


Article

Markov Chain K-Means Cluster Models and Their Use for Companies' Credit Quality and Default Probability Estimation

Nora Gavira-Durón ^{1,*} , Octavio Gutierrez-Vargas ² and Salvador Cruz-Aké ²¹ Department of Banking and Investments, School of Business and Economics, Universidad de las Americas Puebla, Puebla 72810, Mexico² Higher School of Economics, Instituto Politécnico Nacional, Santo Tomas Campus, Ciudad de México 11350, Mexico; ogutierrezv1800@alumno.ipn.mx (O.G.-V.); scruga@ipn.mx (S.C.-A.)

* Correspondence: nora.gavira@udlap.mx

Abstract: This research aims to determine the existence of inflection points when companies' credit risk goes from being minimal (Hedge) to being high (Ponzi). We propose an analysis methodology that determines the probability of hedge credits to migrate to speculative and then to Ponzi, through simulations with homogeneous Markov chains and the k-means clustering method to determine thresholds and migration among clusters. To prove this, we used quarterly financial data from a sample of 35 public enterprises over the period between 1 July 2006 and 28 March 2020 (companies listed on the USA, Mexico, Brazil, and Chile stock markets). For simplicity, we make the assumption of no revolving credits for the companies and that they face their next payment only with their operating cash flow. We found that Ponzi companies (1) have a 0.79 probability average of default, while speculative ones had (0) 0.28, and hedge companies (−1) 0.009, which are the inflections point we were looking for. Our work's main limitation lies in not considering the entities' behavior when granting credits in altered states (credit relaxation due to credit supply excess).

Keywords: homogeneous markov chains; k-means; credit risk; default clusters



Citation: Gavira-Durón, N.; Gutierrez-Vargas, O.; Cruz-Aké, S. Markov Chain K-Means Cluster Models and Their Use for Companies' Credit Quality and Default Probability Estimation. *Mathematics* **2021**, *9*, 879. <https://doi.org/10.3390/math9080879>

Academic Editors:

José Álvarez-García, Oscar V. De la Torre-Torres and María de la Cruz del Río-Rama

Received: 17 February 2021

Accepted: 1 April 2021

Published: 16 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Credit positioning evolution can migrate through stability and instability periods of the institutions that request it. In the first case, loans represent low-risk profitable funds (hedge), which, over time and due to financial risk factor variations, market conditions, and socio-economic issues finally may become speculative and high risk (Ponzi).

It is fundamental to highlight that traditional credit for companies is not safe from the impact produced during crises, and also depends on intrinsic factors such as their level of capitalization, assets, liquidity, and solvency, among others.

In an economy where banks have a close financial relationship with each other, past, present, and future are linked not only to capital assets and workforce but also to financial relationships. This fact was indicated by Minsky [1] in his second theorem of financial instability.

Minsky determines three types of income–debt relations for the firms: hedge, speculative, and Ponzi financing. This research finds transition thresholds between each state proposed by Minsky's hypothesis and classifies the firm's actual state. This research goes far beyond calculating bankruptcy probabilities; it gives actual thresholds to the Minsky hypothesis and may be used to check the industry and financial firm's health.

In this paper, we use the quarterly operations cash flows and the theoretical payment obtained under the assumption of the company paying their long-run debt in ten years, at the companies effective interest rate, (financial cost divided by long-run debt) as the underlying and the strike price of a theoretical call option for each company.

The idea behind this theoretical call is to assess the firm's bankruptcy probability on next quarter in the absence of other credit sources; a stand-alone assumption given by the

Minsky hypothesis (the firm's reason to obtain debt) adapted from a vanilla (European) option framework. The vanilla option framework arises from the idea of the firm covering interests and a part of the long-run debt each quarter or going into bankruptcy.

Instead of arbitrarily classifying the firm's probabilities into groups, we used a clustering procedure to determine the number of groups (using the elbow technique) and their membership in each quarter. We obtained three clusters as the optimal number of groups, which coincides with the Minsky theory. We used the Markov chains to get the transition probabilities through groups and their stationarity states.

We decided to use Markov chains because two main reasons: First, Markov chains are the standard technique when dealing with default probabilities, so their use makes our methodology comparable. Second, Markov chains allow us to find stationary states to compare a particular quarter with the long-run equilibrium and then determine the industry's health or the analyzed economy.

The use of Markov chains also allowed us to set thresholds between each group by taking the average default probability on each group along with the sample, thus giving a clear cut between Minsky's groups.

According to Minsky's theory, adapted to our operational definition, a hedge company can cover its payments and only use debt to start new profitable projects. By their side, speculative firms can cover their interest obligations but not the principal without bridge loans, because they have some unprofitable projects waiting for better economic conditions to become profitable.

Finally, Minsky postulates that Ponzi enterprises do not have sufficient operating cash flow to cover interests nor principal on their debts, which means that they must acquire more debt to cover those obligations or face bankruptcy. The gradual temptation to expand and companies' expenditures increase these debts to generate more income [2].

Dickinson [3] postulates that such financing increase is naturally generated in the business growth cycle, which can be seen as the credit system's evolution. In a hypothetical case, whether financial or not, an entity goes from having profitable hedge credits to having high-risk Ponzi credits, given the necessities of higher financial requirement: the more profitable, the riskier. In some cases, this leads to a financial collapse; whereas, in others, the course is corrected and it returns to low-risk credits.

The literature reviewed for corporate credit analysis does not consider an inflection point among hedge, speculative, and Ponzi credits; to wit, there is no preventive measure that announces a possible change in the firm's credit status. The lack of preventive information occurs for two main reasons: First, the rating companies keep their methodologies as an industrial secret and traditionally make risks assessments only for public companies. Second, market oversight implies very different views and methodologies converging on prices only if the market is efficient and deep.

There are classical models of default probabilities based on Markov chains; however, these models use bond prices and historical recovery rates in their estimation. The main problem in this approach is that companies must be public to apply such models. In this article, we use the Merton [4] model's to calculate companies' default probabilities in the next period using Markov chains and financial quarterly reports (operating cash flows, financial expenses and long-run liabilities), obtained from Economatica, over the period from 1 July 2006 to 30 March 2020 from 35 public companies.

According to Minsky theory, we classify companies using their next quarter default probabilities; calculated with Merton's models' adaptation. It is worth mentioning that this method is not chained to normal distribution. If the researcher decides so, the methodology can use any probability distribution as long as a closed form for a vanilla call exists.

We classify companies, in each quarter, using their default probabilities and then checking their stability in their classification (short and long term) using the Markov chains. This procedure is analogous to that performed in traditional credit analysis, but it is now extendable to any private company without regards to its size or whether it is publicly traded.

The methodology to be used comprises simulations with homogeneous Markov chains and a clustering method (k-means) in order to determine changing thresholds for the different company's credit positions; which, if any, will determine a functional proposal to allow classifying other companies, and the feasible regions for each financing type will be determined.

We believe that the paper's main contribution is the novel application of Markov chains to quantify Minsky's hypothesis in any firm. We used a classical mathematical tool to give a quantitative solution to a classic economic problem. We also give a reference for an industry or economy's health to decision-makers.

This methodology allows the analyst to know companies' solvency and compare it with the industry's risk exposition. Regulators or economic decision-makers could measure and compare a big company's financial health and use it as an indicator for possible third-generation crises, i.e., companies that could systematically all default. It is important to stress that we consider that the cash flows are an ergodic stochastic process.

A remarkable fact about credit models is their dependence on normal or logistic distribution assumptions. Models such as the Altman z-score must be specified according to their industry and company's size. According to Minsky's theory of financial market instability, our application prevents distribution dependence assumptions and is a pioneer in Markov chain application for quantitative classification and determination of credit migration probabilities. With that, we present a mathematical application which enriches economic theory.

We also want to point out that the Merton Model only provides default probabilities under the assumption of full payment of debt upon an option's expiration rate (an unlikely assumption due to the need for financial structure requirements), and the Altman's Z score classifies firms on arbitrary groups using a methodology that must be adapted to each industry and firm's size. The need for adaptation in Altman's methodology creates non-comparable groups. Alongside this, the econometric methods rely on the distribution (normal, logistic, or Poisson) and ergodicity assumptions to only provide default probabilities, but the firm's classification is arbitrary. None of them provide an internally coherent classification methodology or provide further analysis.

As possible future investigation lines, we consider models that could modify companies and the industry's default probabilities through Hidden Markov chains where the external factor could be the regulator's decisions or market's volatility.

In Section 2, we review the literature on the methodologies implemented for corporate credit analysis. In Section 3, we introduce the financial sector characteristics. Section 4 presents the methodology, while Section 5 presents the results. Furthermore, in Section 6, we show the research conclusions, recommendations, and limitations.

2. Literature Review

A significant amount of research exists on transition matrix application to estimate the probability of default due to credit quality impairment on commercial, microcredit, and mortgage portfolios in Latin America; such as that of Aparicio et al. [5] who analyze the credit portfolio of the Peruvian financial system considering the credit transition matrix usage subject to the economic cycle.

For the Colombian study case, Támara-Ayús, Aristizábal, and Velásquez [6] compare discrete and continuous transition matrices. There are also studies with GNP (Gross National Product) as that carried out by McCulloch and Tsay [7] observing that uncertainty about the situation of a given period will depend on the model specifications. By his side, Peña [8] concludes that economic cycles and macroeconomic variables influence credit quality impairment.

Hamilton [9] has proposed default models with Markov chains by using a chain comparable to logarithmic levels and trend constructed from the same time series model. In the case of Nicaragua, Gaitán, and Flores [10] analyze banking institutions' credit portfolio through transition models with Markov chains.

Regarding the Venezuelan case, Porras, Anchundia, and Vieira [11] analyze it through transition matrices for the ex-ante and ex-post periods of the foreign capital influx to determine if the rivalry among financial institutions increases as shown in the Yu et al. [12], Pfeuffer and Reis [13], Leung and Kwok [14] papers.

On the literature review, we found that there are models that join the classification of companies and risk structure and delve into the need for updating measurement models. Córdova, Molina, and Navarrete [15] and García, Bolívar, and Vázquez [16] provide examples of this.

There are applications to assess the risk of individuals such as that of Kavitha [17] who proposes a model based on k-means, which has higher efficiency in accuracy and time compared to the traditional methods. For counterparty credit risk, Zhu, Chan, and Bright [18] apply machine learning techniques to determine and solve problems in XVA and credit profiles, through Monte Carlo simulation and K-means clustering method.

In the credit risk analysis of Small and Medium Enterprises (SMEs)—applied to Bahasa Indonesia—Wahyudin, Djatna, and Kusuma [19] model risk clusters by using K-means and risk measurement by calculating qualitative importance scores combined with sentiment scores. The result shows that the model is adequate for clustering and measuring the risk level.

In the case of Mexico, Ayús, Peña, and Álvarez [20] present an analysis of different studies on the commercial portfolio credit risk. They develop a model to predict the probability of having a debtor's default, through factorial and discriminant analysis.

By their side, Lagunas and Ramírez [21] estimate scenarios to determine the number of operations likely to receive illegal financial transactions through stochastic transition matrices.

Our main work contribution is to propose an algorithm capable of classifying a company's credit health according to its operating flow as a proportion of its financial requirements. To simplify the calculations, we assume that the firm will not invest nor have access to other financing sources; with the correct internal information, the analyst can relax this assumption. In addition, we present a method based on Markov chains to determine the thresholds at which companies move from one credit state to another.

3. The Financial Sector

The financial sector not only can generate a boom—regarding cycles—by providing a larger credit volume but also can induce depression Ferreira [22]. In the event that the stock system collapsed, there would be a credit reduction, and this would lead to an economic debacle according to Angeles and Ortiz [23].

It is even possible to see that, in 2008, the debt equity ratio (company's indebtedness capacity against equity) for the banks was not primarily altered, coinciding with Rodríguez and Venegas [24]. We show this in Figure 1.

There are studies such as Berger and Udell [25] which show that companies are always in a financial cycle that commences with financing. There is also evidence of intermediary companies that help in the economic cycle as the case of Bencivenga and Smith [26]. Papers such as that of Adrian and Shin's [27] show that financial institutions function pro-cyclically. Furthermore, some authors argue that while recessions accompanied by financial problems are more prolonged and profound, recoveries are slightly shorter and stronger Claessens, Kose, and Terrones [28].

Terreno, Sattler, and Pérez [29] observed that a typical company asks for debt for continuing with the operations or carrying projects out; this is part of companies' life cycle.

In this paper, we propose a variation from Merton's methodology in order to calculate companies' short-term default probability. Throughout this work, we analyze whether the company is capable of covering its financial requirements under two assumptions: (a) the company does not have access to other financing sources and (b) the company will not carry out any investment operation in the period.

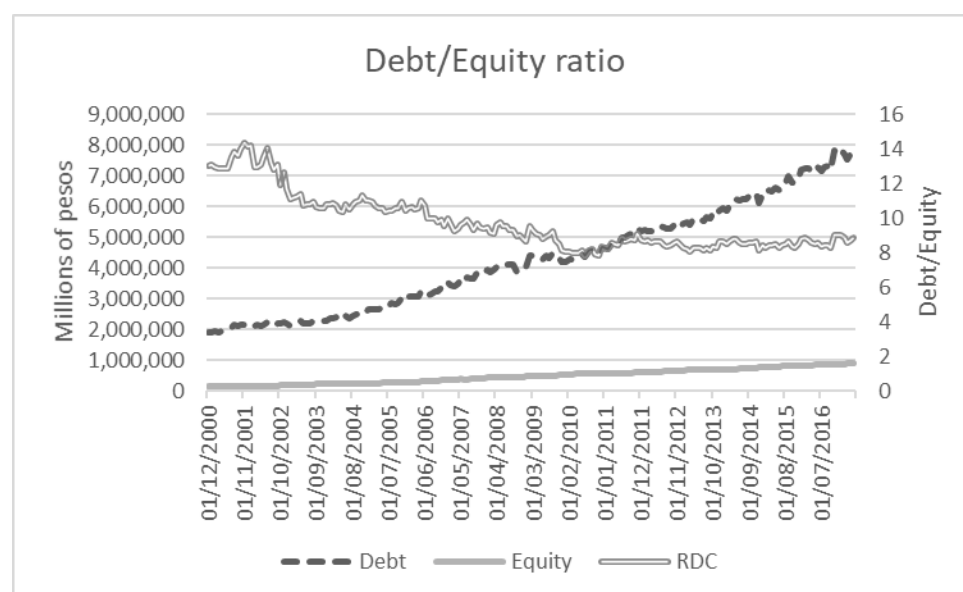


Figure 1. Total annual indebtedness of the Multiple Banking in Mexico. Source: Own elaboration I, with data obtained from the CNBV page (Comisión Nacional Bancaria y de Valores).

Unlike Merton's method, this article focuses on the company's short-term capacity to cover its financial requirements itself only in the following period though the assumption of normality in the innovations.

4. Methodology

Merton's model for corporate default uses the vanilla options rationale to calculate the default probability. Papers from Tudela and Young [30]; Dar, Anuradha, and Qadir [31]; Afik, Arad and Galil [32] are examples of the use of this method.

The dynamics for V , company's operating cash flow over time is described by a stochastic diffusion process where dz is a Brownian motion:

$$dV = Vdt + \sigma Vdz, \quad (1)$$

The cash flow must be enough each period to cover at least the payment of the debt, C ; this implies that we are interested in the section of the distribution where $V_t > C_t$.

Rephrasing this idea, under our assumptions, the shareholders will receive a net cash flow, $S_T = \text{Max}[V_T - C, 0]$ at the end of the trimester, T , from the firm.

Following the Merton's model, at the beginning of the trimester, t , the present value of the expected net cash flow to the shareholders is

$$S_t = V\Phi(d_1) - Ce^{-r(T-t)}\Phi(d_2) \quad (2)$$

where σ is the operational cash flow volatility. Therefore, the probability that the company can make the debt payment in the trimester, given the information contained in the sigma field, F , is

$$P\{V_t > C|F\} = \Phi(d_2) \quad (3)$$

therefore, the probability of default is given by

$$P\{V_t < C|F\} = 1 - \Phi(d_2) \quad (4)$$

4.1. Company Grouping or Clustering

The classification algorithms have become increasingly used because of their potential. Areas of knowledge, such as marketing objectives, medical diagnoses, event detection, categorization, and filters. For more details, see Aggarwal [33].

Due to its widely known classification properties, we use in this paper a dendrogram classification, for details in its use see Berkhin [34]. In particular, we use the AGNES algorithm for R (“Agglomerative Nesting”, [35–37]). We want to emphasize that the algorithm uses a Euclidean distance to classify the objects; this is

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{ip} - x_{jp})^2} \quad (5)$$

We also use a k-means algorithm to check the dendrogram grouping. For details on its use, see Hartigan and Wong [37] or Ledolter [38]. For now, it suffices to say that a k-means algorithm divides the n units into $k \leq n$ different clusters $S = \{S_1, S_2, \dots, S_k\}$, minimizing the within clusters sum of squares until there is no gain in dividing the groups.

$$\arg \min \sum_{j=1}^k \sum_{x_i \in S_j} \|x_i - \bar{m}_j\|^2 \quad (6)$$

Using these methodologies on our data, it is possible to obtain the state transition matrix for the Markov chain; we show it the following section.

4.2. Markov Chains

The literature defines a Markov chain as a system that presents changes between states according to a fixed probabilistic rule (transition probabilities), so the process can be in any state with any past combinations and can take a fixed number of future states.

According to Caballero et al. [39]: given $(\Omega, \mathfrak{F}, \mathbb{P})$ a probability space and E a non-empty, finite, or countable set, it is said that a random variable succession $\{X_n : \Omega \rightarrow E, n = 0, 1, \dots\}$ is referred to as Markov chain with space of E states if it satisfies Markov’s condition, that is, if $\forall n \geq 0$ and for any collection $x_0, x_1, \dots, x_{n-1}, x, y \in E$ is met.

$$\mathbb{P}(x_{n+1} = y | X_n = x, \dots, X_1 = x_1, X_0 = x_0) = \mathbb{P}(X_{n+1} = y | X_n = x) \quad (7)$$

In the context of credit risk, the Markov chain is used to model a debtor’s credit rating migration. For more details, see Norris [40], Venegas [41], Bolívar-Cimé, Notario, and Pérez [42], and Caballero et al. [39]. In order to compare our results, in Table 1 we show the Fitch’s transition matrix for Mexican corporative bonds over the period from 2002 to 2018.

Table 1. Mexico Corporate Finance National Ratings Transition Matrices.

(%) Annual	AAA (mex)	AA (mex)	A (mex)	BB (mex)	BB (mex)	B (mex)	CCC (mex)	CC (mex)	C (mex)	D (mex)
AAA (mex)	96.0	0.6	0.3							
AA (mex)	4.5	88.4	2.5	0.3	0.3		0.3			0.3
A (mex)	0.5	3.9	88.2	2.7	0.3				0.2	0.6
BB (mex)		0.2	8.4	80.7	2.7				0.2	1.2
BB (mex)			1.0	8.1	72.2	1.0	1.0			2.0
B (mex)				2.3	13.6	59.1	9.1	2.3		2.3
CCC (mex)					6.3	31.3	31.3		6.3	12.5
CC (mex)								25.0		25.0
C (mex)						33.3				66.7

Source: Prepared by the authors with data from the Fitch Ratings report [43]: “Study on Transition and Failure of National Ratings of Mexico 2018”.

In this paper, we find the inflection point between hedge and Ponzi companies, regarding the companies’ operating flow and financial indicators; hence, it is more appropriate to use the matrices generated by Markov chains, similarly to Bolívar-Cimé, Notario and Pérez [42].

5. Results

This research uses quarterly operating flows as well as financial requirements (interest plus estimated amortization) of a sample of 35 companies listed on USA, Mexico, Chile, and Brazil, obtained from “Economatica” over the period from 1 July 2006 to 30 March 2020, representing twelve quarters.

We applied the k-mean methodology to each of the four quarters of financial information. In our analysis, we assume no revolving credits to the company nor the chance of using cash reserves, so they face the next payment only with their operating flow.

Hence, there are differences in the default probability calculation in the proposed methodology compared to those made by the rating agencies. We show the results in Table 2.

Table 2. Average default probabilities compared to agency ratings.

#	Company	Rating	Agency	Calculated Probability
1	Boeing Co Com (Chicago, IL, USA)	BBB-	Fitch	0%
2	Embraer ON (São José dos Campos, São Paulo, Brazil)	BB+	Fitch	0%
3	Exxon Mobil Corp Com (Irving, TX, USA)	AA	S&P	2%
4	AT&T Inc Com (Whitacre Tower, Dallas, TX, USA)	A-	Fitch	1%
5	Chevron Corp Com (San Ramon, CA, USA)	AA	S&P	2%
6	Verizon Comm Inc Com (New York City, NY, USA)	A-	Fitch	1%
7	Amazon.Com, Inc Com (Seattle, Washington, DC, USA)	A+	Fitch	1%
8	Walmart Inc Com (Bentonville, AR, USA)	AA	Fitch	2%
9	Unitedhealth Group Inc Com (Minnetonka, MN, USA)	A	Fitch	1%
10	Comcast Corp Com A (Philadelphia, PA, USA)	A-	Fitch	1%
11	Johnson & Johnson Com (New Brunswick, NJ, USA)	AAA	Fitch	1%
12	Intel Corp Com (Santa Clara, CA, USA)	A+	Fitch	1%
13	Procter & Gamble Co Com (Cincinnati, OH, USA)	BBB	Fitch	0%
14	Pfizer Inc Com (New York, NY 10017, USA)	A-	Fitch	1%
15	General Motors Company Com (Detroit, MI, USA)	BBB-	Fitch	0%
16	Intl Business Machines Corp Com (Armonk, NY, USA)	A+	Fitch	1%
17	Walt Disney Co Com (Burbank, CA, USA)	A-	Fitch	1%
18	Cisco Systems, Inc Com (San Jose, CA, USA)			
19	Fomento Econ Mex UBD (Monterrey, Mexico)	A	Fitch	1%
20	Cemex CPO (Monterrey, Mexico)	BB	S&P	0%
21	Alfa A (Monterrey, Mexico)	BBB-	Fitch	0%
22	America Movil L (Mexico City, Mexico)	AAA	Fitch	1%
23	Elektra Gpo. (Mexico City, Mexico)	A	Fitch	1%
24	CCR SA ON (São Paulo, Brazil)	AA+	Fitch	2%
25	Cemig ON (Belo Horizonte, MG, Brazil)	BB-	Fitch	0%
26	Sid Nacional ON (São Paulo, Brazil)	B-	S&P	0%
27	CPFL Energia ON (Campinas, Brazil)	AAA	Fitch	1%
28	Gerdau PN (Porto Alegre, RS, Brazil)	BBB-	Fitch	0%
29	Empresas Cmpc S.A. Ord (Santiago, Chile)	BBB	Fitch	0%
30	Empresas Copelec S.A. Ord (Santiago, Chile)	BBB	Fitch	0%
31	Enel Americas S.A. Ord (Santiago, Chile)	A-	Fitch	1%
32	Enel Generacion Chile S.A. Ord (Santiago, Chile)	A-	Fitch	1%
33	Buenaventura C1 (San Isidro, Perú)	BB+	Fitch	0%
34	BRF SA ON (Itajaí, Santa Catarina, Brazil)	BB	Fitch	0%
35	Ambev S/A ON (São Paulo, Brazil)	BBB	S&P	0%

Source: Own elaboration with data obtained from Economatica.

In order to classify the companies, we will use only their default probability measured with our Merton method variation. Then, we use the AGNES method (Agglomerative Nesting, implemented in R-project) to group them. The AGNES methodology results in the dendrogram showed in Figure 2.

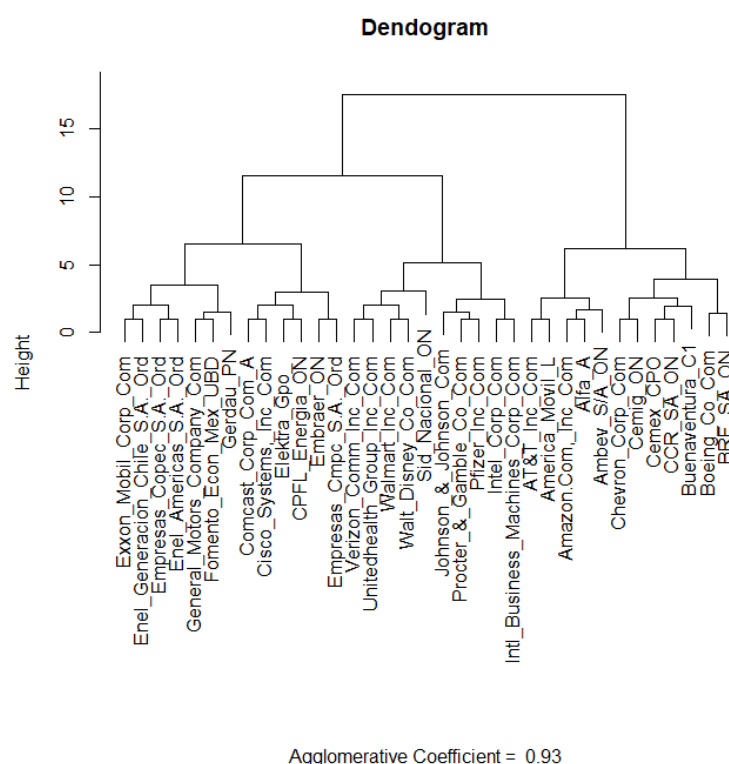


Figure 2. Dendrogram for the companies selected. Source: Own elaboration with R-project and “cluster” package.

Although it is not possible to appreciate the cluster to which each company belongs (we show it in Table 3), it is possible to observe the hierarchy.

As a result of Agglomerative Nesting procedure, AGNES, we can find that companies are aggrouped, according to their default probabilities, in 3 clusters. Clustering will help us determine the possible states for the stochastic process (The Minsky’s classification). The Markov chains determinate transition probabilities and long-run belonging rate for each state in the transition matrix.

Table 3 shows the company grouping results as well as their transition over the periods. We use values from the {1 (Ponzi), 0 (Speculative), −1 (Hedge)} set to show their group, according to their default probability level.

The results show that companies that start in Ponzi (1), generally continue in the same group throughout the different periods; sometimes they may be speculative in any period, but they fail to be a hedge (−1) firm. In the case of ELEKTRA GPO, migration from Ponzi to hedge occurred “jumping” intermediately in speculative periods and hedge.

While most companies that started in the speculative (0) group recovered in the second period, we observed that in some cases after several periods in the hedge classification they either showed impairment or went directly to a Ponzi stage, as in the CEMEXCPO case. Regarding companies that started in a hedge position, the majority stayed there and did show an inflection point when they became speculative before being classified as Ponzi.

We want to emphasize that companies classified as Ponzi belong to the aerospace and beer industries. Speculative companies are focused on commerce, food, and iron production. Hedge companies are focused in several areas.

It is imperative to clarify that the financial requirement coverage analysis eliminates any possible size bias within the sample as it is based on proportion.

Table 3. Membership of companies by default cluster.

Company	Q4 -2018	Q1 -2019	Q2 -2019	Q3 -2019	Q4 -2019	Q1 -2020	Q4 -2018	Q1 -2019	Q2 -2019	Q3 -2019	Q4 -2019	Q1 -2020
BRF SA ON (Itajaí, Santa Catarina, Brazil)	−1	−1	0	1	−1	−1	0%	0%	65%	88%	20%	17%
Ambev S/A ON (São Paulo, Brazil)	−1	−1	−1	−1	−1	−1	0%	0%	0%	5%	0%	0%
Exxon Mobil Corp Com (Irving, TX, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
AT&T Inc Com (Whitacre Tower, Dallas, TX, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Chevron Corp Com (San Ramon, CA, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Verizon Comm Inc Com (New York City, NY, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Amazon.Com, Inc Com (Seattle, Washington, DC, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Walmart Inc Com (Bentonville, AR, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Unitedhealth Group Inc Com (Minnetonka, MN, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Comcast Corp Com A (Philadelphia, PA, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Johnson & Johnson Com (New Brunswick, NJ, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Intel Corp Com (Santa Clara, CA, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Procter & Gamble Co Com (Cincinnati, OH, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Pfizer Inc Com (New York, NY 10017, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
General Motors Company Com (Detroit, MI, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Intl Business Machines Corp Com (Armonk, NY, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Walt Disney Co Com (Burbank, CA, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Cisco Systems, Inc Com (San Jose, CA, USA)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Fomento Econ Mex UBD (Monterrey, Mexico)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Cemex CPO (Monterrey, Mexico)	−1	0	0	−1	−1	−1	1%	25%	60%	16%	14%	3%
Alfa A (Monterrey, Mexico)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
America Movil L (Mexico City, Mexico)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Elektra Gpo (Mexico City, Mexico)	1	−1	−1	1	0	−1	100%	0%	0%	100%	32%	0%
Cemig ON (Belo Horizonte, MG, Brazil)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Sid Nacional ON (São Paulo, Brazil)	0	−1	−1	−1	−1	−1	2%	0%	0%	0%	0%	0%
CPFL Energia ON (Campinas, Brazil)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Embraer ON (São José dos Campos, São Paulo, Brazil)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	2%
Gerdau PN (Porto Alegre, RS, Brazil)	−1	−1	−1	−1	−1	−1	0%	4%	0%	0%	0%	0%
Empresas Cmpc S.A. Ord (Santiago, Chile)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Empresas Copec S.A. Ord (Santiago, Chile)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Enel Americas S.A. Ord (Santiago, Chile)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Enel Generacion Chile S.A. Ord (Santiago, Chile)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%
Buenaventura C1 (San Isidro, Perú)	−1	−1	0	0	−1	−1	0%	1%	62%	16%	19%	0%
Boeing Co Com (Chicago, IL, USA)	1	1	1	1	1	1	100%	100%	100%	100%	100%	100%
CCR SA ON (São Paulo, Brazil)	−1	−1	−1	−1	−1	−1	0%	0%	0%	0%	0%	0%

Source: Own elaboration with R-project and “cluster” package.

Since a Markov chain is regular, if $n \rightarrow \infty$, the exponentials $P^{(n)}$ of the P transition matrix converge to a matrix W in a way that all its rows are equal to the same probability vector $u = (u_x)_{x \in E}$, then, for the case studied $u = \{0.114285, 0.0571428, 0.828571\}$ is obtained. We show the transition matrix in Table 4 and Figure 3.

Table 4. Quarterly transition matrix of the companies from the sample. 01/03/2014 to 01/03/2020.

	1	0	−1
1	0.75	0.25	0
0	0	0.5	0.5
−1	0.0344	0	0.9655

Source: Own elaboration with R-project.

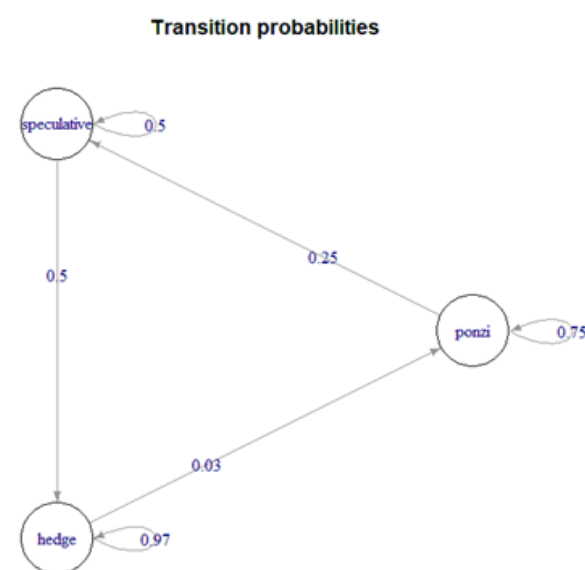


Figure 3. Matrix of graphical transition for companies. Source: Own elaboration with R-project and igraph package.

By carrying out 5000 simulations with the transition matrix, we found that a company could go from a hedge classification, −1, to the Ponzi group, 1, in an average of 34 periods, as shown in Figure 4.

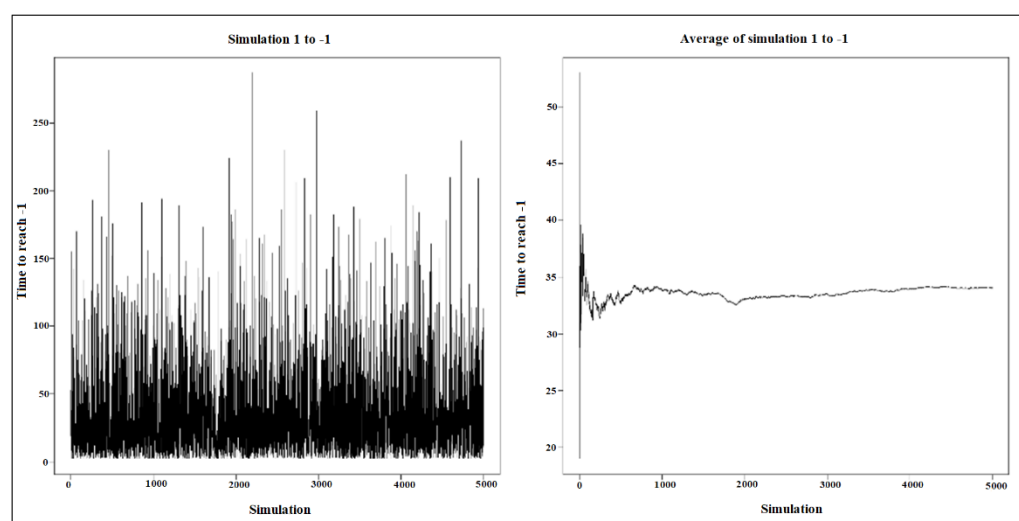


Figure 4. Simulations and average time of state “1”. Source: Own elaboration with R-project.

A significant result obtained from k-means grouping over the default probabilities is that firms in the Ponzi state (1) have an edge probability of default of 0.79, while the speculative (0) ones have 0.28; finally, we found that hedge firms (−1) have an edge default probability of 0.009. Therefore, we can confirm that the inflection point exists. Thus, the resulting classification function generated by our model is

$$f(x) = \begin{cases} -1, & 0 < x \leq 0.28 \\ 0, & 0.28 < x \leq 0.79 \\ 1, & x > 0.79 \end{cases} \quad (8)$$

We want to point out that transition probabilities and asymptotical states are stable along with the sample and on out of the sample trials. This feature, stability, make the asymptotical states a useful tool to assess the industry's health. Its stability also allows us to give cutting edges to the theoretical Minsky groups, one of our main contributions.

Finally, in Table 5 we compare Z Altman versus our application. As we said before, it is not comparable in their totally but also contains three stages. It is shown that Z Altman have several companies which are in the early bankruptcy stage meanwhile our application says another thing. With our application there is, in the last period, one company (Boeing) classified as a Ponzi. Remember that according to Minsky, a Ponzi enterprise does not have enough operating cash flow to cover interest or debts, the enterprise must acquire more debt to cover those obligations and, in our case, we are supposing no revolving credits and going to face next payment with their cash flow. A big difference between our application and Z Altman is that Z Altman uses ratio analysis and multiple discriminant analysis which need a multivariate normal distribution in continuous variables meanwhile our application uses probability theory (Markov chains), financial mathematics (Merton's model), and data mining (clustering).

Table 5. Z Altman vs. Minsky's Groups. 1 March 2014 to 1 March 2020.

Company	Q3 -2019	Q4 -2019	Q1 -2020	Q3 -2019	Q4 -2019	Q1 -2020
Boeing Co Com (Chicago, IL, USA)	1	1	1	2.570504494	2.233507057	1.43531255
Embraer ON (São José dos Campos, São Paulo, Brazil)	−1	−1	−1	0.898896071	0.888065486	0.501569282
Exxon Mobil Corp Com (Irving, TX, USA)	−1	−1	−1	3.686010408	3.489752209	3.055181343
AT&T Inc Com (Whitacre Tower, Dallas, TX, USA)	−1	−1	−1	1.240510776	1.227069878	1.064776635
Chevron Corp Com (San Ramon, CA, USA)	−1	−1	−1	3.602881718	3.581614816	3.244899774
Verizon Comm Inc Com (New York City, NY, USA)	−1	−1	−1	1.808870621	1.760786898	1.724445429
Amazon.Com, Inc Com (Seattle, Washington, DC, USA)	−1	−1	−1	7.738432004	7.586699588	9.474626187
Walmart Inc Com (Bentonville, AR, USA)	−1	−1	−1	4.591543813	4.684393948	4.816751716
Unitedhealth Group Inc Com (Minnetonka, MN, USA)	−1	−1	−1	3.773413778	4.012850064	3.770474019
Comcast Corp Com A (Philadelphia, PA, USA)	−1	−1	−1	1.809844627	1.747225324	1.611141332
Johnson & Johnson Com (New Brunswick, NJ, USA)	−1	−1	−1	4.972185446	5.378295088	5.673101051
Intel Corp Com (Santa Clara, CA, USA)	−1	−1	−1	5.227025032	5.872670718	4.724158649
Procter & Gamble Co Com (Cincinnati, OH, USA)	−1	−1	−1	6.063299178	6.142447306	5.463732498
Pfizer Inc Com (New York, NY 10017, USA)	−1	−1	−1	2.90145047	2.85006229	3.019380193
General Motors Company Com (Detroit, MI, USA)	−1	−1	−1	1.059018867	0.970756385	0.859555422
Intl Business Machines Corp Com (Armonk, NY, USA)	−1	−1	−1	2.491670635	2.517623184	2.371848396
Walt Disney Co Com (Burbank, CA, USA)	−1	−1	−1	2.971125915	2.967232846	2.366680775
Cisco Systems, Inc Com (San Jose, CA, USA)	−1	−1	−1	4.320716755	4.433723659	4.063902584
Fomento Econ Mex UBD (Monterrey, Mexico)	−1	−1	−1	11.18952678	11.021056	7.94217879
Cemex CPO (Monterrey, Mexico)	−1	−1	−1	1.417641949	1.464219447	0.997691805
Alfa A (Monterrey, Mexico)	−1	−1	−1	1.564214225	1.582884514	1.176101053
America Movil L (Mexico City, Mexico)	−1	−1	−1	1.563825184	1.585347478	1.338311027
Elektra Gpo (Mexico City, Mexico)	1	0	−1	1.988999187	1.987660035	15.21261735
CCR SA ON (São Paulo, Brazil)	−1	−1	−1	1.898926374	1.7351045	1.392020105
Cemig ON (Belo Horizonte, MG, Brazil)	−1	−1	−1	1.316805993	1.327685596	1.015047328
Sid Nacional ON (São Paulo, Brazil)	−1	−1	−1	0.66058143	0.653956514	0.523369495
CPFL Energia ON (Campinas, Brazil)	−1	−1	−1	2.023567102	2.207687855	1.804125106
Gerdaul PN (Porto Alegre, RS, Brazil)	−1	−1	−1	1.204286126	1.261666239	0.954862088
Empresas Cmpc S.A. Ord (Santiago, Chile)	−1	−1	−1	1.954458606	1.904796194	1.660744623
Empresas Copec S.A. Ord (Santiago, Chile)	−1	−1	−1	2.524339455	2.45181136	2.10348955
Enel Americas S.A. Ord (Santiago, Chile)	−1	−1	−1	0.842415216	0.754965233	0.801867622
Enel Generacion Chile S.A. Ord (Santiago, Chile)	−1	−1	−1	1.108866147	1.006618261	0.928291146
Buenaventura Cl (San Isidro, Perú)	0	−1	−1	1.672448948	1.395253837	1.054983844
BRF SA ON (Itajaí, Santa Catarina, Brazil)	1	−1	−1	1.523005691	1.512703362	1.138336218
Ambev S/A ON (São Paulo, Brazil)	−1	−1	−1	9.344636071	7.834986815	4.914414625

Source: Own elaboration with R-project and yahoo finance.

6. Discussion & Conclusions

The paper's main objective is to determine the existence of an inflection point in which the credit risk of a company listed on American Stock Exchanges Prices goes from being minimum (hedge) to very high (Ponzi) by using an adaptation of Merton's default probability methodology, data mining classification tools and Markov chains. We also found transition probabilities for Minsky's classifications and use them to get stationary states that give us an insight into the long-run composition of economic apparatus.

When assessing the default probabilities, compared to existing methods, a determinant feature of this methodology is that our method can be used on any company without regard if it is public or private, previous works are focused on public companies. We also think that is worth to mention that this methodology may be adapted to any probability distribution as long as the researcher can use a European call option; this opens a broad scope of analysis that goes from the closed forms of valuation to the adaptable numerical methods.

The study demonstrated the existence of this inflection point; besides, determined migration thresholds among clusters and a functional proposal to classify firms based on their credit risk.

The Markov chains proved that the number of states determined by the clustering methodology is stable along with the sample because the transition probabilities remain similar in the long-run, giving confidence about the states' long-run distribution and the validity of their derived inferences.

The paper shows a very ductile methodology that does not rely on any distribution assumption or require market data to perform the analysis; it suffices with financial statement information (operating cash flow, long-run debt, and financial expenses).

The methodology also allows us to; naturally, group the firms following a sound economic theory, giving cutting edges to each group. The paper is the first one making such classifications. This unique feature also permits analysis of the behavior of a firm's financial health compared to industry or time. It also permits the decision-makers to analyze the general health of an industry or country (this is a line for future work).

It is essential to mention that the methodology relies on the ergodic assumption to provide trustable long run states. On its current form, the paper does not consider the entities' behavior when granting credits in altered states (credit supply excess or other economic turmoil). This problem may be solved by the inclusion of Poisson jumps (positive or negative) on the Markov chain; we also consider this improvement as a future research line.

Author Contributions: Conceptualization, Methodology, Formal Analysis, Investigation, Writing-Original Draft, Preparation and Writing-Review & Editing, O.G.-V., N.G.-D., S.C.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Minsky, H.P. The Financial Instability Hypothesis 4. The Jerome Levy Economics Institute Working Paper No. 74. 1992. Available online: <https://ssrn.com/abstract=161024> (accessed on 4 March 2021). [CrossRef]
2. Minsky, H.P. Can "it" happen again? A reprise. *Challenge* **1982**, *25*, 5–13. [CrossRef]
3. Dickinson, V. Cash Flow Patterns as a Proxy for Firm Life Cycle. *Account. Rev.* **2011**, *86*, 1969–1994. [CrossRef]
4. Merton, R.C. On the pricing of corporate debt: The risk structure of interest rates. *J. Financ.* **1974**, *29*, 449–470.
5. Aparicio, C.; Gutiérrez, J.; Jaramillo, M.; Moreno, H. Indicadores Alternativos de Riesgo de Crédito en el Perú: Matrices de Transición Crediticia Condicionadas al Ciclo Económico. Superintendencia de Banca, Seguros y Administradoras Privadas de Fondos de Pensiones 2003. pp. 1–17. Available online: https://www.sbs.gob.pe/Portals/0/jer/ddt_ano2013/SBS-DT-001-2013.pdf (accessed on 4 March 2021).

6. Támara-Ayús, A.; Aristizábal, R.; Velásquez, E. Matrices de transición en el análisis del riesgo crediticio como elemento fundamental en el cálculo de la pérdida esperada en una institución financiera colombiana. *Rev. Ing. Univ. Medellín* **2012**, *11*, 105–114.
7. McCulloch, R.E.; Tsay, R.S. Statistical analysis of economic time series via Markov switching models. *J. Time Ser. Anal.* **1994**, *15*, 523–539. [CrossRef]
8. Peña, E.L. *Matrices de Transición del Crédito en Nicaragua*; Banco Central de Nicaragua: Managua, Nicaragua, 2013. Available online: https://www.bcn.gob.ni/estadisticas/estudios/2014/DT-31_Matrices_de_Transicion_del_Credito_en_Nicaragua.pdf (accessed on 4 March 2021).
9. Hamilton, J.D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econ. J. Econom. Soc.* **1989**, *57*, 357–384. Available online: <https://www.jstor.org/stable/1912559> (accessed on 4 March 2021). [CrossRef]
10. Gaitán, M.M.; Flores, M.M. Riesgo Crediticio Bancario: Análisis de la Cartera de Crédito del Banco MGM, SA, Mediante El empleo de Modelos de Transición con Cadenas de MARKOV en el Periodo Terminado 2015. Ph.D. Thesis, Universidad Nacional Autónoma de Nicaragua, Managua, Nicaragua, 2 September 2017. Available online: <https://core.ac.uk/download/pdf/94852394.pdf> (accessed on 4 March 2021).
11. Porras, E.R.; Anchundia, S.; Vieira, M. Competencia, Rivalidad y Entrada del Capital Extranjero en la Banca Venezolana. Observatorio de la Economía Latinoamericana. 2002. Available online: <http://www.eumed.net/cursecon/ecolat> (accessed on 4 March 2021).
12. Yu, F.H.; Lu, J.; Gu, J.W.; Ching, W.K. Modeling Credit Risk with Hidden Markov Default Intensity. *Comput. Econ.* **2018**, *54*, 1213–1229. [CrossRef]
13. Pfeuffer, M.; Reis, G.D. Capturing Model Risk and Rating Momentum in the Estimation of Probabilities of Default and Credit Rating Migrations. *arXiv* **2018**, arXiv:1809.09889.
14. Leung, K.S.; Kwok, Y.K. Counterparty risk for credit default swaps: Markov chain interacting intensities model with stochastic intensity. *Asia-Pac. Financ. Mark.* **2009**, *16*, 169. [CrossRef]
15. Córdova, J.F.; Molina, E.C.; Navarrete, P. Lógica difusa y el riesgo financiero. Una propuesta de clasificación de riesgo financiero al sector cooperativo. *Contad. Y Adm.* **2017**, *62*, 1670–1686. [CrossRef]
16. García, J.C.; Bolívar, H.R.; Vázquez, F.A. Actualización del modelo de riesgo crediticio, una necesidad para la banca revolvente en México. *Rev. Finanzas Y Política Econ.* **2016**, *8*, 17–30. [CrossRef]
17. Kavitha, K. Clustering loan applicants based on risk percentage using K-means clustering. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2016**, *66*, 162–166.
18. Zhu, S.; Chan, J.; Bright, D. Applying Machine Learning for Troubleshooting Credit Exposure and xVA Profiles (16 June 2019). Available online: <https://ssrn.com/abstract=3404863> (accessed on 4 March 2021).
19. Wahyudin, I.; Djatna, T.; Kusuma, W.A. Cluster Analysis for SME Risk Analysis Documents Based on Pillar K-Means. *Telecommun. Comput. Electron. Control.* **2016**, *14*, 674–683. [CrossRef]
20. Ayús, A.L.; Peña, A.M.; Álvarez, Y.C. El análisis factorial y el análisis discriminante en la estimación de la pérdida esperada para una institución financiera. *Rev. Cienc. Estrateg.* **2017**, *25*, 425–440.
21. Lagunas, P.S.; Ramírez, P.J. Expectativas para operaciones financieras en los sectores vulnerables mediante Matrices de Transición. *Rev. Mex. Econ. Y Finanzas* **2017**, *12*, 71–101. [CrossRef]
22. Ferreira, J. Inestabilidad financiera y los ciclos económicos: La contribución de Hyman Minsky. In *Modelos Macroeconómicos de Crecimiento y de Fluctuaciones Cíclicas*; Ferreira, J., Ed.; Universidad Autónoma Metropolitana: Mexico City, Mexico, 1995; pp. 113–123.
23. Angeles, C.G.; Ortiz, J. Dos Crisis Económicas Recientes en México: Causas y Perspectivas. Available online: <https://www.researchgate.net/publication/267220509> (accessed on 26 October 2018).
24. Rodríguez, N.A.; Venegas, M.F. Liquidez y apalancamiento de la banca comercial en México. *Anal. Econ.* **2012**, *27*, 73–96.
25. Berger, A.N.; Udell, G.F. The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *J. Bank. Financ.* **1998**, *22*, 613–673. [CrossRef]
26. Bencivenga, V.R.; Smith, B.D. D. Financial intermediation and endogenous growth. *Rev. Econ. Stud.* **1991**, *58*, 195–209. [CrossRef]
27. Adrian, T.; Shin, H.S. Liquidity, monetary policy, and financial cycles. *Curr. Issues Econ. Financ.* **2008**, *14*, 1–7. Available online: <https://ssrn.com/abstract=1090550> (accessed on 4 March 2021).
28. Claessens, S.; Kose, M.A.; Terrones, M.E. How do business and financial cycles interact? *J. Int. Econ.* **2019**, *87*, 178–190. [CrossRef]
29. Terreno, D.D.; Sattler, S.A.; Pérez, J.O. Las etapas del ciclo de vida de la empresa por los patrones del estado de flujo de efectivo y el riesgo de insolvencia empresarial. *Rev. Dep. Acad. Cienc. Adm.* **2017**, *12*, 22–37.
30. Tudela, M.; Young, G. A Merton-model approach to assessing the default risk of UK public companies. *Int. J. Theor. Appl. Financ.* **2005**, *8*, 737–761. [CrossRef]
31. Dar, A.A.; Anuradha, N.; Qadir, S. Estimating probabilities of default of different firms and the statistical tests. *J. Glob. Entrep. Res.* **2019**, *9*, 27. [CrossRef]
32. Afik, Z.; Arad, O.; Galil, K. Using Merton model for default prediction: An empirical assessment of selected alternatives. *J. Empir. Financ.* **2016**, *35*, 43–67. [CrossRef]
33. Aggarwal, C.C. *Data Classification: Algorithms and Applications*; CRC Press: Boca Raton, FL, USA, 2014.

34. Berkhin, P. A survey of clustering data mining techniques. In *Grouping Multidimensional Data*; Springer: Heidelberg/Berlin, Germany, 2006; pp. 25–71. [[CrossRef](#)]
35. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K. Cluster: Cluster Analysis Basics and Extensions. *R Package Version 2.0.7-1* **2018**, *1*, 56.
36. Leonard, K.; Peter, J.R. Finding groups in data: An introduction to cluster analysis. Probability and Mathematical Statistics. In *Applied Probability and Statistics*; Wiley Series; John Wiley & Sons: Hoboken, NJ, USA, 1990.
37. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C* **1979**, *28*, 100–108. [[CrossRef](#)]
38. Ledolter, J. *Data Mining and Business Analytics with R*; John Wiley Sons Inc.: Iowa City, IA, USA, 2013.
39. Caballero, M.E.; Rivero, V.M.; Uribe, G.; Velarde, C. *Cadenas de Markov. Un Enfoque Elemental*; Series Textos; Aportaciones Matemáticas: Ciudad de México, Mexico, 2017.
40. Norris, J.R. Introduction. In *Markov Chains*; Cambridge University Press: Cambridge, UK, 1998; pp. xiii–xvi.
41. Venegas, M.F. *Riesgos Financieros y Economicos/Financial and Economical Risks: Productos Derivados Y Decisiones Economicas Bajo Incertidumbre*; Cengage Learning: México City, Mexico, 2008.
42. Bolívar-Cimé, A.; Notario, C.; Pérez, A. Modelos de Markov para la trayectoria académica de estudiantes de la UJAT. *Misc. Mat.* **2016**, *62*, 29–43. Available online: https://miscelaneamatematica.org/welcome/default/download/tbl_articulos.pdf2.8c525e9a283566e3.363230332e706466.pdf (accessed on 4 March 2021).
43. FitchRatings. Estudio Sobre Transición e Incumplimiento de Calificaciones Nacionales de Fitch México 2018. Available online: <https://www.fitchratings.com/research/es/structured-finance/transition-study-fitch-national-ratings-mexico-2018-28-03-2019> (accessed on 4 March 2021).