium term, which will in turn be beneficial in the face of future systemic crises. Kaufman, Kraay, and Zoido-Lobaton (1999) define Quality of Institutions as an index. The index and captures six dimensions of institutional quality: democracy, political instability, rule of law, bureaucratic regulation,

government effectiveness, and corruption. The higher the index value, the better is the quality of institutions (Galindo and Micco, 2007) examine the relationship between creditor protection and credit responses to macroeconomic distresses. Their database comprises of 79 countries with variables collected over a period of 14 years (between the years 1990 and 2004) and includes legal elements of finance on aggregate credit growth. Findings of their work show that credit is more susceptible to external shocks in countries with weak legal creditor protection and weak enforcement.
In his paper, Hatem ElFeituri provides insights into the complex relationships between banking stability, institutional quality, market concentration, competition, and political conflict in the MENA region. The author argues that policymakers should focus on strengthening institutions, promoting competition, and reducing political conflict in order to promote banking stability and support economic growth and development (ElFeituri, 2022). In fact, the author investigates the relationship between banking stability, institutional quality, market concentration, competition, and political conflict in the Middle East and North Africa (MENA) region.

ElFeituri argues that banking stability is crucial for economic growth and development, and that it is influenced by a variety of factors. First, the author examines the role of institutional quality in promoting banking stability. He finds that countries with strong institutions, such as those that enforce property rights and contracts, tend to have more stable banking systems. In contrast, countries with weak institutions, such as those that are prone to corruption and political conflict, tend to have less stable banking systems. Second, the author explores the impact of market concentration on banking stability.

He argues that high levels of market concentration can lead to lower levels of competition, which can in turn lead to greater risk-taking by banks and lower levels of stability. However, he also notes that moderate levels of market concentration can be beneficial for banking stability, as they can promote economies of scale and reduce the costs of regulation. Finally, the author examines the role of political conflict in influencing banking stability. He argues that political conflict, such as civil wars and insurgencies, can have a negative impact on banking stability by disrupting economic activity, reducing investment, and undermining the credibility of government institutions (ElFeituri, 2022).

G. Machine Learning on Predicting Banking Crisis

a. Machine Learning Overview

Machine learning can be defined as sets of algorithms and statistical models that find patterns in data, learn from the data, and attempt to make predictions using new or unseen data (Langley, 1996).

Machine learning describes the systems' capacity to learn from problem-specific data and automate the analytical model building process. Machine learning has countless applications and across numerous industries. Machine learning models can be categorized as such (Janiesch, Zschech, & Heinrich 2021):

Machine learning is said to be supervised when the different classes of the target variable are known, and the goal is to study many labeled examples to be able to make predictions about future unknown data points. Within supervised learning, there are different learning algorithms:

Distance-based learning such as K-Nearest Neighbors (KNNs) and Support Vector Machines (SVMs), Probabilistic Classification Models such as the Naïve Bayes Model (multinomial and Gaussian) and Tree-Based Models such as Classification trees, Decision trees, Random Forests, and Gradient Boosting Trees.

Machine learning is said to be unsupervised when labels are unknown, and the machine's goal is to organize data in a way such that it can derive a certain structure or pattern. Unsupervised learning aims to group data into clusters or tries to find different ways of making complex data appear simple.

Reinforcement learning where the algorithm chooses the appropriate action or decision to each data point, and this is commonly applied in robotics as reinforcement learning fits naturally with Internet of Things applications. The learning algorithm receives a reward signal after a decision is made as an indicator of the decision's performance and based on the signal received, the algorithm modifies the strategy to achieve the highest reward.

Janiesch, Zschech, and Heinrich (2021) illustrate machine learning concepts in the blow figure.

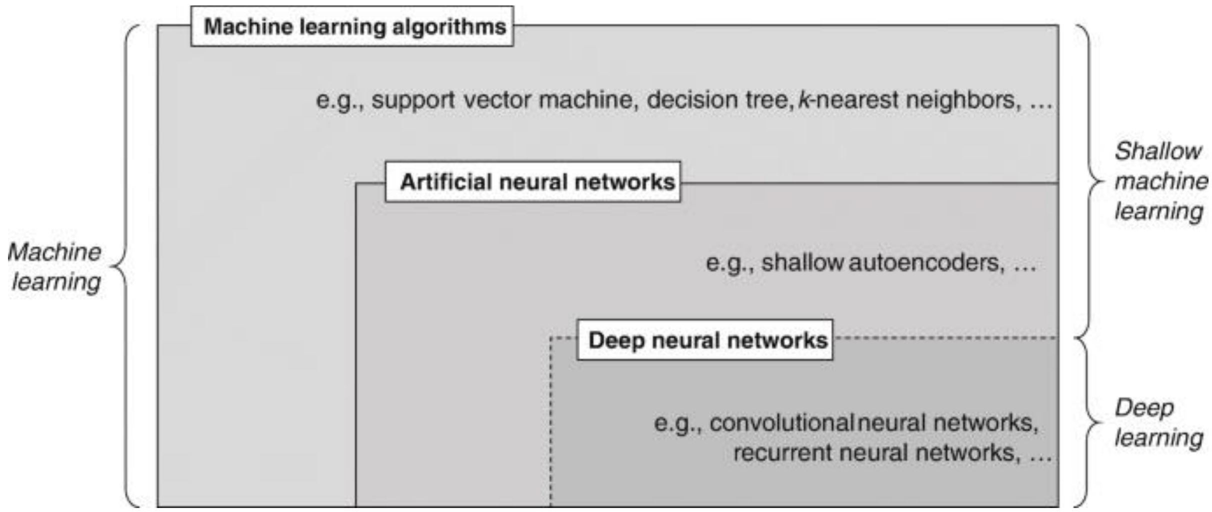


Figure 3: Diagram of Machine Learning Concepts (ref: Goodfellow et al. 2016)

b. <u>Machine Learning Research in Predicting Banking Crises</u>

Research has already been conducted to predict Banking Crises. In fact, Beutel et al. (2019) question in their research whether Machine Learning can predict banking crises. In this study, a sample of systemic banking crises that occurred in developed economies over a timeframe of around 50 years was examined (panel dataset). The researchers then estimated the probabilities of crises occurring using logistic regression, K-Nearest Neighbors (KNNs), Decision Trees, Random Forests, Support Vector Machines (SVMs) and Neural Networks[1].

[1] Beutel et al. discuss that binary models (like logit) are commonly used in literature that studies early warning tools. Yet, they find that Machine Learning methods seem to allow for nonlinearities and appear to be more flexible in terms of distributional assumptions, which may be useful when forecasting extreme events such as systemic banking crises.

Since Logit approach is common in studies of early warning indicators, Beutel et al. use it as a benchmark to compare the predictive performance of the different machine learning models employed. Findings of the research showed that machine learning models did not outperform the benchmark in predicting crises when the models were applied on new or unseen data. Therefore, it is observed that the Logit approach seems to efficiently use the available information and would have been able to predict the occurrence of the 2008 global financial crisis based on data retrieved from many countries.

The models in this research identify they key predictive variables of banking crises to be credit expansion, asset price booms and external imbalances which validates the theories which claim that banking crises are rooted
in macroeconomic fundamentals rather than depositor fears.

Discussing macroeconomic fundamentals, Drehmann and Juselius (2014) argue that early warning indicators of banking crises should be evaluated on their performance vis-à-vis policy makers' problem. In their research, they assess costs and benefits of different policy measures, timing, and stability requirements of the early warning indicators and translate them into statistical evaluation criteria.

When the statistical criteria were applied to potential warning indicators, it was uncovered that the credit-to-GDP gap and the debt service ratio consistently outperform other measures. The credit-to-GDP gap has the best performance on the long term whereas the debt service ratio performs better on the short term (Drehmann and Juselius, 2014).

Expanding on their work, Drehmann et al. (2018) find that household and international debt represent sources of vulnerabilities that could lead to banking crises.

H. Overview of the Lebanese Financial, Banking & Currency Crisis

Since October of 2019, Lebanon has been witnessing several simultaneous crises: an economic crisis, a banking and financial crisis, a currency crisis, a healthcare crisis, and finally a devastating explosion at the Port of Beirut on the 4th of August of 2020. While all crises are significant and have detrimental effects on societies individually, the compounded and simultaneous effect has been catastrophic.

The World Bank ranked the Lebanese financial crisis among the top three most severe crises since the mid-19th century[2]. The World Bank claims that the 37% drop in GDP per capita is usually witnessed in times of conflict or even war. Focusing on the banking sector in Lebanon, October 17th of 2019 represented a schismatic date dividing banking practices before and after that date.

[2] Spring 2021 Lebanon Economic Monitor

In fact, and over the last few decades, the Lebanese banking sector was funding the Lebanese economy by financing its sovereign debt. The sector's heavily concentrated and increasing investment in the state was funded by deposits.

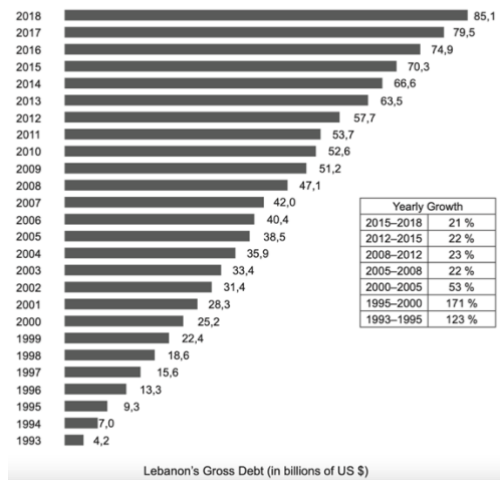


Figure 4: Lebanon's Gross Debt (in billions of US Dollars), (source:[12])

The informal capital controls triggered a state of fear among depositors as limits were being set and continuously changed on cash withdrawals in foreign and local currencies, limitations were enforced on transfers abroad. On the credit side, loan requests and disbursements were halted – in Retail banking and Corporate & Small to Medium Enterprises (SME) banking which had detrimental impacts on the economy.

In addition, the sharp devaluation of the Lebanese Lira significantly affected business owners on an operational level and on the financial level. Consequently, businesses were only able to function if cash was available and as a result, many firms downsized, limited operations, moved operations abroad, and took other measures to overcome the hardships.

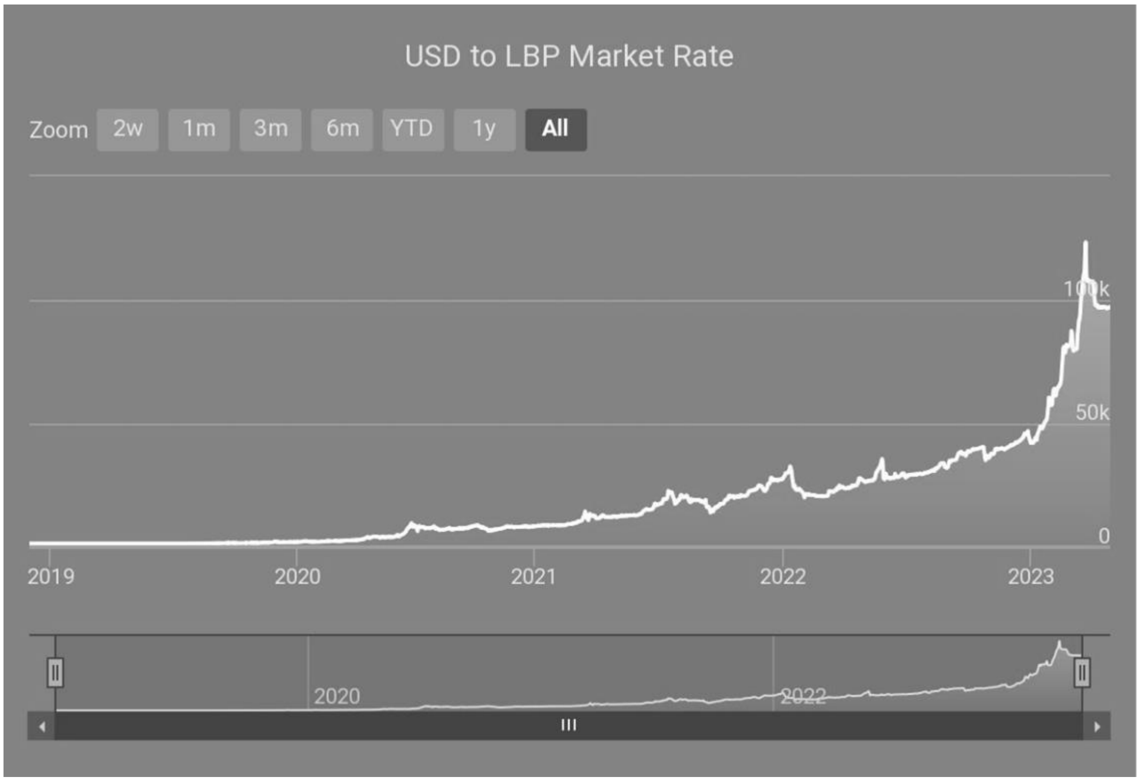


Figure 5: Black Market Exchange Rate since July, 2019 and Q1 of 2023 (source: lirarate.org)

Kaminsky & Reinhart (1999) study the twinning of crises during the time where turmoil hit the Mexican and Asian markets. They investigate whether there are potential associations between banking and currency crises.

Their findings curtailed that a banking crisis usually heads a currency crisis, and the resulting currency crisis seems to worsen the depth of the banking crisis, creating as such a vicious cycle. These patterns coincide with an economic recession, following growth boosted by credit expansion, capital inflows, and an overvalued currency.

The work of Baz, Cathcart and Michaelides (2019) provides a comprehensive analysis of the economic crisis in Lebanon and proposes a range of policy solutions to address the country's challenges. The authors highlight the urgent need for action to prevent further deterioration of the economy and restore confidence in the financial system. In fact, the authors examine the economic crisis in Lebanon and proposes potential solutions to address the country's challenges. The authors begin by providing a brief overview of Lebanon's economic history and the factors that have contributed to its current crisis, including political instability, corruption, and unsustainable economic policies. They note that the crisis has been exacerbated by the COVID-19 pandemic and the explosion at the Port of Beirut in August 2020. The authors then propose a range of policy measures to address the crisis, including restructuring the banking sector, implementing fiscal reforms, and promoting private sector growth. They argue that restructuring the banking sector is essential to restoring confidence in the financial system and rebuilding the economy.

The authors also propose introducing a new currency, which they call the "Lollar," (a term which was initially introduced by banking professional Dan Azzi[3]) to replace the Lebanese pound. They argue that the Lollar would help to address the country's currency crisis and provide a stable medium of exchange for businesses and consumers. Finally, the authors discuss the potential role of the international community in supporting Lebanon's economic recovery. They suggest that international organizations and donor countries could provide financial assistance and technical expertise to help implement the proposed reforms.

Also, in the paper “A Primer on the Financial Crisis in Lebanon: A Historical and Cross-Country Perspective”, Nada Mora delves into the financial crisis in Lebanon. Mora (2020) explains that the crisis only became visible when dollar liquidity became scarce in the summer of 2019 but deteriorated with the political events that took place in October of 2019. Mora (2020) argues that despite the fact that the visibility of the crisis is only recent, it has been actually developing over the past decades. Having commonalities with previous crises in other countries, Mora (2020) explains that the balance sheets, both in terms of currency and maturity, of the government, the banking sector, the central bank of Lebanon and the private sector, were in mismatch. Another similarity with past crises that was highlighted by Mora (2020) is the fixed or pegged exchange rate, which makes it susceptible to speculation, resonating with 2nd and 3rd generation crises (discussed earlier in this paper).

[3] Lollar by Dan Azzi: https://danazzi.com/

Mora (2020) discusses further that Lebanon’s large depositor base (local, regional and expatriate) and its loyalty has resulted in imbalances in the balance sheets to build up and makes the recovery more complex since a halt in domestic depositor funding spreads rapidly across all balance sheets, making it a major contributor to the systemic liquidity freeze and to the adverse impact on the economy.

The Yale School of Management blog post titled "Crisis in Lebanon: Economic Free-Fall, IMF Negotiations, and Beirut Explosion" discusses the ongoing economic crisis in Lebanon, which has been exacerbated by the COVID-19 pandemic and the August 2020 explosion in Beirut. The post is the first part of a series on the crisis and focuses on the economic factors that led to the current situation. The post explains that Lebanon's economy has been in decline for years, and the country was already struggling with high levels of debt, corruption, and political instability before the pandemic hit. The pandemic caused a sharp contraction in the economy, and the subsequent lockdowns and restrictions further exacerbated the crisis. In addition, the explosion in Beirut destroyed a large part of the city's infrastructure and further hampered economic activity. The post also discusses the International Monetary Fund (IMF) negotiations that took place in 2020 and 2021, as Lebanon sought to secure a loan to help stabilize its economy. However, the negotiations were stalled due to political disagreements and concerns about the country's ability to implement the necessary reforms. The post concludes by noting that Lebanon's economic crisis is far from over, and that the country will need significant international support and cooperation to overcome its challenges (“Part I of Crisis in Lebanon: Economic ‘Free Fall,’ IMF Negotiations, and Beirut Explosion” 2021).

In the second part of the blog post discusses the interrelated challenges that contributed to the economic and political crisis in Lebanon. The post identifies the complex web of factors, including a high level of public debt, a large trade deficit, a bloated public sector, and corruption, that created the conditions for the crisis. The authors highlight the impact of the Syrian conflict on Lebanon, including the influx of refugees and the spillover of violence. The post also explores the role of sectarianism in Lebanese politics and the challenges it presents for governance and economic development. Finally, the post outlines the steps that Lebanon must take to address the crisis, including fiscal reforms, a restructuring of the banking sector, and the establishment of a transparent and accountable government ("Part II of Crisis in Lebanon: Buildup of Interrelated Challenges" 2021).

Research conducted on predicting banking crises using machine learning models has been done on data retrieved from developed economies and countries. Therefore, the purpose of this research study is to attempt and apply machine learning techniques on data retrieved from Lebanon to examine and investigate the banking, financial and currency crisis of 2019. This research also aims to shed light on the different benefits that such models can bring to the banking and financial industry, especially in a country with heavy political inertia, significant association between policy makers and bank owners and a hindered independence of the financial regulator. The studies conducted by Demirgüç-Kunt et al. (2005), Babecký et al. (2014), Drehmann et al. (2014), and Beutel et al. (2019), are used as a steppingstone to apply several Statistical, Econometric, and Machine Learning methods that aim to predict banking crises and develop early warning indicators in the case of Lebanon.

Potentially effective models could serve as essential tools rather than just being statistical models and algorithms: models and early warning indicators could be used across the Lebanese banking industry - in a generalized and standardized manner - such that they overcome the limitations of the different approaches adopted by the different banks operating in Lebanon. After validation of these models, the goal is to eventually implement them in the risk management departments of different Lebanese banks. By doing so, this research will open the door for further research to understand and evaluate the efficiency and accuracy of the predictive power of machine learning tools in the Lebanese banking sector as well as the efficiency of early warning indicators. Should these tools prove to be efficient, risk managers should then be able to optimize their hedging and provisioning strategies against systemic banking crises that might arise in the future.

Being armed with effective tools requires enforcement which highlights the importance of the role of policy and policymakers. Reiterating, Laeven (2011) affirms that vulnerabilities in a banking system can be the result of policies at both micro and macro levels. Therefore, it is crucial to delve into policy reform opportunities as it is evident that Lebanon’s financial crisis is intertwined with its political one. In the “Report on The Third Webinar Series Lebanon’s Economic Crisis: Is There a Way Forward?” (2022), provides an overview of the discussions and insights shared during several webinars. In fact, the Olayan School of Business at the American University of Beirut organized webinars and brought together a range of experts, including economists, policymakers, and civil society representatives, to discuss the ongoing economic crisis in Lebanon and potential solutions to address the country's challenges.

The report highlights several key themes that emerged during the webinar, including the need for political and economic reforms, the importance of international support, and the potential role of the private sector in promoting economic growth. The experts discussed the impact of the crisis on different sectors of the economy, including the banking sector, the labor market, and the informal economy. They noted that the crisis has had a profound impact on the most vulnerable populations, including refugees, women, and youth. The experts also discussed potential policy solutions to address the crisis, including fiscal and monetary reforms, structural reforms, and the promotion of private sector-led growth. They emphasized the need for a comprehensive and coordinated approach to address the crisis and restore confidence in the economy.

The authors of the report begin by presenting the agreement that was signed between Lebanon and the IMF in April of 2022. The agreement is for a four-year Extended Fund Facility (EFF) for a value of $3 billion. The agreed-upon program is based on five main pillars: restructuring the financial sector, implementing fiscal reform along with a debt restructuring strategy, reforming state-owned enterprises (SOEs), strengthening the state's governance, anti-corruption, and anti-money laundering framework, and exchange rate unification and moving to a credible exchange rate regime (Faour and Jamali, 2022). Diving deeper into the first pillar, the authors explain that this pillar necessitates the restructuring of the financial sector along with a loss-allocation – which is challenging from a political aspect. Restructuring and loss-allocation require a cabinet approval of a bank restructuring strategy, legislation of a bank resolution framework, and initiating an externally aided Asset Quality Review (AQR) of the 14 largest banks in Lebanon.

The second pillar focuses on fiscal reform and debt restructuring to achieve fiscal and debt sustainability. The plan aims to achieve a primary deficit of 4% in 2022 through a combination of fiscal reforms and import taxation at a unified exchange rate. However, the implementation of this measure has faced resistance, which could make it challenging to achieve fiscal sustainability. On the debt restructuring front, there has been little progress in initiating negotiations with creditors due to the lack of progress in implementing an IMF program, which is typically a precondition for meaningful negotiations. As a result, market expectations for a potential conclusion of debt restructuring are pessimistic, and the country's Eurobonds are currently trading at between 6 and 7 cents to the dollar, with a projected recovery value of zero if an IMF program does not materialize.

The third pillar of the economic reform plan for Lebanon focuses on the reform of state-owned enterprises, particularly the state-owned electricity sector that has accumulated a debt of $40 billion. The plan aims to make the sector self-sufficient and not a drain on the country's fiscal position by increasing tariffs to account for the market exchange rate and higher prices of fossil fuels. The fourth pillar concerns institutional reforms aimed at strengthening the governance, anti-corruption, and anti-money laundering framework of the country. This involves reforming the decades-old bank secrecy law to align it with international standards, which may entail undoing bank secrecy altogether. However, the main risk lies in the implementation of the legislation, given the country's track record in this regard.

The fifth pillar involves reforming the exchange rate regime by unifying the exchange rate and moving to a flexible, market-determined exchange rate system while implementing a capital controls law. However, three years after the economic collapse began, a capital controls law is yet to be legislated due to political bickering (Faour and Jamali, 2022).

The report further discusses the challenges Lebanon faces in implementing its agreement with the IMF to address its economic crisis. The country's economic woes are attributed to elite capture, extractive institutions, weak governance, and rampant corruption. The report cites a World Bank report from 2017 that warned of the potential consequences of a sudden stop scenario, which is now unfolding in the country. Despite the agreement with the IMF, the risk of the plan not being implemented or going off track is high due to Lebanon's political context. The country's economic model did not serve the middle class or the poor well, and the ruling elites instituted a patronage system that allowed them to capture the state and its associated economic rents. The UNHCR has accused the Lebanese state of being responsible for human rights violations and the unnecessary immiseration of the population. The report also presents a series of webinars hosted by the Olayan School of Business at the American University of Beirut, which discussed the ongoing economic, banking and financial crises in Lebanon. In the first webinar, Dr. Marwan Barakat discussed the causes and current outlook of the crisis. He highlighted that the crisis began in 2014 when the Central Bank's net foreign exchange position turned negative, and it was exacerbated in 2016. He projected two scenarios for real GDP growth in 2022: an optimistic scenario with 5% growth if there is an agreement with the IMF, and a pessimistic scenario with a 6% contraction if there is no agreement.

Dr. Barakat suggested that restructuring the banking sector is necessary, and it should entail shrinking the sector's asset base and deposit base and delinking the balance sheets of the banking sector from that of the Central Bank. In the second webinar, Dr. Wissam Harake discussed how the contraction in economic output began before the crisis broke in 2019, and how the financial crisis exerted the largest effect on the economy, leading to the disintegration of the pillars of Lebanon's political economy.

Dr. Harake noted that the crisis is likely to rank among the top ten, and possibly the worst three crises globally, based on the Crisis Severity Index. Dr. Harake argued that the status quo can lead to national fragmentation and a breakdown of social peace, and emphasized the need to urgently adopt and implement a credible, comprehensive, equitable macro-financial stabilization strategy if Lebanon is to avoid complete destruction of its social and economic networks. The strategy would be predicated on restructuring the financial sector, achieving unification and stability in the exchange rate, debt restructuring, fiscal adjustment, and enhanced social protection. The third webinar hosted Mr. Mike Azar, who discussed the challenges of addressing losses in Lebanon's banking sector and the negotiations with the IMF. He emphasized the complexity of the bank restructuring process due to the link between the balance sheets of the central bank and the banking sector. He argued that addressing the gap in the financial system would require a large proportion of depositors to take a hit, and that the gap was a consequence of mismanagement in the public sector and central bank over several decades. Mr. Azar also discussed Lebanon's macroeconomic conditions and the negotiations with the IMF. He noted that the current account deficit was widening due to limited production capacity, which could lead to full dollarization as the LBP loses value.

He argued that the IMF's SLA agreement with Lebanon was very generic and did not address difficult questions, and that it was unlikely for Lebanon to obtain an IMF program. In the fourth webinar, Dr. Sami Geadah discussed that while Lebanon and the IMF have agreed to an SLA, a full-fledged Extended Fund Facility (EFF) program will require approval from the IMF's executive board.

An EFF aims to stabilize the country, place debt on a sustainable footing, and address fiscal and external imbalances while allowing the government to provide essential public services. To be eligible for exceptional access to financing from the IMF, additional criteria have to be met. An EFF will provide Lebanon with direct financial support, building credibility with donors and bond holders, debt restructuring, convincing domestic parties that the reforms are appropriate, and obtaining policy advice and technical assistance from the IMF. However, the prior actions spelled out in the SLA may delay executive board approval of the EFF. Dr. Geadah suggested that Lebanon should embark on the reforms that do not need parliamentary approval to initiate the reforms in the interest of time. The IMF will monitor the program's progress using specific performance criteria and benchmarks. In the fifth and final webinar, Dr. Alain Bifani discussed the optimal utilization of oil and gas resources in Lebanon to maximize economic returns. He emphasized the importance of discussing Lebanon's potential hydrocarbon windfall and noted that the costs associated with extraction include setting up infrastructure, drilling, and building platforms for gas treatment and pipelines.

Dr. Bifani argued that the revenues from gas exports should not be used to repay debt but should be placed in a Sovereign Wealth Fund that invests in assets outside Lebanon to avoid the Dutch disease. He also stated that gas revenues will not be a game changer for Lebanon, and structural reforms will still be necessary (Faour and Jamali, 2022).

Effective policy is crucial for resolving financial and banking crises. Policy decisions can determine the speed and effectiveness of a response to a crisis, as well as the long-term stability of the financial and banking systems. Policies can include measures to address imbalances in the economy, improve governance and transparency, strengthen financial regulation and supervision, and ensure that financial institutions are adequately capitalized. Policy makers must also address issues related to debt sustainability and restructuring, and work to restore confidence among investors and the public. Without effective policies, financial and banking crises can continue to worsen, leading to a downward spiral that can be difficult to reverse.

CHAPTER III

DATA

A. The Data

1. Sources

Building the work conducted of Beutel et al. (2019), we construct a dataset of systemic banking crises for 40 economies over the past 12 years, using Laeven and Valencia's database (latest update: 2020) and using the database for systemic banking crises established by the European System of Central Banks (ESCB) and the European Systemic Risk Board (ESRB) used by Beutel et al. (2019). Countries included in our dataset were selected based on Data Availability.

Also, macroeconomic, and financial data is collected, on quarterly basis, from several sources: The IMF's International Financial Statistics Database, The OECD's Database, The World Bank's Database, The BIS' Database, Individual Countries' Central Banks' Statistics, Thomson Reuters' Data Stream, Statista website and Blom Invest BRITE Database. Laeven and Valencia's latest database contains detailed information on banking and financial crises:

TABLE NAME	INFORMATION
CRISIS YEAR	165 Countries Systemic Banking Crises: Start Year Currency Crisis: Year Sovereign Debt Crises: Year Sovereign Debt Restructuring: Year
MONTHLY CRISIS DATES	Banking crisis (start year and month) Sovereign debt crisis (start year and month) Currency crisis (start year and month)
CRISIS FREQUENCY	Frequency of Individual Crises per Year Frequency of Twin Crises per Year Frequency of Triple Crises
BANKING CRISIS RESOLUTION AND OUTCOMES	Information by country on: Crisis Start and End dates & Outcomes
ADDITIONAL INFORMATION ON BANKING CRISES	Further details per country over time regarding Banking Crises

Table 1: Summary of Laeven & Valencia's Database

2. Variables

Our dependent variable is extracted from Laeven and Valencia’s database. It takes a value of One from the Start Date of the Crisis occurrence and until its end. It takes a value of Zero otherwise.

a. Credit Quality

As discussed in the literature, banking crises usually follow credit booms. In fact, when the debt to the private sector increases, risks to the financial system increase as well, especially when booms in asset prices are financed by debt (Beutel et al., 2019). When borrowers near default, asset prices decrease and borrowers are further unable to repay their debt (Kindleberger and Aliber, 2005; Jordà et al., 2015).

As a result of the above, banks may be forced to reduce leverage which may in turn prompt a credit crisis and potentially induce a recession. Furthermore, Beutel et al. (2019) discuss that runs on banks take place when banks' net worth decreases and depositors lose confidence in the institutions or in the entire sector (Allen and Gale, 2007).

To capture Credit Quality, we include changes in Credit to the Private Sector (non-financial institutions) and changes in Nonperforming Loans (as a percentage of total loans granted) and debt service ratios.

b. Macroeconomic Environment

Changes in the macroeconomic environment are closely related to the credit environment. According to Beutel et al. (2019), fast economic growth increases risk appetites and ultimately credit growth (Drehmann et al., 2011; Kindleberger and Aliber, 2005; Minsky, 1982).

Additionally, real economic slowdowns can lead to difficulties in repayment from the borrowers' side inducing as such difficulties in the financial sector (Allen and Gale, 2007).

To capture changes in the real economic environment, we look at changes Broad Money Supply, changes in Public Debt, and Inflation. Furthermore, we include three-month interest rates to investigate investors and banks' willingness to take on excessive risks when interest rates are low (Maddaloni and Peydró, 2011; Allen and Gale, 2007; Rajan, 2006). Contrariwise, a sudden increase in interest rates may pressure banks (Minsky, 1982).

c. External and Global Imbalances

As discussed by Beutel et al. (2019), the external contributed significantly to the literature on early warning indicators (Frankel and Rose, 1996; Kaminsky and Reinhart, 1999) and balance-of-payments crises. In Beutel et al. (2019)'s work, classic balance-of-payment crises are of a lesser concern since the countries considered in their paper have advanced economies. Yet, external imbalances may still add to vulnerabilities.

The reasoning behind credit quality and external imbalances are somewhat similar, as conveyed by Beutel et al. (2019): they discuss that large capital inflows from abroad may boost asset price booms but create a reverse effect when inflows drop or stop altogether (Kaminsky and Reinhart, 1999; Calvo, 1998). Similarly, regarding credit growth, inflows from abroad to the banking sector could support credit growth. Hence, to capture these effects, Narrow Money Supply is included, as well as the Real Effective Exchange Rate, the Nominal Effective Exchange Rate. Finally, global events may impact domestic banking systems through contagion, financial sector interconnectedness and trade links (Kaminsky and Reinhart, 2000).

So, we add oil prices as an indicator for changes in the global environment. Kaminsky and Reinhart's paper, "The Twin Crises: The Causes of Banking and Balance of Payments Problems," explores the relationship between financial crises and balance of payments crises.

The authors argue that these two types of crises are often interrelated, with one crisis leading to the other. They also identify a number of common causes of both types of crises, including capital inflows, overvaluation of the currency, and high levels of external debt. The paper draws on a dataset of 20 countries that experienced twin crises in the period from 1970 to 1993 and uses statistical analysis to identify the key factors that contributed to these crises. The authors find that financial liberalization and large capital inflows are important predictors of twin crises, as are high levels of external debt and overvaluation of the currency. The paper's findings suggest that policymakers need to pay close attention to both the banking sector and the balance of payments when designing policies to prevent financial crises. They also emphasize the importance of early intervention in addressing potential problems in these areas, as delays can exacerbate the crisis and lead to greater economic damage. Overall, the paper provides valuable insights into the causes and consequences of financial and balance of payments crises and highlights the need for effective policy responses to address these challenges (Kaminsky and Reinhart, 1999).

B. Data Transformation & Processing

a. Countries' Selection

To compile the data for this research, Laeven and Valencia's Database is used as the main database since it contains detailed information about each country's crisis type, date and other relevant information. The initial data collection started with a total of around 120 countries which had witnessed Banking, Sovereign Debt, Currency Crises, twin crises (any combination of the aforementioned three crises) or a triple crisis, between the year 1970 and the year 2017. Then, data was collected on the variables of interest over a time-span ranging from the first quarter of the year 2010 and until the last quarter of the year 2021. During this phase, the following issues were encountered: missing series (entirety of the data for a certain variable) for many countries, data available in different frequencies for variables and countries (data of interest is compiled on a quarterly basis) and missing values within series. As a consequence, data was restricted to a list of forty countries, as detailed in the table hereafter.

Following the methodology employed by Beutel et al. (2019), and since it is more important to have sufficient data than to have a complete set of all indicators, missing values will be excluded from the dataset.

b. Variables' Transformation

Macroeconomic data can be found in different forms: Values, period-to-period change, percentage change, as a percentage of GDP, among other forms. Data examined in Beutel et al. (2019)'s work is expressed either in real terms or as a share of GDP.

Concerning the dates, data is collected starting Q1 2010 and until Q4 2021. However, it is worth noting that data availability before and after said dates is an issue in certain countries. As for the variable regarding banking crises' occurrence, it is a binary variable that takes the value of 1 during the quarters where the crisis occurs or is ongoing and 0 otherwise[4]. Credit to Private Sector: represents the value of total credit extended to the private sector, excluding financial institutions and is expressed as a percentage of GDP[5]. Debt Service Ratio (DSR): debt service ratio of the private, non-financial sector, expressed as a percentage of exports. This is a common metric used to measure an entity's ability to cover its debt payments, so the higher it is, the better[6]. Nonperforming loans (NPL): the value of total nonperforming loans expressed as a percentage of total loans outstanding.

The lower this ratio is, the better[7]. Narrow Money: is a classification of money supply that includes physical money (coins and paper), demand deposits, and other liquid forms held by the central bank[8]. Broad Money: is a measure of the amount of money in a national economy, including highly liquid forms (narrow money) and less liquid forms[9]. Public Debt: represents total borrowings by a government[10].Net Foreign Assets: represents the value of a country's assets abroad, less the value of domestic assets owned by foreigners, and relates to the balance of payments. [11] Real Effective Exchange Rate (REER): represents a country's weighted average currency relative to an index or basket of other currencies. Weights are

[4] Source: Laeven & Valencia's Database
[5] Source: Bank for International Settlements
[6] Source: Bank for International Settlements
[7] Source: The World Bank
[8] Source: OECD
[9] Source: OECD
[10] Source: The IMF
[11] Source: Corporate Finance Institute

determined by comparing the relative trade balance of a country's currency against that of each country in the index.[12] Nominal Effective Exchange Rate (NEER): is an unadjusted weighted average rate at which a country's currency exchanges against a basket of multiple foreign currencies. This rate represents the amount of domestic currency required to purchase foreign currency.[13] 3-Month Interest Rates: the rates at which short-term government paper is issued or traded in the market[14]. Consumer Price Index (CPI): measures the average change over time in prices paid by consumers for a market basket of consumer goods and services[15]. Oil Prices (BRENT Oil) represent an indicator for global developments [16].

[12] Source: Bank of Canada
[13] Source: Bank of Canada
[14] Source: OECD
[15] Source: US Bureau of Labor Statistics
[16] Source: Beutel et al. (2019)

VARIABLE NAME	DEFINITION	SOURCE
Banking Crisis	Dummy Variable That Equals One If There Is a Banking Crisis and Zero Otherwise	Laeven & Valencia's Database (2020)
Credit Expansion	Debt To Private Sector in Real Terms	BIS Database, BDL
Financial Measures	Debt Service Ratio Expressed as a % of Exports Level of Non-Performing Loans to Total Loans 3-Months Interest Rates	BIS Database FRED, STATISTA, BRITE, BDL
Money Supply	Levels Of Narrow and Broad Money Indices	FRED, OECD, BRITE, US Data.Gov IMF IFS, BDL, OECD Data Malaysia Central Bank
Public Debt	Levels Of Public Debt in Real Terms	FRED, BRITE, Office of National Statistics (UK), National Bank of Belgium Australian Government (Taxation Office), Bank of Canada, Nasdaq Data, The Global Economy
Net Foreign Assets	Value Of Foreign Assets Abroad Less Value of Domestic Assets Owned by Foreigners	World Bank Data, BRITE, US Data.Gov FRED, BIS, BDL, IMF IFS
Inflation	Consumer Prices Index	FRED, BRITE, Office of National Statistics (UK), National Bank of Belgium
External Imbalances	Real Effective Exchange Rate, Nominal Effective Exchange Rate, Oil Prices (Brent Oil)	FRED, BRITE, Argentina's Central Bank, World Bank, US Data.Gov

Table 2: Description of Variables

Country	Systemic Banking Crisis (starting date)	Currency Crisis	Sovereign Debt Crisis (year)	Sovereign Debt Restructuring (year)
Argentina	1980, 1989, 1995, 2001	1975, 1981, 1987, 2002, 2013	1982, 1989, 2001, 2014	1993, 2005, 2016
Australia				
Austria	2008			
Belgium	2008			
Brazil	1990, 1994	1976, 1982, 1987, 1992, 1999, 2015	1983	1994
Canada				
Chile	1976, 1981	1972, 1982	1983	1990
China, P.R.	1998			
Colombia	1982, 1998	1985		
Czech Republic	1996			
Denmark	2008			
Finland	1991	1993		
France	2008			
Germany	2008			
Greece	2008	1983	2012	2012
Hungary	1991, 2008			
India	1993			
Indonesia	1997	1979, 1998	1999	2002
Ireland	2008			
Italy	2008	1981		
Japan	1997			
Korea	1997	1998		
Lebanon	1990	1984, 1990		
Luxembourg	2008			
Malaysia	1997	1998		
Mexico	1981, 1994	1977, 1982, 1995	1982	1990
Netherlands	2008			
New Zealand		1984		
Norway	1991			

Country	Systemic Banking Crisis (starting date)	Currency Crisis	Sovereign Debt Crisis (year)	Sovereign Debt Restructuring (year)
Poland	1992		1981	1994
Portugal	2008	1983		
Russia	1998, 2008	1998, 2014	1998	2000
Spain	1977, 2008	1983		
Sweden	1991, 2008	1993		
Switzerland	2008			
Thailand	1983, 1997	1998		
Turkey	1982, 2000	1978, 1984, 1991, 1996, 2001	1978	1982
United Kingdom	2007			
United States	1988, 2007			

Table 3: List of Countries of Interest

C. The Final Dataset

The final dataset consists of 40 countries and 15 exogenous variables with observations over approximately 48 quarters per observation and country. So, a total of 720 observations are collected per country for a total of 25,500 observations across the dataset, after resolving the issue of missing values.

Country coverage and crisis dates

Country	Data Availability		Crisis Dates *	
	Start	End	Start	End
Argentina	Q1 2010	Q4 2021	-	-
Australia	Q1 2010	Q4 2021	-	-
Austria	Q1 2010	Q4 2021	Q1 2010	Q2 2012
Belgium	Q1 2010	Q4 2021	Q1 2010	Q2 2012
Brazil	Q1 2010	Q4 2021	-	-
Canada	Q1 2010	Q4 2021	-	-
Chile	Q1 2010	Q4 2021	-	-
China	Q1 2010	Q4 2021	-	-
Colombia	Q1 2010	Q4 2021	-	-
Czech Rep.	Q1 2010	Q4 2021	-	-
Denmark	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Finland	Q1 2010	Q4 2021	-	-
France	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Germany	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Greece	Q1 2010	Q4 2021	Q1 2010	Q4 2012
Hungary	Q1 2010	Q4 2021	Q1 2010	Q4 2012
India	Q1 2010	Q4 2021	-	-
Indonesia	Q1 2010	Q4 2021	-	-
Ireland	Q1 2010	Q4 2021	Q1 2010	Q4 2012
Italy	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Japan	Q1 2010	Q4 2021	-	-
Korea	Q1 2010	Q4 2021	-	-
Lebanon	Q1 2010	Q4 2021	Q3 2019	Q4 2021
Luxembourg	Q1 2010	Q4 2021	Q1 2010	Q4 2012
Malaysia	Q1 2010	Q4 2021	-	-
Mexico	Q1 2010	Q4 2021	-	-

Netherlands, The	Q1 2010	Q4 2021	Q1 2010	Q1 2010
New Zealand	Q1 2010	Q4 2021	-	-
Norway	Q1 2010	Q4 2021	-	-
Poland, Rep. of	Q1 2010	Q4 2021	-	-
Portugal	Q1 2010	Q4 2021	Q1 2010	Q4 2012
Russian Federation	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Spain	Q1 2010	Q4 2021	Q1 2010	Q4 2012
Sweden	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Switzerland	Q1 2010	Q4 2021	Q1 2010	Q1 2010
Thailand	Q1 2010	Q4 2021	-	-
Turkey	Q1 2010	Q4 2021	-	-
United Kingdom	Q1 2010	Q4 2021	Q1 2010	Q4 2011
United States	Q1 2010	Q4 2021	Q1 2010	Q4 2011

Table 4: Countries & Crises in the Dataset

* Countries with crisis start date as Q1 2010 actually have an earlier crisis start date, being Q3 of 2008 (corresponding to the start of the global crisis). Countries with no crisis in our selected timeframe have had at least one crisis as these countries are recorded in Laeven and Valencia’s database (last update:2020)

D. Exploratory Data Analysis

The aim of this section is to explore the data and potential relationships or interactions between variables, enabling further insightful exploration and analysis. Data is visualized geographically, showing the countries where crisis occurred.

Then, total occurrences are visualized. Data shows that less than 10% of the observations represents Crises Occurrences: the total number of crises occurrences is 126, corresponding to 7% of total Data and as such, the dataset is considered imbalanced. For the visualizations, please refer to the Appendix.

CHAPTER IV

METHODOLOGY

We use a total of fifteen exogenous variables which capture major economic changes that might impact the banking sector and hence the likelihood of the occurrence of a systemic banking crisis. All variables are either indices, expressed in real terms, or expressed as a share of GDP. For our model specifications, we use different datasets comprising of different combinations of the selected variables. The first dataset uses all of the available variables. Then, we specify smaller models using subsets of variables in order to illustrate the relative performance of methods across datasets depending on complexity. Generally, reduced complexity reduces models' predictive performance. Yet, this may be offset by information gained from fitting models using fewer variables.

The below table summarizes the different datasets:

DATASET	**VARIABLES**
Dataset	All Variables
Dataset 1	Credit & Credit Quality (Credit to Private Sector, DSR & NPL)
Dataset 2	Money Supply & Foreign Assets
Dataset 3	3-months Interest Rates, Public Debt & Broad Money
Dataset 4	External Imbalances (CPI, REER, NEER & Oil Prices)

Table 5: Description of Datasets

In order to guarantee comparability across the different datasets, we use the same sample for all datasets. Following the methodology of Beutel et al. (2019), these variables were specified a priori to allow understanding of the data in each dataset. In addition, limiting the research to a set of pre-specified variables, relevant to the field of study, reduces the potential risks of data mining and overfitting, as explained by Inoue and Kilian (2005). This is particularly critical for early-warning models used to implement policies since an overestimation of models' accuracy might suggest that a financial sector is secure whereas it might not actually be.

A. Data Preprocessing

Our data naturally contains a time trend (the exception being inflation and money market rates). To remove the time trend, we employ a growth rate approach, as it is one of the two of the most frequently used methods in literature on early warning (the other being the Hodrick-Prescott (HP) filtering approach), by computing changes between periods.

The formula for period-to-period growth rate is: *(Q+1 value – Q value) / Q value * 100*

Where:

Q value is the initial value of the variable being measured

Q+1 value is the value of the variable being measured at the following quarter

While doing so, a significant loss in data occurs and so we opt to keep the data as is, pre-transformation.

B. Description of Models

Method Name	Hyperparameter Name	
Logit	Exogenous Variables	
Trees	Complexity Parameter	
KNNs	K: Number of Nearest Neighbor	
Random Forests	Node Size: Minimum # of Observations per Final Node of Each Tree in the Forest	Number of Variables Randomly Sampled as Candidates at Each Split
SVMs	Parameter in Radial Basis Function	Cost of Violation of Soft Margin Constraint
Neural Networks	Learning Rate Parameter Controlling the Rate of Convergence of the Learning Algorithm	NeN Size: Number of Nodes in Hidden Layers

Table 6: Summary of Models Employed

1. Logistic Regression (Logit)

As mentioned by Beutel et al. (2019), Logistic Regression (Logit) models are the most commonly used models in early warning literature (Frankel and Rose, 1996; Bussière and Fratzscher, 2006; Lo Duca and Peltonen, 2013).

Such models are based on two assumptions: the first assumption is that the dependent binary variable is driven by a latent process which is in turn related to the explanatory variables in a linear way. The second assumption is that the latent process is assumed to be linked to the binary variable by a logistic transformation.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \varepsilon_i$$

Equation 1: Logit Equation

a. Logistic Regression using SMOTE

Using SMOTE oversampling technique, the length of the new oversampled data is 2,226 which are randomly generated and perfectly balanced with 50% of the data representing crisis occurrences at 1,113 records and the remaining 50% representing non-crisis occurrences, also at 1,113 records.

We run the logit model and find that the significant predictors of banking crises are: Credit to Private Sector, Debt Service Ratio, Narrow Money Supply, Real Public Debt, Real Broad Money Supply, Real Net Foreign Assets, Real Effective Exchange Rate, Nominal Effective Exchange Rate, 3-months interest rates and Oil Prices (see Appendix).

b. <u>Logistic Regression using Random Oversampling</u>

Using the oversampling technique, we find that the significant predictors are: Credit to Private Sector, Narrow Money Supply, Real Public Debt, Real Broad Money Supply, Real Effective Exchange Rate, Nominal Effective Exchange Rate, and 3-months interest rates (see Appendix). Then, we construct a subset of the data such that the proportion of crisis to no-crisis is maintained. We rerun the logit model and find that the significant predictors are: Credit to Private Sector, Debt Service Ratio, Narrow Money Supply, Real Effective Exchange Rate, Nominal Effective Exchange Rate and Oil Prices (see Appendix). Comparing both oversampling techniques, we observe better results with Random Oversampling.

2. *Trees*

Trees are made up of a root node, decision nodes (branches) and final nodes (leaf nodes). A decision tree has a flow-chart-like structure: at the Root Node, the entirety of the sample is represented. A Decision Node is the node where a sub-node is split into further sub-nodes. A Leaf Node represents a class label or a class distribution. Splitting is the process of separating a node into two or more sub-nodes.

Binary decision trees essentially group data into different clusters by successively comparing the values of variables to a certain threshold. The major task performed by the tree is to determine these thresholds in an optimal way, and to select the sequence in which variables will be compared; this process determines the ultimate form of the tree. The selection of the hyperparameter thus decides on the complexity of the tree.

Lower complexity costs imply a decrease in classification error on the training data but at the same time implies an increase in the potential for overfitting. Split can be determined based on

Entropy, $$E(S) = -\sum_{i=1}^{c} p_i \log_2 (p_i)$$

Equation 2: Entropy Equation

Information Gain, $$IG(F) = E\left(S_{before}\right) - E(S_{after})$$

Equation 3: Information Gain Equation

or the Gini Impurity $$G = \sum_{i=1}^{C} p_i(1 - p_i)$$

Equation 4: Gini Impurity Equation

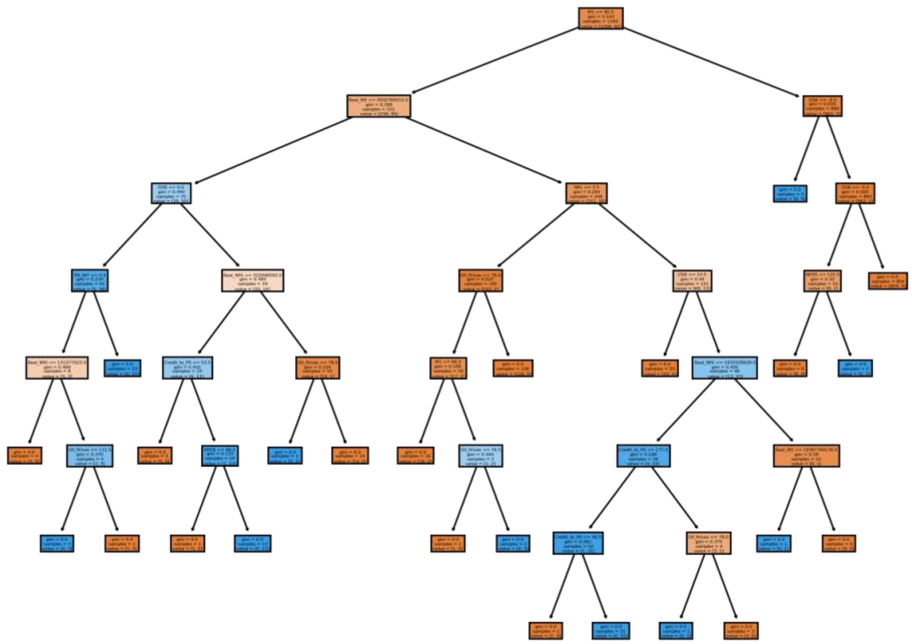


Figure 6: Decision Tree

Above is the pruned tree with 5 end nodes: the significant predictors of banking crises are: Credit to Private Sector, Real Effective Exchange Rate, Broad Money Supply and Oil Prices.

3. KNNs:

The k-nearest neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used for both classification and regression. KNN is non-parametric but rather stores the available data and classifies a new data point based on similarity to the nearest neighborhood. Distance between classes can be measured by either the Manhattan Distance, Euclidean Distance or the Maximum Distance:

$$\|X - X'\| = \sqrt{\sum_{i=1}^{d} (Xi - Xi')}$$

Equation 5: Euclidean Distance Equation

4. Random Forests: XGB Classifier

Random forests (Breiman, 1996, 2001) represent a large number of different decision trees. Forests generalize trees by averaging the predictions of the trees which can reduce the variance of estimates and ultimately reduce prediction errors. To improve stability and accuracy, Random Forests generate heterogeneity from several homogenous tree models by bagging (bootstrap aggregating, random sub-sampling) or boosting (adding models sequentially to improve the final model).

As mentioned by Beutel et al. (2019), Random Forests have been the most frequently employed machine learning technique in early warning literature (Alessi and Detken, 2018; Holopainen and Sarlin, 2017; Tanaka et al., 2016).

However, they explain that, reducing variance and improving out-of-sample predictive performance depends on achieving a reduced correlation between the randomly generated trees (Breiman, 2001). Beutel et al. (2019) claim that attaining a sufficiently low degree of correlation could be challenging in the presence of correlation in the training data. In our work, we attempt to improve the random forest classifier by boosting: we employ Extreme Gradient Boosting which yielded high accuracy rates.

5. SVMs

Support vector machines (SVMs) are supervised learning methods that can be used for Classification, Regression and Outlier Detection. SVMs use a hyperplane to separate the data points and choose the hyperplane with maximum margins. The optimization criterion in this case is the width of the margin (the distance) between the classes and the ultimate goal is to find the hyperplane with the largest margin by finding the support vectors and their weights.

Equation of the hyperplane: $\boldsymbol{w}Tx + b = 0$

Where b is a single number representing bias: if hyperplane passes through origin: $b = 0$ and if b > 0 this means it moves parallel to itself in the direction of $\boldsymbol{w}$.

Distance of a point xn from a hyperplane (can be positive or negative)

$$\gamma_n = \frac{\boldsymbol{w}^T x_n + b}{||\boldsymbol{w}||}$$

$\boldsymbol{w}$ is the Euclidean norm (distance from origin to vector $\boldsymbol{w}$).

6. Neural Networks

As the name indicates, Artificial Neural Networks are inspired by the human brain and represent a network of simple network nodes (neurons), arranged sequentially. When taken together, they can approximate functional forms and link the input and the output of the neural network. The neurons are the building blocks of the network. As a deep learning method, it learns hierarchal representation from data and scales with more data. The architecture of the network (i.e., the number of hidden layers and the number of neurons in each hidden layer) are predetermined as hyperparameters. The algorithm's main task is to estimate the weights connecting the neurons of adjacent layers (Beutel et al. (2019)). Usually, weights are estimated with the goal of minimizing a certain loss function. When no hidden layers are determined, Neural networks incorporate a simple logit model.

For all models, we proceed to train the models on training datasets which represents a subset of the data. Then, we proceed to test the models' performance on the test set. Next, we run the models on the entire dataset and test for Lebanon as unseen data.

C. Model Comparison: Advantages and Disadvantages

	Benefits	Drawbacks
logit	Explicit probabilistic foundations	Pre-specified functional form
	High interpretability	
knn	Simple approach	Strong curse of dimensionality
trees	Automatic variable selection	Instability across time/samples
	Intuitive approach	
rf	More stable than trees	Risk of overfitting
	Improves on tree accuracy	Complex drivers of predictions
svm	Flexible nonlinear fitting	Risk of overfitting
	Computationally efficient	Ad hoc in probabilistic setups
		Difficult to communicate
nen	Flexible functional form	Risk of overfitting
	Recent advances in the literature	Computationally expensive
		Difficult to communicate

Table 7: Comparison of Models (source: Beutel et al., 2019)

Method	Hyperparameter Name	Value	Comment
KNN	k (number of nearest neighbors to use for each prediction)	K = 5	K-means inertia Using the Elbow Method, we determine that K = 5
XGBoost	Booster: Gradient Boosting Objective: Binary Evaluation Metric: Log Loss	# of Boosting Rounds: 1000 Early Stopping Rounds: 10	Early stopping set at 10 to conserve memory
SVM	Kernel Cost Function	Radial Basis Function C = 0.01	Following the methodology of Beutel et al. (2019)
NEN	Learning rate parameter controlling the rate of convergence of the learning algorithm	Learning Rate = 0.001 Epochs = 100 Optimizer = SGD	Most Commonly Used
K-Means (Clustering)	K (number of clusters)	K = 5	K-means inertia K determined using Elbow Method

Table 8: Models & Hyperparameters

CHAPTER V

FINDINGS

Based on our data and the models employed on different subsets of the data, we conclude that machine learning techniques – namely KNNs and Extreme Gradient Boosting Tree models (XGBOOST) – are able to classify an occurrence of a crisis with a certain degree of accuracy. Unlike the results reached by Beutel et al. (2019) that simple logit models outperform all machine learning methods considered under a variety of circumstances, we find that, based on our datasets, machine learning methods are able to achieve accurate classification results.

Following the methodology employed by Beutel et al. (2019), and to test our results, we use different combinations of data by creating different datasets which include different combinations of predictors. However, it is worthy to note that the timing of crises differs between countries, depending on how quickly it was resolved.

Finally, and based on data availability, our dataset includes countries that are not particularly similar in terms of size, economic growth, political aspect, and other characteristics.

Beutel et al. (2019) discuss the evaluation of predictions using different performance measures. The performance of early warning models can be evaluated with respect to either binary signals or probabilities. Their work presents four standard performance measures: relative usefulness, F-measure, area under the curve (AUC), and Brier probability score (BPS). The authors deliberately do not use cross-validation to evaluate models due to potential over-estimation of performance. The performance measures are based on the recursive predictions for the out-of-sample part of the dataset.

Performance evaluation using different measures

Results	F1	AUC	Accuracy
LOGIT	**0.1**	**0.52**	**0.93**
KNN	0.61	0.85	0.95
XGBOOST	**0.9**	**0.99**	**0.98**
SVM	0	0.49	0.94
NEN	0.1	0.65	0.78
LOGIT1	0	0.5	0.92
KNN1	**0.67**	**0.88**	**0.96**
XGBOOST1	0.83	0.986	0.97
SVM1	**0**	**0.76**	**0.91**
NEN1	0.1	0.56	0.89
LOGIT2	0	0.5	0.92
KNN2	0.32	0.82	0.92
XGBOOST2	0.75	0.97	0.96
SVM2	0	0.44	0.91
NEN2	0	0.82	0.91
LOGIT3	0	0.5	0.93
KNN3	0.21	0.81	0.91
XGBOOST3	0.8	0.95	0.97
SVM3	0	0.48	0.94
NEN3	**0**	**0.87**	**0.92**
LOGIT4	0	0.5	0.92
KNN4	0.34	0.81	0.88
XGBOOST4	0.29	0.92	0.9
SVM4	0	0.7	0.92
NEN4	0	0.69	0.91

Table 9: Performance Evaluation

The table shows the performance of different machine learning algorithms in predicting a binary outcome (in this case, a banking crisis) using various evaluation metrics. The algorithms include logistic regression (LOGIT), k-nearest neighbors (KNN), XGBoost, support vector machines (SVM), and neural network (NEN).

The evaluation metrics used in the table include F1 score, AUC (area under the receiver operating characteristic curve), and accuracy.

The F1 score is a harmonic mean of precision and recall, and it measures the balance between the true positive rate and the false positive rate. A value of 1 represents perfect precision and recall, while a value of 0 represents the worst possible score. AUC measures the overall performance of the classifier, and it ranges from 0.5 (random guessing) to 1 (perfect classification). Accuracy measures the proportion of correctly classified instances out of the total number of instances. Looking at the results in the table, we can see that some algorithms perform better than others in predicting the binary outcome. For example, XGBoost has the highest F1 score, AUC, and accuracy, indicating that it performs better than other algorithms in predicting the outcome of a crisis. On the other hand, some algorithms, such as SVM and NEN, have low scores across all evaluation metrics, indicating that they perform poorly in predicting the outcome.

In addition, and as a robustness check of our models, we can conclude that the best performing models are indeed the XGBoost and the KNN classifiers since they are the only models to have an F1-score which is greater than 0 across all datasets. This means that these models were able to classify crises occurrences accurately, to a certain extent.

The Receiver Operating Characteristic (ROC) curve is a graphical representation of the performance of a binary classifier system that compares the trade-off between true positive rate (sensitivity) and false positive rate (1-specificity) at different threshold settings. The ROC curve plots the true positive rate on the y-axis and the false positive rate on the x-axis. The Area Under the Curve (AUC) is a scalar metric that summarizes the ROC curve's overall performance. It measures the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example. The AUC ranges between 0 and 1, where a value of 0.5 represents a random classifier, and a value of 1 indicates a perfect classifier that correctly classifies all examples. Generally, an AUC score above 0.8 is considered to be a good performance for a classifier system, while an AUC score below 0.5 indicates a classifier system that performs worse than random guessing. If the ROC curve goes along the diagonal, the AUC would be 0.5 and it would not be possible to derive insights from the prediction. As ROC curves move closer to the upper left corner (i.e., with an increasing area under the curve AUC), the prediction becomes more insightful. For ROC curves below the diagonal, the prediction is considered to be performing worse than random guessing.

Specificity, sensitivity, false positive rate, false negative rate, true positive rate, true negative rate, and the F-1 score are all metrics used to evaluate the performance of a binary classification model, i.e., a model that predicts the outcome of a binary event (in our case being the occurrence of a crisis).

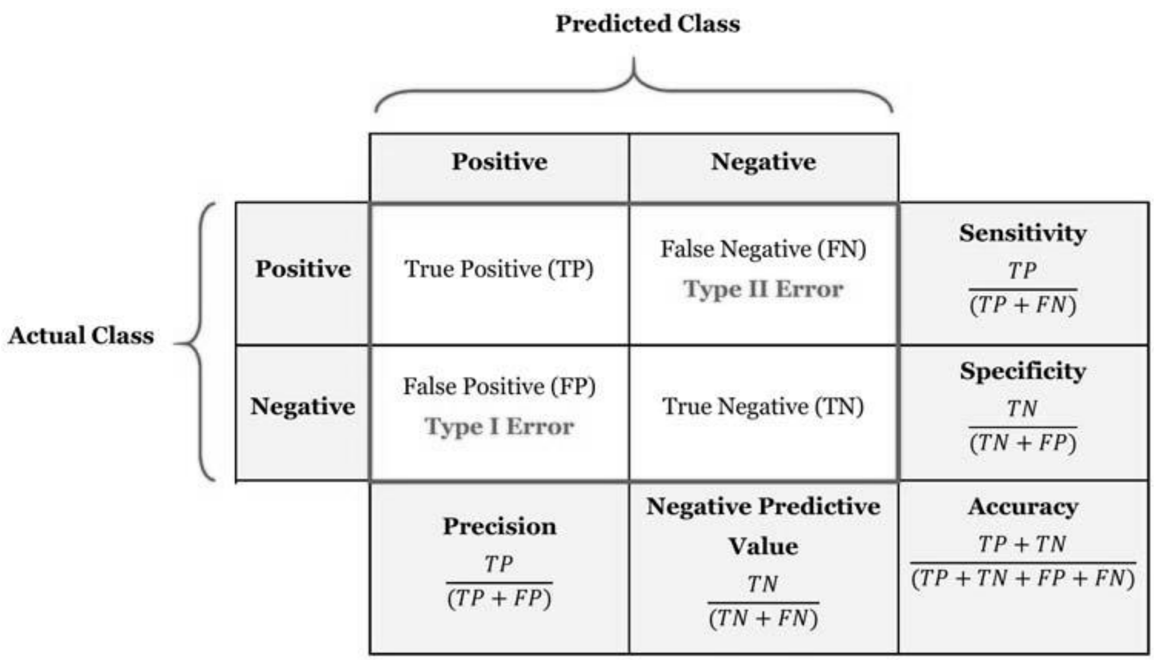


Figure 6: Confusion Matrix

Sensitivity is the proportion of true positive results that are correctly identified by the model, i.e., the model's ability to correctly identify positive cases. It is calculated as true positives divided by the sum of true positives and false negatives. Sensitivity ranges from 0 to 1, with higher values indicating better performance.

Specificity is the proportion of true negative results that are correctly identified by the model, i.e., the model's ability to correctly identify negative cases. It is calculated as true negatives divided by the sum of true negatives and false positives. Specificity ranges from 0 to 1, with higher values indicating better performance.

False positive rate (FPR) is the proportion of negative cases that are incorrectly identified as positive by the model, i.e., the model's tendency to generate false positives. It is calculated as false positives divided by the sum of false positives and true negatives. FPR ranges from 0 to 1, with lower values indicating better performance. False negative rate (FNR) is the proportion of positive cases that are incorrectly identified as negative by the model, i.e., the model's tendency to generate false negatives. It is calculated as false negatives divided by the sum of false negatives and true positives. FNR ranges from 0 to 1, with lower values indicating better performance.

True positive rate (TPR), also known as sensitivity or recall, is the proportion of true positive results that are correctly identified by the model, i.e., the model's ability to correctly identify positive cases. It is calculated as true positives divided by the sum of true positives and false negatives. TPR ranges from 0 to 1, with higher values indicating better performance.

True negative rate (TNR), also known as specificity, is the proportion of true negative results that are correctly identified by the model, i.e., the model's ability to correctly identify negative cases. It is calculated as true negatives divided by the sum of true negatives and false positives. TNR ranges from 0 to 1, with higher values indicating better performance.

F-1 score is a weighted average of precision and recall (sensitivity) and provides a single value to summarize the model's performance. It ranges from 0 to 1, with higher values indicating better performance.

F-1 score is calculated as follows:

*2 * (precision * recall) / (precision + recall),*

where precision is the proportion of true positives among all positive predictions, and recall (sensitivity) is the proportion of true positives among all actual positive cases.

$$F_1 = \frac{TP}{TP + \frac{(FP+FN)}{2}}$$

Equation 6: F-1 Score Equation

In general, a good model should have high sensitivity and specificity, low false positive and false negative rates, and a high F-1 score. However, the relative importance of these metrics depends on the specific context and the costs and benefits associated with each type of error. In our work, as much as it is important that models yield high performance, it is also very important to look at the recall value, as we are interested in accurately classifying and predicting the occurrence of crises, which is a rare event.

This paper builds on the research conducted by Beutel et al. (2019) that has been motivated by the 2008 global financial crisis. In their research, Beutel et al. (2019) present the work done by prominent researchers in the field such as Alessi and Detken (2011), Lo Duca and Peltonen (2013), and Drehmann and Juselius (2014) on early warning models. In fact, Alessi and Detken (2018) introduced Random Forests in early warning models related to systemic banking crises on a country level whereas Tanaka et al. (2016, 2018) employed them to predict individual banks' failures. Predictive power of models is then evaluated using cross-validation techniques. Beutel et al. (2019) discuss that their findings could be justified by comparing results to results of other studies that employ machine learning techniques to predict financial crises. As previously mentioned, machine learning methods usually allow for the employment of a larger number of parameters than statistical models and consequently allow for more flexibility in estimation. This flexibility might return better in-sample results, yet, it also holds the risk of overfitting and so hinders the performance on out-of-sample data. Beutel et al (2019) discuss that cross-validation might show interesting results as compared to the conventional out-of-sample evaluation, yet, it has its weaknesses: as discussed by Holopainen and Sarlin (2017), cross-validation might return inflated performance metrics and that favor machine learning models over statistical models.

Our findings reveal, based on the data available, that machine learning models do in fact outperform the logit model in classifying crisis occurrences. More specifically the tree-based model XGBoost achieved near perfect results. However, looking at the different performance measures is important in assessing the performance of each model, specifically when attempting to predict rare events.

A model could have high accuracy and low F-1 score while predicting all occurrences as a non-crisis event as the frequency of the crisis occurrence is low when compared to the volume of the data. To test the performance of the models, we select the best performing models: KNNs and XGBoost. We exclude the data related to Lebanon and keep it as a test set (out-of-sample set). We train the models and validate the results on the data related to the remaining countries in our database and introduce the data for Lebanon as unseen data. We find that the KNN model achieves near perfect results on the in-sample data but yields an accuracy of 68% with an AUC of 0.95 and an F-1 score of 0. Similarly, the XGBoost model also achieves near perfect results on the in-sample data but yields an accuracy of 77.3%, an AUC of 0.18 and an F-1 score of 0. These results illustrate that machine learning algorithms are prone to overfitting when it comes to out-of-sample data.

In-sample results for the best performing models

Results	**F1**	**AUC**	**Accuracy**
KNNs	0.64	0.9	0.96
XGBoost	0.86	0.987	0.98

Table 10: In-Sample Results

Out-of-sample results for the best performing models

Results	**F1**	**AUC**	**Accuracy**
KNNs	0	0.95	0.68
XGBoost	0	0.18	0.77

Table 11: Out-of-Sample Results

Further, we aim to highlight the economic intuition driving the predictions, and more specifically the logit model, as discussed by Beutel et al. (2019). This model, despite having lower predictive power than other models, was able to provide insights on the most important variables: Credit to Private Sector, Debt Service Ratio, Narrow Money Supply (M1), Real Public Debt, Real Broad Money Supply, Real Net Foreign Assets, Real Effective Exchange Rate, Nominal Effective Exchange Rate and Oil Prices. Findings are consistent with existing theoretical and empirical evidence, which attributes the weakness of the banking sector to many of these different factors.

We also find that the majority of coefficients have the expected sign which is explained in theory and evidence: for example, increase in credit growth increases the probability a crisis occurrence. Conversely, an increase in the Debt Service Ratio is negatively associated with a crisis occurrence which means that so long that debt is being serviced, the probability of a crisis occurrence is reduced.

Finally, we attempt look further into the classification of crises occurrences. Therefore, we develop a K-means model which clusters crises occurrences based on the following attributes: Credit to Private Sector, Debt Service Ratio and Nonperforming Loans. Then, we attempt to classify the case of Lebanon based on the resulting model.

K-means clustering is a popular unsupervised machine learning algorithm that is used to group similar data points into clusters based on their features or characteristics. It aims to partition a given dataset into K clusters, where K is a pre-defined number of clusters.

The algorithm works by first randomly initializing K cluster centers, and then iteratively assigning each data point to the nearest cluster center and updating the cluster centers based on the mean of the data points assigned to them. This process is repeated until the cluster centers no longer move significantly or a maximum number of iterations is reached.

One key challenge in K-means clustering is determining the optimal number of clusters. There are several methods to estimate the optimal number of clusters, such as the elbow method or the silhouette score. K-means clustering is widely used in various applications, such as market segmentation, customer segmentation, anomaly detection (in this case a crisis can be considered as an anomaly), and bioinformatics.

Our findings reveal that Lebanon falls in the cluster of the following countries: Austria, Canada, Chile, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, and the United Kingdom. As the clusters were segmented based on Credit Quality indicators (Credit to Private Sector, Debt Service Ratio and Nonperforming loans), Lebanon's credit level and quality appears to be at par with the aforementioned countries within the same cluster, which did not suffer from banking crises in the timeframe 2019-onwards. Results might be more insightful had data related to countries which share similar characteristics been available. Also, further segmentation could be conducted using different features such as public debt levels, money supply levels and changes in interest rate levels.

CHAPTER VI

CONCLUSION

This paper has presented an analysis of prediction of systemic banking crises. Research was done on a dataset covering 40 different countries, including advanced and developing economies, over the period 2010–2021. We assess and compare the predictive performance of different machine learning methods against the benchmark, being the logit approach, on out-of-sample data. Findings showed that the logit approach does not seem to outperform all machine learning methods based on different performance metrics and using different datasets and variables. Yet, it showed to be robust to the choice of crisis variable and data transformation. Machine learning techniques enabled us to properly classify crises occurrences while the logit model enabled us to identify the most significant predictors of banking crises. Our findings showed more stability in the performance of machine learning methods across the different variations in data as compared to the logit model. As discussed by Beutel et al. (2019), these results raise doubts in the performance of machine learning methods in "real-world out-of-sample prediction situations" (Beutel et al. (2019)). Actually, Beutel et al. (2019) claim that potential modifications for improving machine learning stability and performance in early warning applications on out-of-sample data are yet to be explored.

The limitations that are to be addressed are mainly data availability, integrity, and quality. In fact, data availability and data quality are crucial factors that can impact the performance of machine learning models.

Insufficient data availability can lead to models with poor accuracy and reliability, as the lack of data can result in underfitting, where the model fails to capture the complexity of the underlying patterns in the data. On the other hand, a large amount of low-quality data can lead to overfitting, where the model fits too closely to the noise in the data and fails to generalize well to new data. Data quality issues, such as missing data, outliers, and data errors, can also negatively impact the performance of machine learning models. For example, missing data can lead to biased or incomplete analysis, while outliers can skew the results and lead to inaccurate predictions. Data errors can also result in incorrect modeling assumptions and can negatively affect the reliability and validity of the model. Therefore, it is crucial to ensure that data is both available and of high quality for machine learning models to perform effectively. This involves properly collecting, cleaning, and preprocessing the data to ensure that it is representative, accurate, and complete.

In the case of Lebanon, future research could investigate the impact of political and governance variables on financial stability. Depositor Trust and the Involvement of Politically Exposed Persons in the Banking Sector would be interesting to examine, as part of a larger sample of similar countries (similar in terms of economic development and economic / political characteristics).

In discussions on machine learning and predicting financial crises, it is worth examining crises using both Network models and Early Warning Models. In fact, the recent financial turmoil that shook the world in 2008 ascertained how intertwined financial systems are. In such interconnected markets, banks and financial institutions are also linked in many ways (borrowing, lending, holding deposits, etc…). Consequently, they are subject to contagion risk (Babus, 2016). Samitas et al. (2020) study the detection of potential contagion risks in financial networks using network analysis and machine learning algorithms. With increasing openness in markets, and specifically in financial markets, banking systems are more and more interconnected, linkages between banks carry the risk of contagion (Babus, 2016). Minoiu et al. (2015) explored the prediction of systemic banking crises through the connectedness of global financial networks. These studies support the theories mentioned by Laeven (2011) that systemic banking crises could be a result of contagion. As such, future research could explore employing both Network and Early Warning models to detect potential crises (based on Early Warning Models) and their potential spillover (based on Network Models), providing policymakers with data driven tools for economic and financial policy related decision making.

APPENDIX A

A. List of Countries in our Database

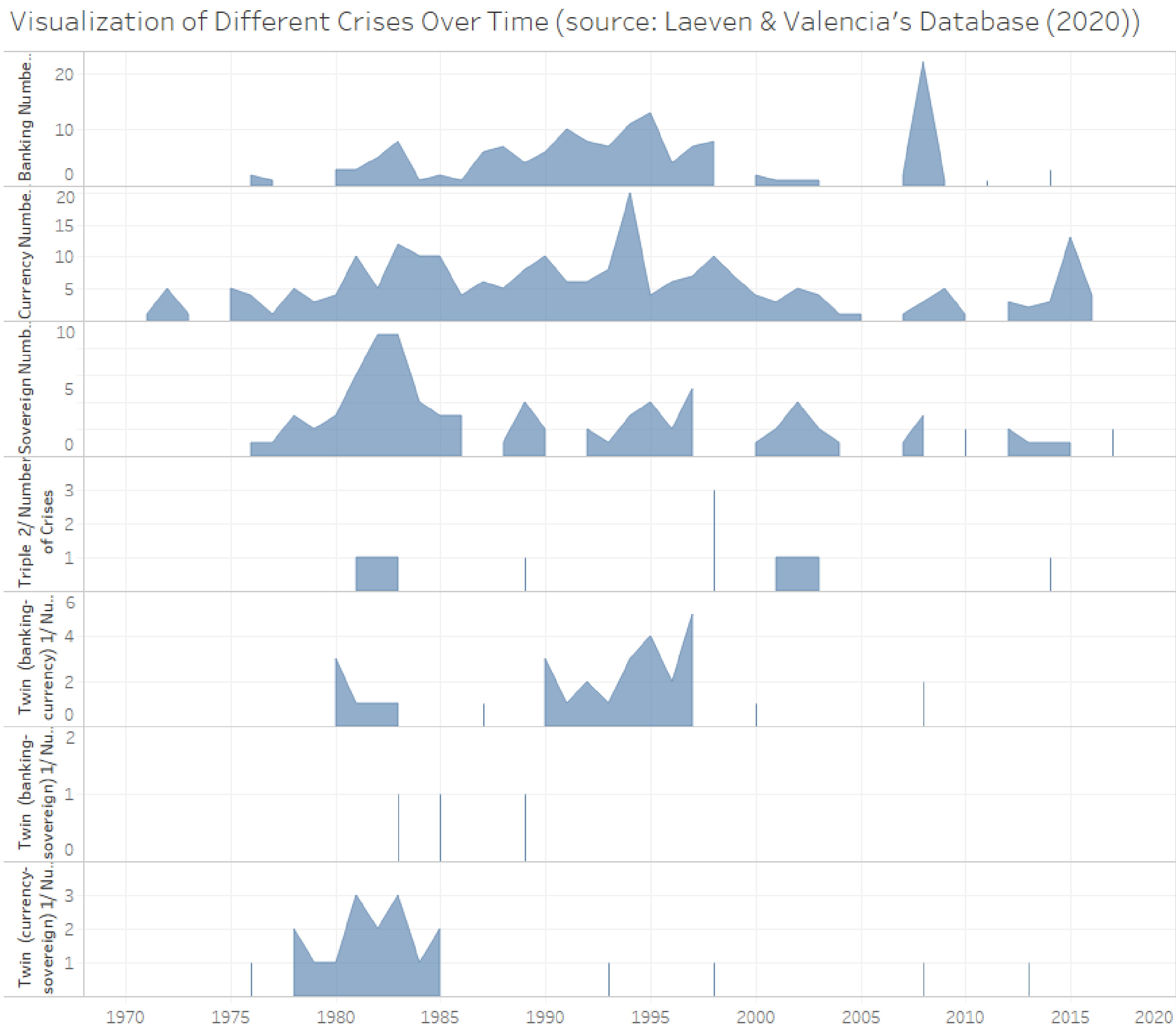
COUNTRIES	
ARGENTINA	ITALY
AUSTRALIA	JAPAN
AUSTRIA	KOREA
BELGIUM	LEBANON
BRAZIL	LUXEMBOURG
CANADA	MALAYSIA
CHILE	MEXICO
COLOMBIA	NETHERLANDS
CZECH REP.	NORWAY
DENMARK	POLAND
FINLAND	PORTUGAL
FRANCE	RUSSIA
GERMANY	SPAIN
GREECE	SWEDEN
HUNGARY	SWITZERLAND
INDIA	THAILAND
INDONESIA	TURKEY
IRELAND	UNITED KINGDOM

B. VARIABLES AND SOURCES

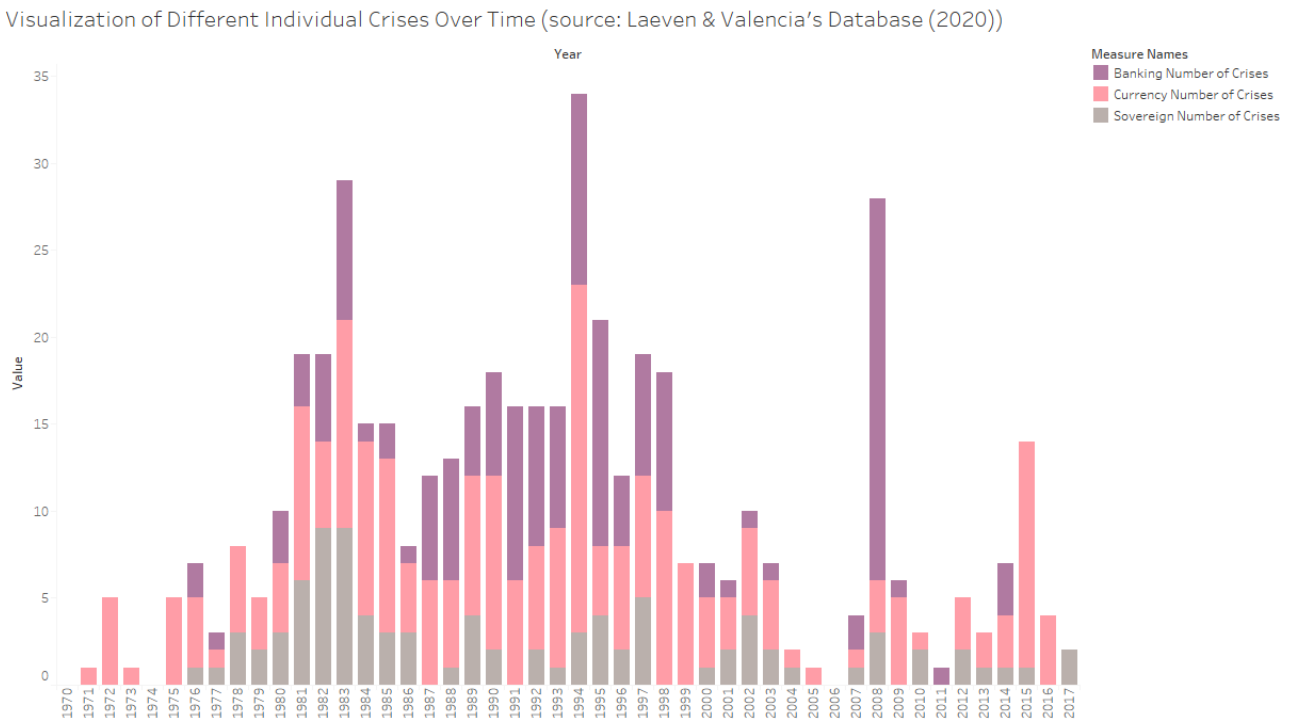
VARIABLE	SOURCE
3-MONTHS INTEREST RATE	FRED
BANKING CRISIS	Laeven & Valencia's Database
CONSUMER PRICE INDEX	FRED, Australian Government (Taxation Office), Bank of Canada, Nasdaq Data, The Global Economy
CREDIT TO PRIVATE SECTOR	BIS, BDL, BRITE
DEBT SERVICE RATIO	BIS, FRED, Our World in Data
DEMAND DEPOSITS	FRED, IMF IFS, BDL, BRITE, OECD
MONEY SUPPLY (BROAD MONEY)	FRED, Central bank of Argentina, Malaysia Central Bank, OECD
NET FOREIGN ASSETS	FRED, BIS, BDL, BRITE, IMF IFS
NONPERFORMING LOANS	BIS, BDL, BRITE, Statista
PUBLIC DEBT	FRED, IMF IFS, BDL, BRITE, UK Central Bank, National Bank of Belgium
REAL EFFECTIVE EXCHANGE RATE	Datastream, BIS, Central Bank of Argentina
NOMINAL EFFECTIVE EXCHANGE RATE	Datastream, BIS, Central Bank of Argentina
OIL PRICES	FRED

APPENDIX B

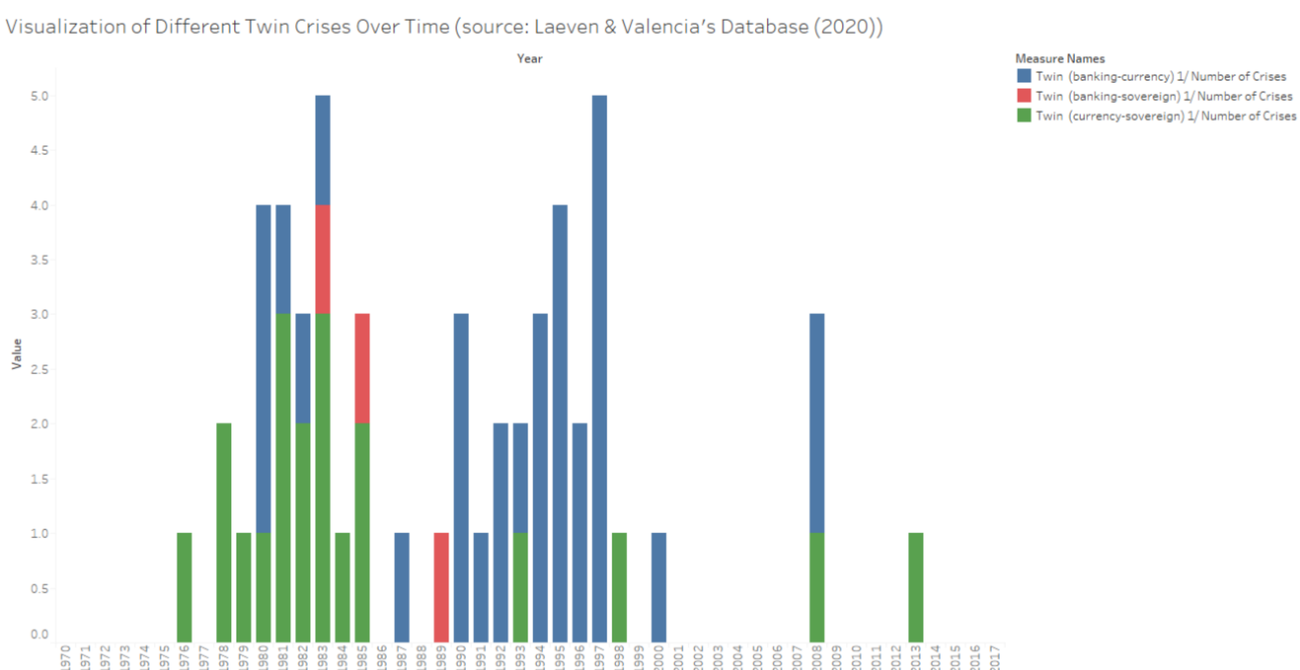
Diagram Visualizing Data Collection



Visualization of Different Crises Over Time



Visualization of Individual Crisis Occurrences Over Time



Visualization of Twin Crises Occurrences Over Time

Visualizing Laeven & Valencia's database, we find that it is common for multiple crises to occur at the same time. Looking at the peaks in the bar chart, we note 3 periods where the occurrence of crises was the highest:

1980 – 1983:

Several crises occurred worldwide between 1980 and 1985. Some of the notable ones are:

Latin American Debt Crisis (1982): This crisis started when Mexico announced that it could no longer service its foreign debt. It soon spread to other Latin American countries, as investors became concerned about their ability to repay their loans.

Savings and Loan Crisis (US, 1980s): This crisis was caused by a combination of factors, including risky lending practices, poor regulation, and economic recession. It led to the failure of many savings and loan institutions in the US and cost taxpayers billions of dollars.

Third World Debt Crisis (1980s): This crisis was characterized by the inability of many developing countries to service their foreign debt. It was caused by a combination of factors, including high interest rates, low commodity prices, and poor economic policies.

Australian Banking Crisis (1983): This crisis was caused by the collapse of several major banks in Australia, which had engaged in risky lending practices. The crisis led to the nationalization of some of these banks and a period of economic recession in Australia.

Oil Glut (1980s): This crisis was caused by an oversupply of oil on the global market, which led to a sharp decline in oil prices. This had a significant impact on oil-exporting countries, many of which experienced economic recession and financial crisis.

Early to mid 90s:

Many crises occurred worldwide in the 1990s but some of the most notable ones are:

Nordic banking crisis (1987-1993): This was a series of banking crises that occurred in the Nordic countries of Sweden, Norway, Finland, and Denmark. The crisis was triggered by a combination of factors (mainly: deregulation of the banking sector, real estate market speculation, and economic recession).

Japanese asset price bubble (1986-1991): This was a speculative bubble in Japan's real estate and stock markets that peaked in 1991. When the bubble burst, it caused a prolonged period of economic stagnation in Japan known as the "Lost Decade".

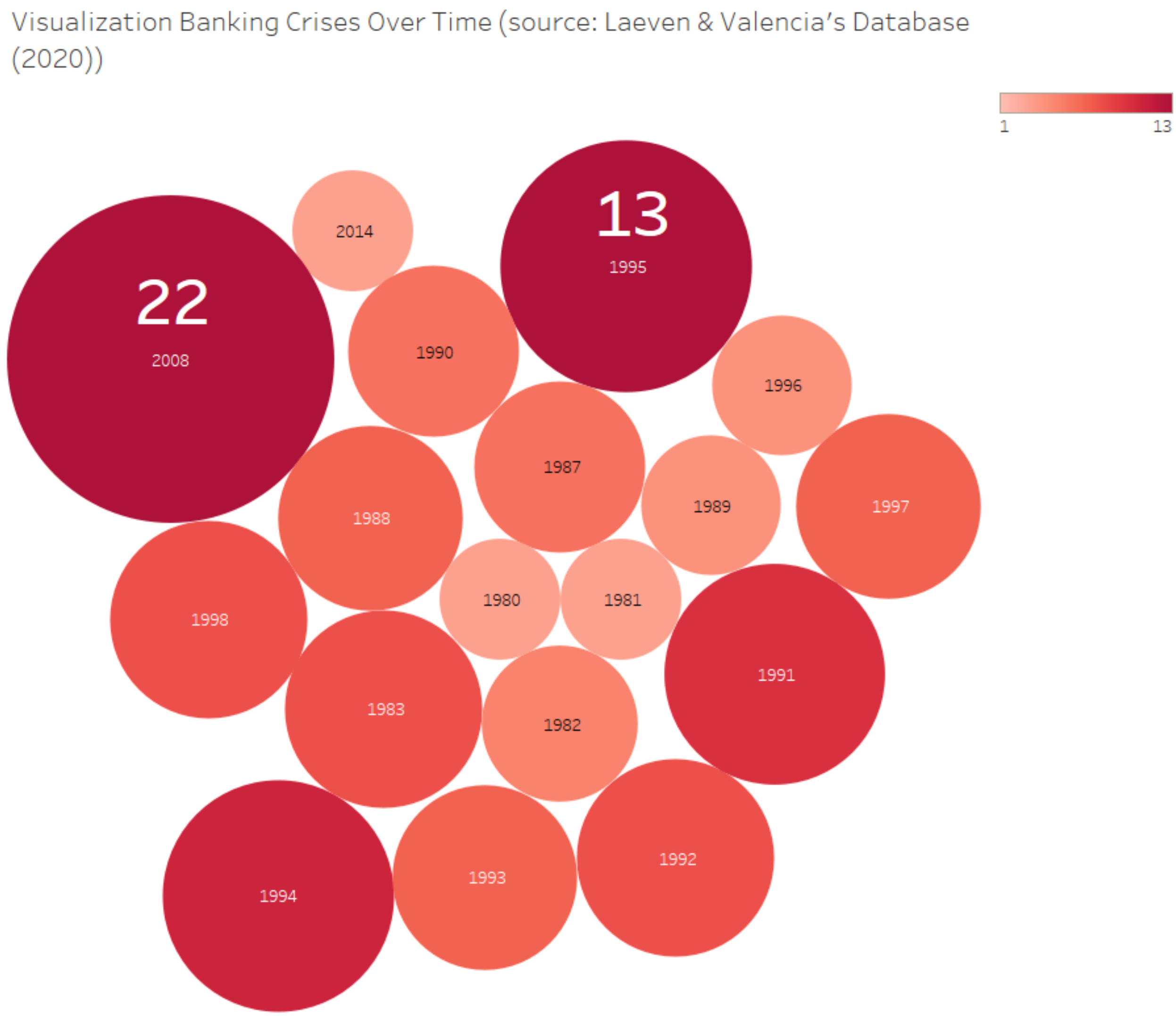
Mexican peso or Tequila crisis (1994-1995): This was a currency crisis that occurred in Mexico when the Mexican government devalued the peso. The crisis was caused by a combination of factors, including large fiscal deficits, political instability, and a decrease in foreign investment.

2009 (2007-2009)

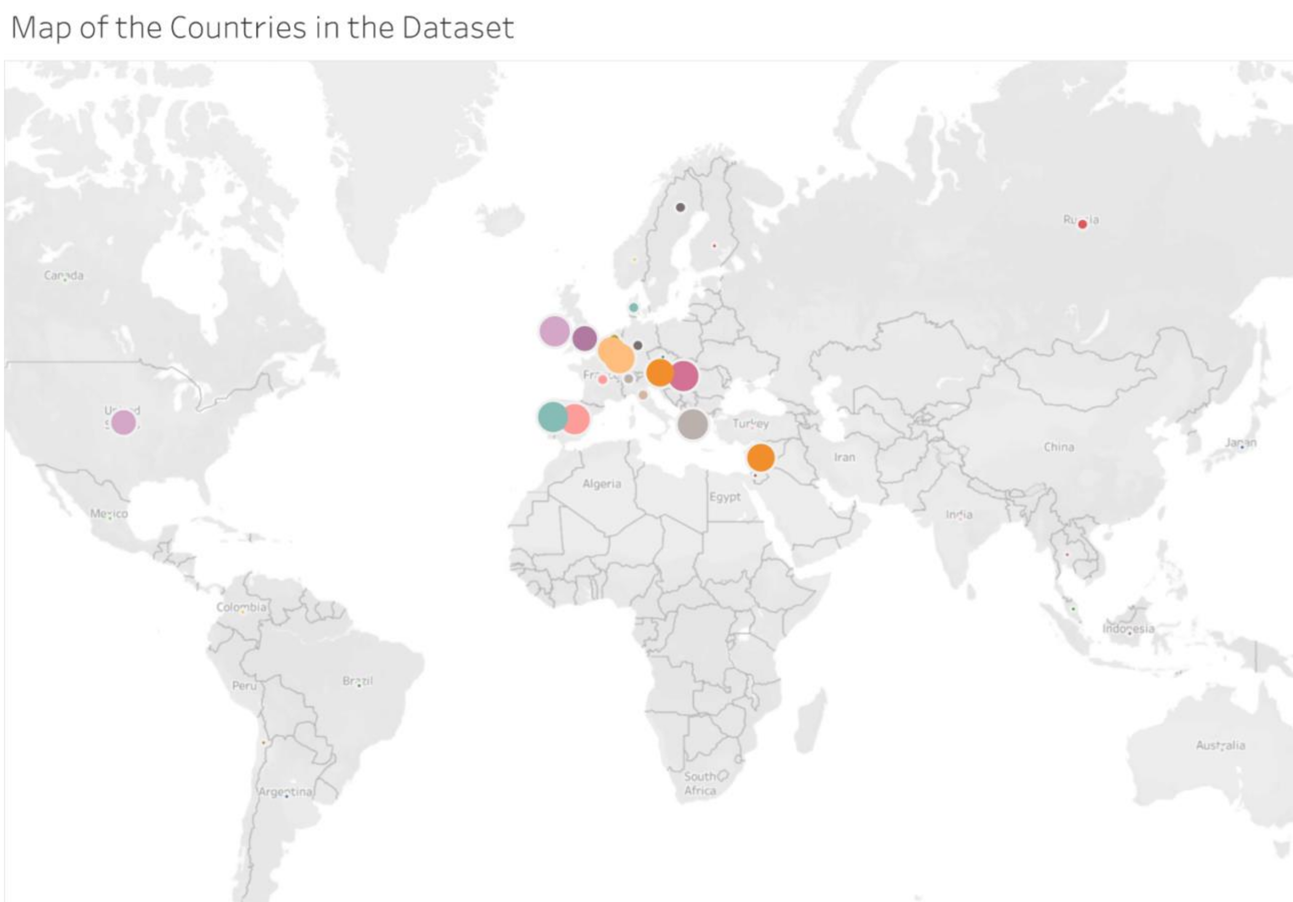
The financial crisis of 2007-2009, also known as the Great Recession, was a global economic downturn that began in the United States and quickly spread to other countries. The crisis was triggered by a number of factors, including the collapse of the US housing market, risky lending practices by banks, and excessive risk-taking by investors. In the early 2000s, US banks began issuing mortgages to borrowers with poor credit histories, known as subprime mortgages. These mortgages were bundled together into complex financial instruments called mortgage-backed securities (MBS) and sold to investors around the world. As housing prices began to fall in 2006, many borrowers defaulted on their mortgages, causing the value of MBS to plummet. Banks and other financial institutions that had invested heavily in MBS suddenly found themselves facing enormous losses. This sparked a wave of bank failures and government bailouts. The crisis quickly spread to other sectors of the economy, causing a global recession. The crisis also revealed weaknesses in the regulatory system and led to widespread calls for reform. Governments around the world implemented a range of measures to stabilize the financial system and prevent similar crises from happening in the future. These measures included tighter regulation of banks and financial markets, increased transparency, and stronger consumer protections.

Data Sample

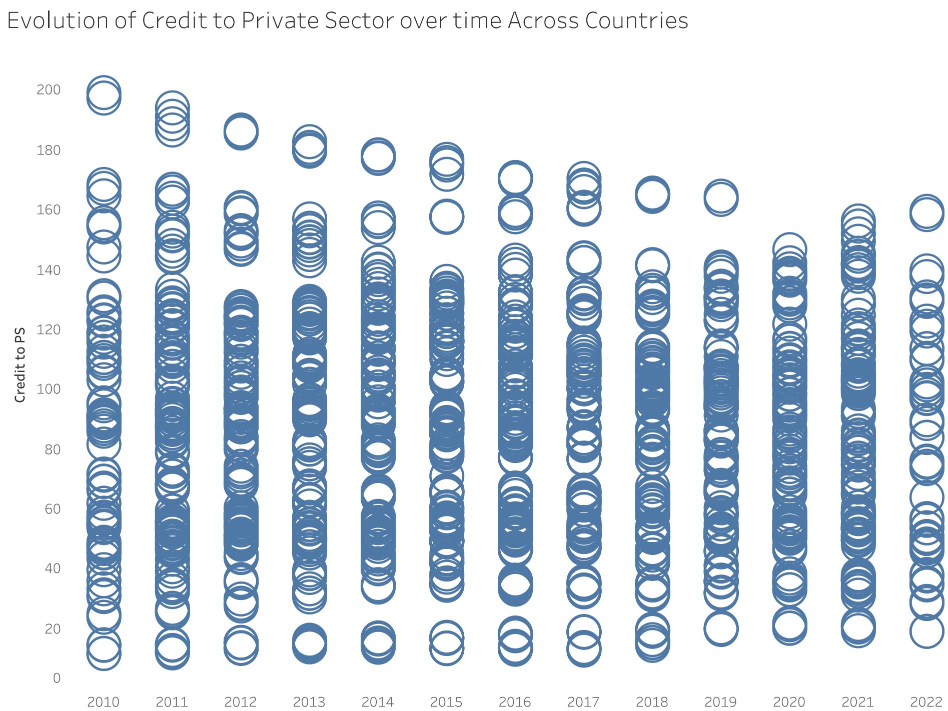
index	Credit to PS	DSR	NPL	M1	Public Debt	M3	NFA	REER	NEER	Oil Prices	3M INTEREST	CPI
std	38	9.4	7.2	60	857839199939	7662637374883	302849744196	10.6	16	26	4.2	578
min	12	-21	0.0	24.0	0.0	32965910	-39366451	16.0	22.0	31.0	-1.0	59.0
mean	86	12.5	5.15	112	206927702456	2374416052390	38643624995	95.24	96	76	2.5	200
max	199	31	52	697.0	9575843584106	54959024152180	6398042209704	126.0	198.0	119.0	40.0	5099
75%	112	19	5	129.0	5841678404.	137532200395	1003376299	101.0	103.0	108.0	3.0	114
50%	87	14	3	102.0	740574526	15829316565	239741791	98.0	99.0	73.0	1.0	101
25%	54	5	2	79.0	64199049	3532886380	10360301	91.0	92.0	52.0	0.0	98



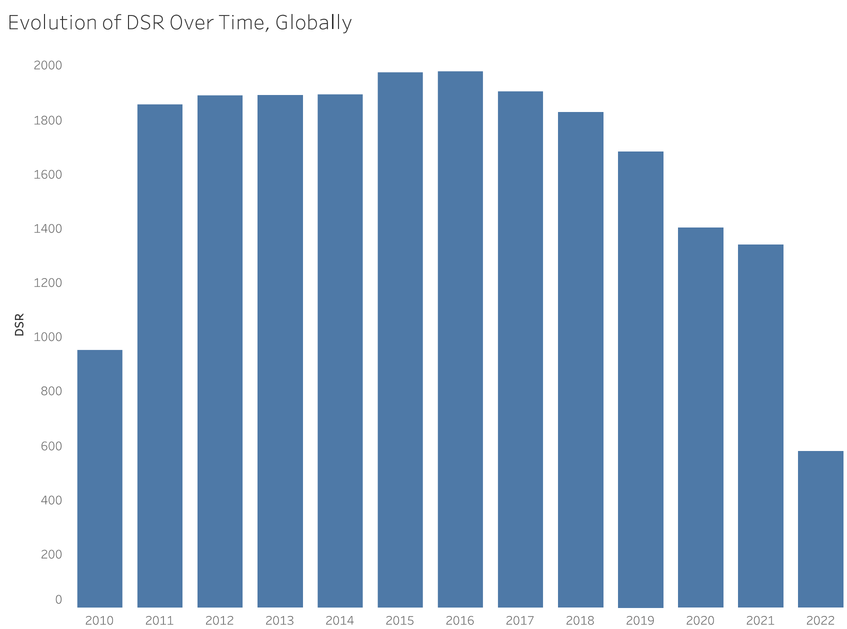
Bubble Chart Visualizing the Years with the Highest Count of Crises Occurrence



Cross-Country Crisis Occurrence



Evolution of Credit to Private Sector Over Time and Across Countries

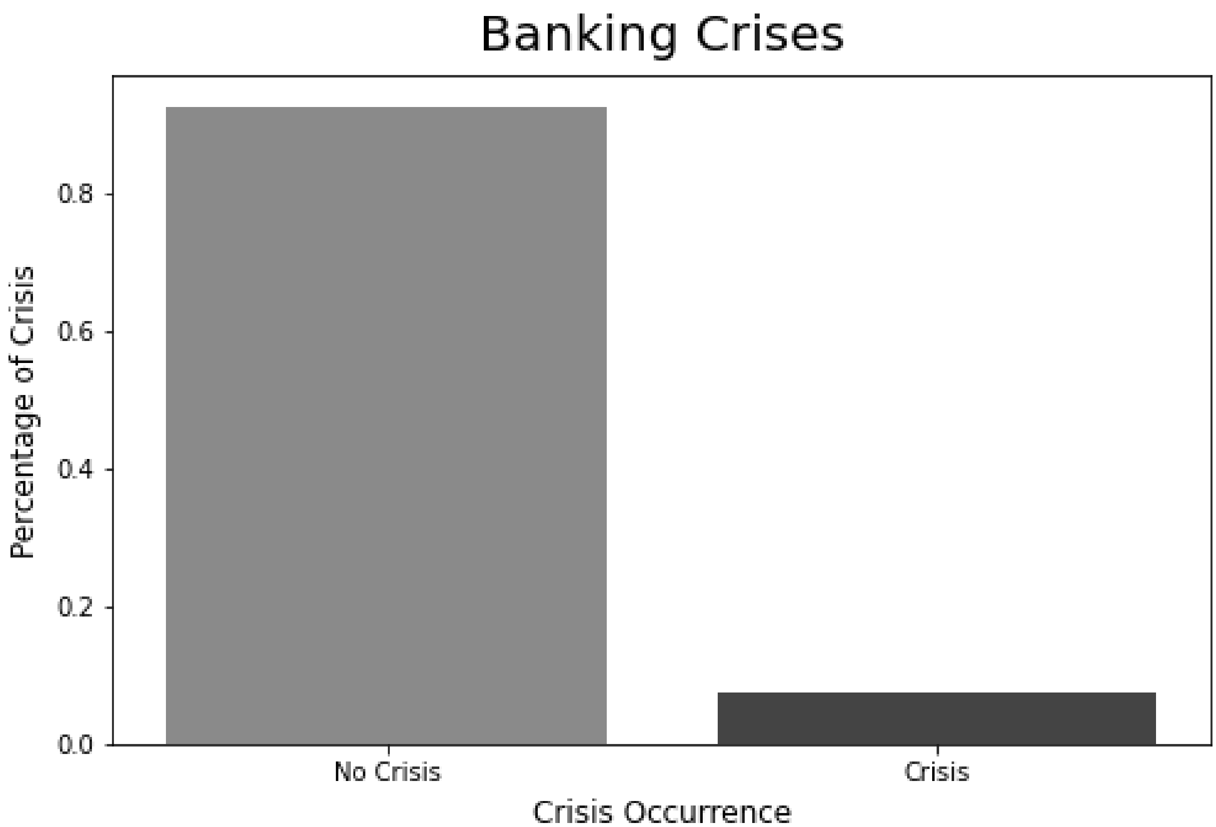


Evolution of Debt Service Ratio Over Time, Globally

Average NPL Per Country

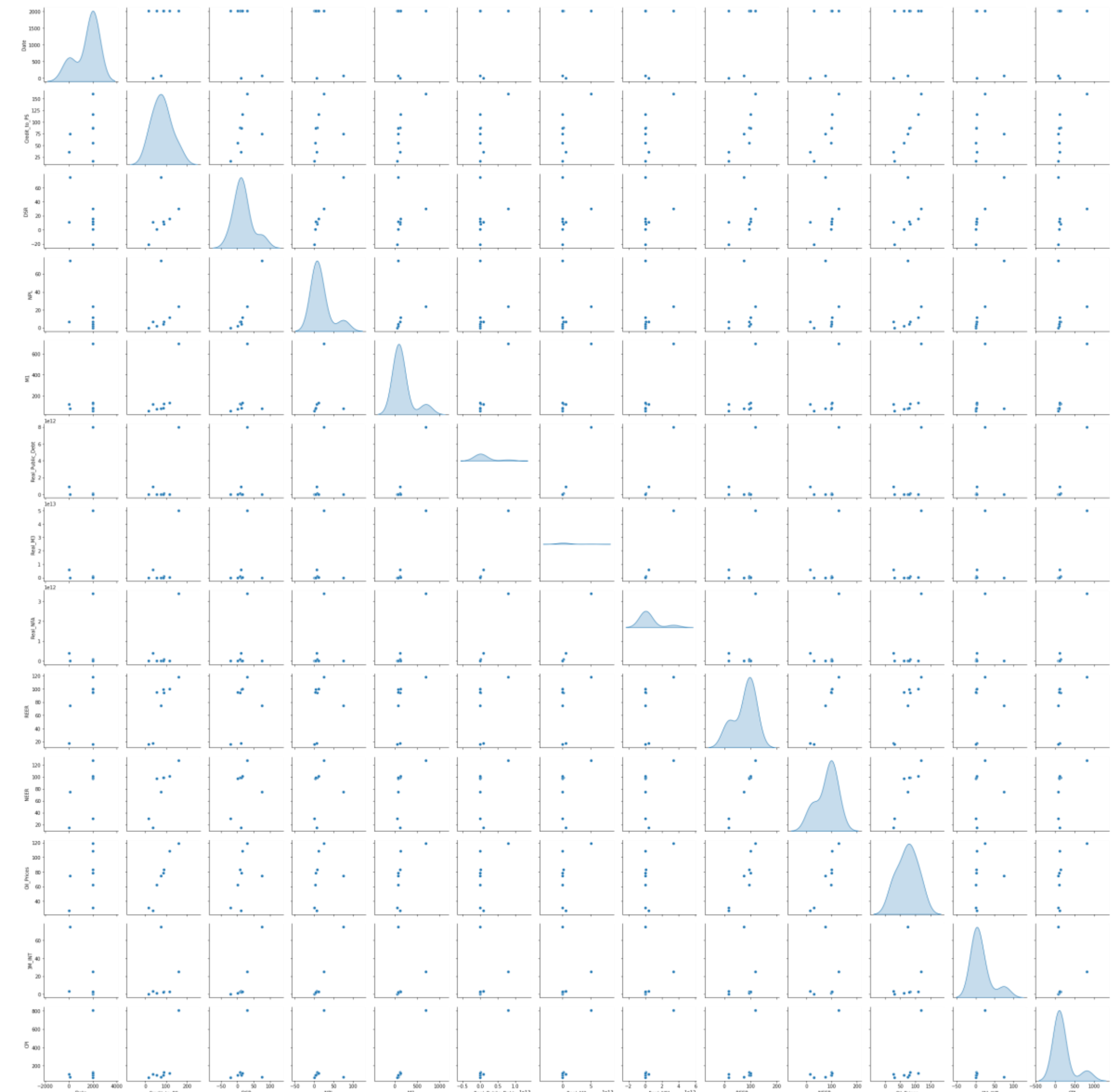
Average NPL Per Country

Average NEER & REER



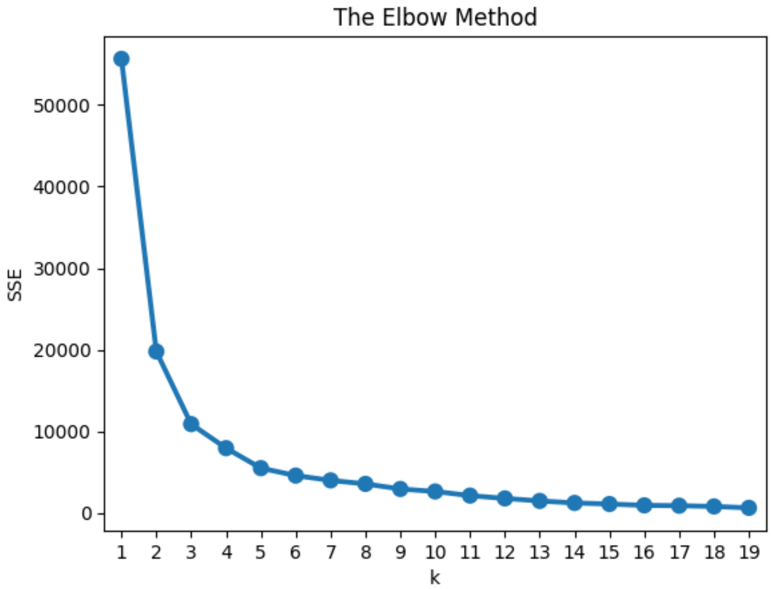
Percentage of Crises Occurrences

As an extraordinary event, it is expected to find that the number of crises occurrences in our dataset be smaller than the number of tranquil periods. In fact, the variable crisis = 1 represents only 7% of the total variable crisis occurrences.

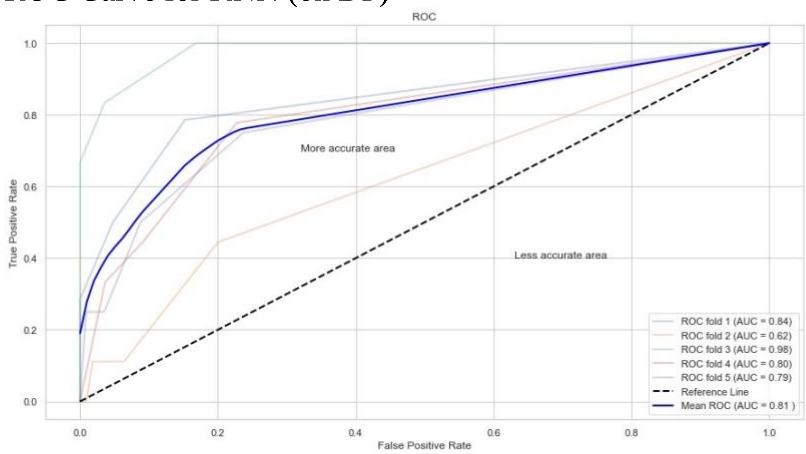


Visualizing Correlations between Numerical Variables

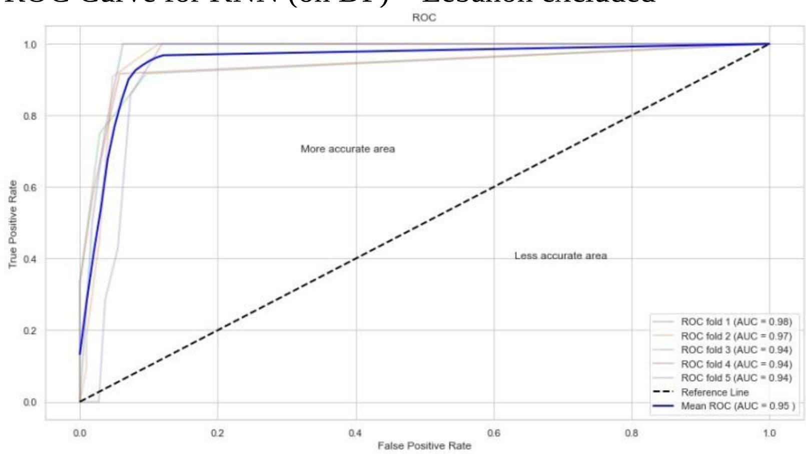
The Elbow Method Curve



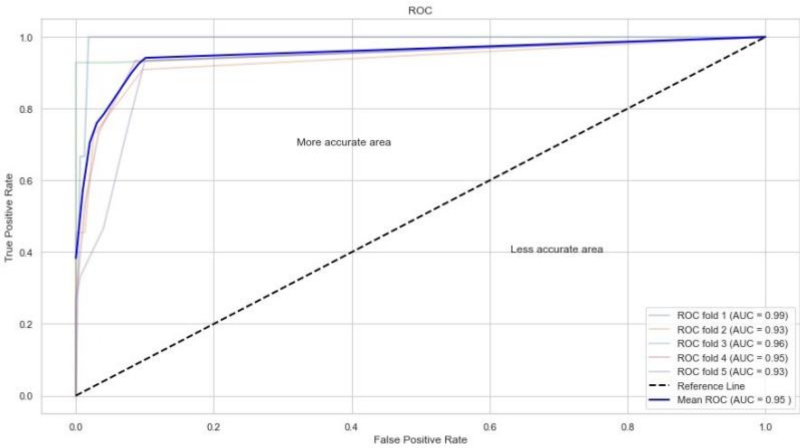
ROC Curve for KNN (on DF)



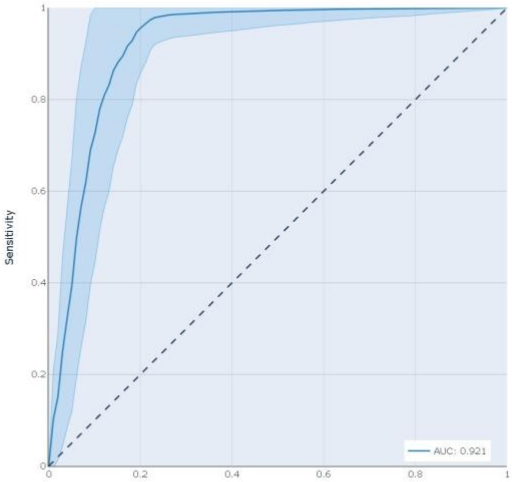
ROC Curve for KNN (on DF) – Lebanon excluded



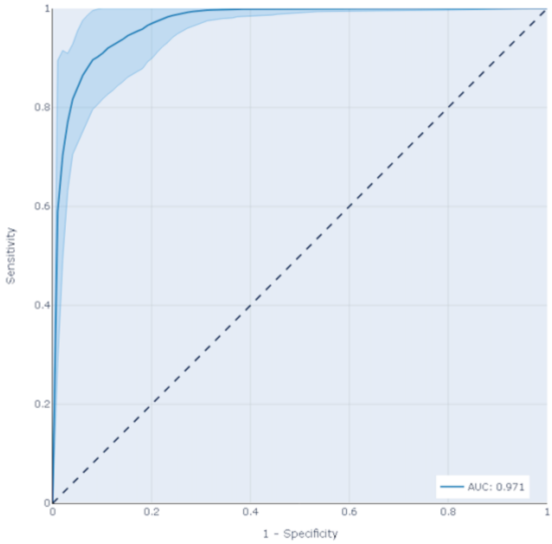
ROC Curve for KNN – Testing on Lebanon


ROC
More accurate area
Less accurate area
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC fold 1 (AUC = 0.99)
ROC fold 2 (AUC = 0.93)
ROC fold 3 (AUC = 0.96)
ROC fold 4 (AUC = 0.95)
ROC fold 5 (AUC = 0.93)
Reference Line
Mean ROC (AUC = 0.95)

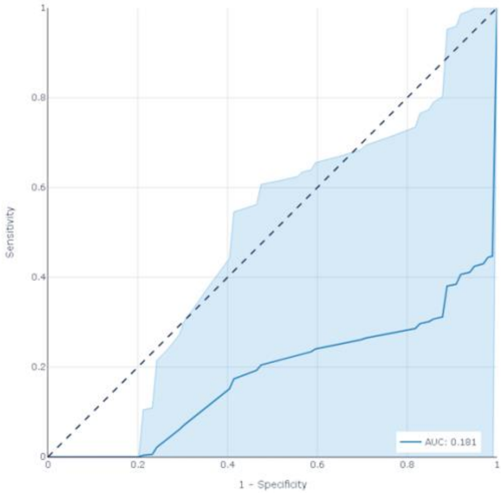
ROC Curve for XGBoost (on DF)



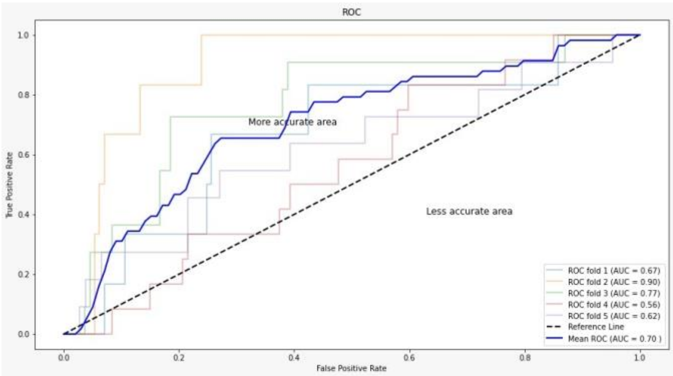
ROC Curve for XGBoost (on DF) – Lebanon excluded



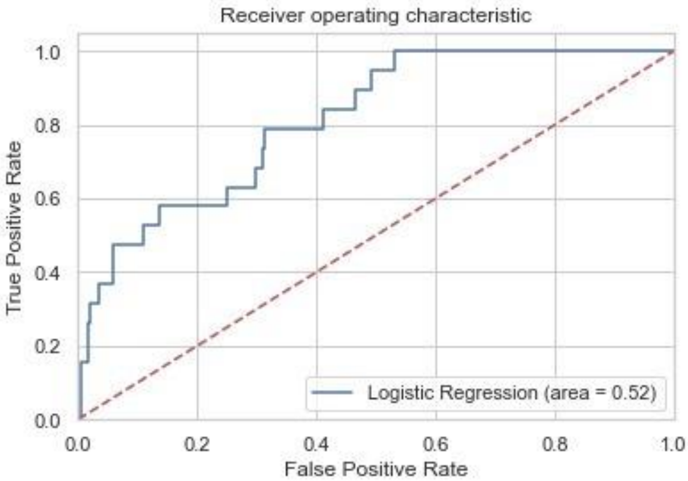
ROC Curve for XGBoost – testing on Lebanon



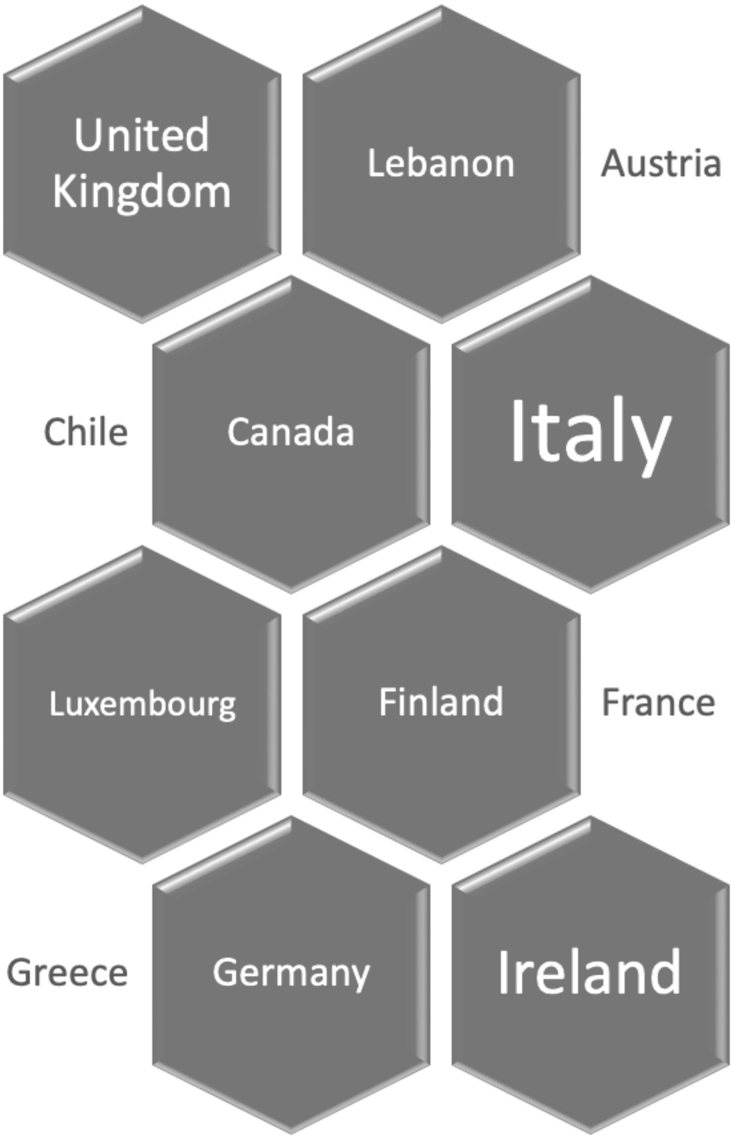
ROC Curve for SVM (on DF)



ROC Curve for LOGIT (on DF)



K-means clustering



Coding Language

Python has a rich ecosystem of libraries and frameworks that can be used for various purposes. Here are some of the most commonly used Python libraries:

NumPy: A library for scientific computing in Python that provides support for large, multi-dimensional arrays and matrices.

Pandas: A library for data manipulation and analysis, providing data structures for working with tabular data, time series data, and more.

Matplotlib: A library for creating static, animated, and interactive visualizations in Python.

Scikit-learn: A library for machine learning in Python that provides algorithms for classification, regression, clustering, and more.

TensorFlow: An open-source platform for building machine learning models that supports both deep learning and traditional machine learning approaches.

Keras: A high-level neural networks API, written in Python and capable of running on top of TensorFlow, Theano, or Microsoft Cognitive Toolkit.

PyTorch: An open-source machine learning library for Python, based on Torch, which is used for applications such as computer vision and natural language processing.

Django: A web framework for Python that provides a set of tools and libraries for building high-performance web applications.

Flask: A micro web framework for Python that is lightweight and easy to use.

Requests: A library for making HTTP requests in Python, allowing you to interact with APIs and web services.

These libraries are just a few examples of the many available in the Python ecosystem. They can help simplify complex tasks, reduce the amount of code needed, and enable developers to create applications and solutions more efficiently.

In our work, we mainly used the following Python Libraries: NumPy, Pandas, Matplotlib, Scikit-learn and Keras.

REFERENCES

1. “Has Financial Development Made the World Riskier?” 2022. ResearchGate.
2. “Keynesian Approaches to Financial Crisis.” 2022. ResearchGate. 2022.
3. “Part I of Crisis in Lebanon: Economic ‘Free Fall,’ IMF Negotiations, and Beirut Explosion.” 2021. Yale.edu. Yale School of Management. December 21, 2021. https://som.yale.edu/blog/part-i-of-crisis-in-lebanon-economic-free-fall-imf-negotiations-and-beirut-explosion.
4. “Part II of Crisis in Lebanon: Buildup of Interrelated Challenges.” 2021. Yale.edu. Yale School of Management. December 21, 2021. https://som.yale.edu/blog/part-ii-of-crisis-in-lebanon-buildup-of-interrelated-challenges.
5. A Manageable Suggestion.” Journal of Post Keynesian Economics 31 (1): 125–38.
6. Aldasoro, Iñaki, Claudio, and Mathias Drehmann. 2018. “Early Warning Indicators of Banking Crises: Expanding the Family.” Ssrn.com. March 11, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3139160.
7. Babecký, Jan, Tomáš Havránek, Jakub Matějů, Marek Rusnák, Kateřina Šmídková, and Bořek Vašíček. 2014. “Banking, Debt, and Currency Crises in Developed Countries: Stylized Facts and Early Warning Indicators.” Journal of Financial Stability 15 (December): 1–17. https://doi.org/10.1016/j.jfs.2014.07.001.
8. Banking Crises: A Review. (Laeven, 2015). Annual Reviews.
9. Baumann, Hannes. 2019. “The Causes, Nature, and Effect of the Current Crisis of Lebanese Capitalism.” Nationalism and Ethnic Politics 25 (1): 61–77.
10. Baz, Salim, Lara Cathcart, and Alexander Michaelides. 2022. “Lebanon: From Dollars to Lollars.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4254934.
11. Beutel, Johannes, Sophia List and Gregor von Schweinitz. “Does machine learning help us predict banking crises?” Journal of Financial Stability 45 (2019): 100693.

12. Calomiris, Charles W., and Joseph R. Mason. "Consequences of bank distress during the Great Depression." American Economic Review 93, no. 3 (2003): 937-947.

13. Caprio, Gerard, James A. Hanson, and Robert E. Litan, eds. Financial Crises: Lessons from the Past, Preparation for the Future. Brookings Institution Press, 2005.

14. Cerutti, Eugenio, Galina Hale, and Camelia Minoiu. 2015. “Financial Crises and the Composition of Cross-Border Lending.” Journal of International Money and Finance 52 (April): 60–81. https://doi.org/10.1016/j.jimonfin.2014.11.013.

15. Charles, Sébastien. 2008. “Teaching Minsky’s Financial Instability Hypothesis”

16. Demirgüç-Kunt, Asli, and Enrica Detragiache. “Cross-Country Empirical Studies of Systemic Bank Distress: A Survey.” National Institute Economic Review 192 (2005): 68–83. doi:10.1177/002795010519200108.

17. Drehmann, Mathias, and Mikael Juselius. 2014. “Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements.” International Journal of Forecasting 30 (3): 759–80. https://doi.org/10.1016/j.ijforecast.2013.10.002.

18. Drehmann, Mathias, Mikael Juselius, and Anton Korinek. 2018. “Going with the Flows: New Borrowing, Debt Service and the Transmission of Credit Booms,” April. https://doi.org/10.3386/w24549.

19. Dooley, Michael. 2003. “Title: Managing Currency Crises in Emerging Markets Title: Introduction to ‘Managing Currency Crises in Emerging Markets’” Publisher: 0–226. https://www.nber.org/system/files/chapters/c9643/c9643.pdf.

20. Elfeituri, Hatem. 2022. “Banking Stability, Institutional Quality, Market Concentration, Competition and Political Conflict in MENA.” *Journal of International Financial Markets, Institutions and Money* 76 (January): 101476. https://doi.org/10.1016/j.intfin.2021.101476.

21. Elia, Jean, and Candidate. n.d. “LEBANESE BANKS: A FACTOR of the CURRENT LEBANESE FINANCIAL CRISIS (2019-2020).”

22. Faour. 2022. “REPORT on the THIRD WEBINAR SERIES LEBANON’S ECONOMIC CRISIS: IS THERE a WAY FORWARD?” Accessed March 18, 2023. https://www.aub.edu.lb/osb/facultyresearch/Events/Documents/Lebanon%E2%80%99s%20Economic%20Crisis-Report%20on%20a%20Webinar%20Series%203.pdf.

23. Glick, Reuven, and Michael Hutchison. n.d. "Pacific Basin Working Paper Series BANKING and CURRENCY CRISES: HOW COMMON ARE TWINS? Center for Pacific Basin Monetary and Economic Studies Economic Research Department Federal Reserve Bank of San Francisco."

24. Gómez, S. (2019). Centrality in Networks: Finding the Most Important Nodes. In: Moscato, P., de Vries, N. (eds) Business and Consumer Analytics: New Ideas. Springer, Cham. https://doi.org/10.1007/978-3-030-06222-4_8

25. Hausenblas, Václav, Ivana Kubicová, and Jitka Lešanovská. 2015. "Contagion Risk in the Czech Financial System: A Network Analysis and Simulation Approach." Economic Systems 39 (1): 156–80. https://doi.org/10.1016/j.ecosys.2014.07.001.

26. Jacklin, Charles J., and Sudipto Bhattacharya. "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications." Journal of Political Economy 96, no. 3 (1988): 568–92.

27. Janiesch, Christian, Patrick Zschech, and Kai Heinrich. 2021. "Machine Learning and Deep Learning." Electronic Markets 31 (3): 685–95. https://doi.org/10.1007/s12525-021-00475-2.

28. Kaminsky, Graciela L., and Carmen M. Reinhart. "The twin crises: the causes of banking and balance-of-payments problems." American economic review 89, no. 3 (1999): 473-500.

29. Laeven, Luc, and Fabián Valencia. 2008. "Systemic Banking Crises: A New Database." Ssrn.com. September 2008. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1278435.

30. Mian, Atif R., and Amir Sufi. 2018. "Finance and Business Cycles: The Credit-Driven Household Demand Channel." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3124001.

31. "Mora Lebanon Primer Jan 2020.Pdf." 2020. Google.com. 2020. https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxkcm5hZGFtb3JhfGd4OjRlNjI1YWU0NjAxMjgzMjE.

32. Nasr, Salim. "Backdrop to Civil War: The Crisis of Lebanese Capitalism." MERIP Reports, no. 73 (1978): 3–13. doi:10.2307/3012262.

33. Reinhart, Carmen M., and Kenneth S. Rogoff. "The aftermath of financial crises." American Economic Review 99, no. 2 (2009): 466-72.

34. “Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2022.” n.d. https://www.nobelprize.org/uploads/2022/10/advanced-economicsciencesprize2022-2.pdf.

35. Sturzenegger, Federico, and Jeromin Zettelmeyer. “Creditors’ Losses versus Debt Relief: Results from a Decade of Sovereign Debt Crises.” Journal of the European Economic Association 5, no. 2/3 (2007): 343–51.

36. Sufi, Amir and Taylor, Alan M., Financial Crises: A Survey (August 2021). NBER Working Paper No. w29155, Available at SSRN: https://ssrn.com/abstract=3905719

37. Tongyu Wang, Shangmei Zhao, Guangxiang Zhu & Haitao Zheng (2021) A machine learning-based early warning system for systemic banking crises, Applied Economics, 53:26, 2974-2992, DOI: 10.1080/00036846.2020.1870657

38. Wicker, E. (1980). A Reconsideration of the Causes of the Banking Panic of 1930. The Journal of Economic History, 40(3), 571-583.