

**MÉTODO PARA LA DETERMINACIÓN DE INSOLVENCIA FINANCIERA A  
PARTIR DE ALGORITMOS DE CLASIFICACIÓN BORROSA  
SUPERVISADA Y NO SUPERVISADA**

Héctor Albeiro Dussan Montoya, Claudia Briceida Coy Coy



UNIVERSIDAD  
La Gran Colombia

Vigilada MINEDUCACIÓN

Maestría en Contabilidad, Facultad de Ciencias Económicas y Empresariales

Universidad La Gran Colombia

Bogotá D.C.

2021

**Método para la determinación de insolvencia financiera a partir de algoritmos de  
clasificación borrosa supervisada y no supervisada**

**Héctor Dussan Montoya, Claudia Briceida Coy Coy**

**Trabajo de Grado presentado como requisito para optar al título de Magister en  
Contabilidad**

**Fabián Alberto Castiblanco Ruiz (Director)**



**Maestría en Contabilidad, Facultad Ciencias Económicas y Empresariales**

**Universidad La Gran Colombia**

**Bogotá D.C.**

**2021**

**TABLA DE CONTENIDO**

<b>Resumen</b> .....	<b>12</b>
<b>Abstract</b> .....	<b>13</b>
<b>Introducción</b> .....	<b>14</b>
<b>1. Marco conceptual</b> .....	<b>22</b>
<b>2. Marco jurídico</b> .....	<b>25</b>
2.1. LEY 1116 INSOLVENCIA .....	25
<b>3. Marco teórico</b> .....	<b>29</b>
3.1. LINEA DE TIEMPO.....	31
3.2. ALGUNOS MODELOS DE PREDICCIÓN DE INSOLVENCIA Y/O QUIEBRA.....	37
3.2.1. Modelos Univariables .....	37
Modelo de Beaver (1966) .....	37
3.2.2. Análisis Discriminante.....	42
Modelo Altman (1968) .....	43
Modelo de Edmister Robert (1972) .....	46
Modelo Deakin (1972, 1976).....	56
Modelo Blum (1978).....	61
3.2.3. Análisis Factorial .....	64

Modelo de Robert Libby (1975) .....	65
Modelo Pinches, Eubank, Mingo y Caruthers (1975).....	68
Modelo de Ibarra (2001) .....	76
3.2.4. Modelos de Puntuación Z .....	81
Modelo Lis (1972) .....	81
Modelo Taffler (1974) .....	82
Modelo Tisshaw (1976) .....	83
Modelo Taffler (1977) .....	84
Modelo del Banco de Inglaterra (1979).....	86
Modelo Mason y Harris (1978).....	88
Modelo de Dataestream (1980).....	89
Modelo Betts y Belhoul (1982 y 1983).....	90
Modelo Hennaway y Morris (1983) .....	92
Análisis de Jackknise .....	95
Modelo Jackknise modelo no cotizado (1983) .....	95
La transformación PAS-score .....	96
El enfoque Z - score en la práctica.....	97
Modelo Hernández (2014) .....	97
3.2.5. Regresión Logística y Binaria Logit - Probit.....	103

Modelo Ohlson (1980).....	104
3.2.6. Predicción de Bancarrota: Una Comparación de Técnicas Estadísticas y de Aprendizaje Supervisado para Computadora. Modelo Z – Score y Procesos Gaussianos .....	107
3.2.7. Modelos y Aplicaciones para Colombia.....	111
Modelo Rosillo (2002).....	111
Modelo de Martínez (2003) .....	113
Modelo Pérez, González, Lopera (2011) .....	116
3.2.8. Redes Neuronales Artificiales (RNA) .....	119
Modelo Pessoa de Oliveira y González Pascual (2016) .....	119
3.2.9. Machine Learning - Aprendizaje Automático .....	128
Aprendizaje no supervisado.....	129
El aprendizaje supervisado .....	138
<b>4 Metodología.....</b>	<b>143</b>
4.1 TIPO DE INVESTIGACIÓN .....	144
4.2 DISEÑO DE LA METODOLOGÍA.....	145
Etapa 1: Diseño de la Muestra.....	150
Etapa 2: Selección y toma de datos .....	154
Etapa 3: Aprendizaje No Supervisado.....	157

Etapa 4: Aprendizaje Supervisado.....	161
Etapa 5: Análisis y Resultados .....	164
<b>5 Conclusiones y trabajos futuros .....</b>	<b>168</b>
<b>6 Bibliografía.....</b>	<b>172</b>

**Lista de tablas**

Tabla 1	<i>Tasas de Error de los modelos de predicción de insolvencia de Altman y Beaver 1968 Modelo de Deakin .....</i>	<i>57</i>
Tabla 2	<i>Resultados promedio de clasificación de los datos del Seguro de Depósito de los Estados Unidos Modelo de Peña, Martínez, Abudu .....</i>	<i>110</i>

**Lista de figuras**

Figura 1	<i>Línea de tiempo Modelos de Insolvencia y/o quiebra</i> .....	32
Figura 2	<i>Lista de Razones Probadas por Beaver</i> .....	39
Figura 3	<i>Lista Reducida de Razones Modelo de Beaver</i> .....	41
Figura 4	<i>Ratios usados por Altman (1968)</i> .....	45
Figura 5	<i>Razones seleccionadas para el modelo de Edmister</i> .....	49
Figura 6	<i>Resumen de Ratios usados por Edmister</i> .....	51
Figura 7	<i>Función Lineal de Edmister</i> .....	52
Figura 8	<i>Valores de los indicadores Relevantes - Edmister</i> .....	56
Figura 9	<i>Razones Financieras usadas en el Modelo de Deakin 1972</i> .....	58
Figura 10	<i>Tasas de error usando funciones discriminantes en datos mixtos Modelo de Deakin</i> .....	60
Figura 11	<i>Razones Financieras usadas en el Modelo de Deakin 1976</i> .....	61
Figura 12	<i>Razones usadas en el Modelo de la Empresa Fallida Blum</i> .....	63
Figura 13	<i>Razones Utilizadas en el modelo de Libby</i> .....	66
Figura 14	<i>Razones Financieras para Empresas Industriales Modelo de Pinches, Eubank, Mingo y Caruthers</i> .....	69
Figura 15	<i>Clasificación Jerárquica de las Razones Financieras para Empresas Industriales Modelo de Pinches, Eubank, Mingo y Caruthers</i> .....	72

Figura 16	<i>Razones Financieras Identificadas como predictores significativos Modelo de Pinches, Eubank, Mingo y Caruthers</i> .....	75
Figura 17	<i>Selección de Razones Financieras Aplicadas al Modelo de Ibarra</i> .....	78
Figura 18	<i>Razones utilizados por el Modelo de Lis (1972)</i> .....	81
Figura 19	<i>Razones utilizadas por el Modelo de Taffler (1974)</i> .....	82
Figura 20	<i>Razones utilizados por Tisshaw (1977)</i> .....	84
Figura 21	<i>Razones utilizadas por Taffler (1977)</i> .....	85
Figura 22	<i>Razones utilizadas por Taffler (1979)</i> .....	87
Figura 23	<i>Razones utilizados por Taffler (1978)</i> .....	88
Figura 24	<i>Razones Utilizadas en el Dataestream</i> .....	89
Figura 25	<i>Razones Utilizadas en el Modelo de Betts y Belhoul</i> .....	90
Figura 26	<i>Razones Finales usados en el Modelo Betts y Belhoul</i> .....	91
Figura 27	<i>Razones utilizados por Hennaway y Morris (1983)</i> .....	92
Figura 28	<i>Razones utilizados por Hennaway y Morris (1983)</i> .....	93
Figura 29	<i>Razones posteriores utilizados por Hennaway y Morris (1983)</i> .....	94
Figura 30	<i>Razones utilizados por Hennaway y Morris (1983)</i> .....	96
Figura 31	<i>Modelo Score Z de Altman para empresas de Capital Abierto Manufactureras Modelo de Hernández</i> .....	99
Figura 32	<i>Modelo Score Z de Altman para Empresas de Capital Cerrado Manufactureras Modelo de Hernández</i> .....	100

Figura 33	<i>Modelo Score Z de Altman para Empresas de Capital Cerrado en General (Manufactureras y no Manufactureras) Modelo de Hernández.</i> .....	101
Figura 34	<i>Puntos de corte o Límites de los Puntajes Z de Altman Modelo de Hernández</i> .....	102
Figura 35	<i>Razones Utilizadas en el modelo de Ohlson</i> .....	105
Figura 36	<i>Razones Financieras Predicción de la Bancarrota Utilizadas por Peña, Martínez, Abudu</i> .....	109
Figura 37	<i>Indicadores utilizados por Rosillo (2002)</i> .....	112
Figura 38	<i>Razones utilizadas en el modelo de Martínez</i> .....	114
Figura 39	<i>Razones Financieras usadas en el Modelo de Pérez, González y Lopera</i> .....	117
Figura 40	<i>Proporción de empresas clasificadas correctamente Modelo LOGIT y PROBIT modelo de Pérez, González y Lopera</i> .....	118
Figura 41	<i>Razones seleccionadas para el modelo estático Pessoa de Oliveira</i> .....	122
Figura 42	<i>Ratios Financieros modelo dinámico Pessoa de Oliveira</i> .....	126
Figura 43	<i>Cluster Fuzzy C-Means</i> .....	136
Figura 45	<i>Distintos vecindarios para diferentes valores de k</i> .....	142
Figura 46	<i>Diseño Metodológico del Modelo de Clasificación de Insolvencia</i> .....	147
Figura 47	<i>Diagrama de flujo de la propuesta procedimiento clasificación de la insolvencia financiera</i> .....	148

Figura 48	<i>Crecimiento de la economía total en Colombia por sectores económicos para el año 2018</i> .....	150
Figura 49	<i>Cantidad de Empresas por Sector Económico con Registro Mercantil</i> .....	151
Figura 50	<i>Muestra Preliminar</i> .....	153
Figura 51	<i>Muestra de Empresas en proceso de Insolvencia</i> .....	153
Figura 52	<i>Muestra Final para el Modelo de Clasificación de Insolvencia</i> .....	154
Figura 53	<i>Lista de Ratios Seleccionados para el Modelo de Clasificación de Insolvencia</i> .....	155
Figura 54	<i>Numero de clústeres vs índices de partición Rstudio</i> .....	159
Figura 55	<i>Gráfica de número óptimo de clústeres para clasificación de insolvencia en muestra de empresas de construcción</i> .....	160
Figura 56	<i>Matrices de confusión para k vecinos más cercanos</i> .....	162
Figura 57	<i>Valores ajustados de los ratios de las nuevas empresas sin clasificar</i> .....	163
Figura 58	<i>Clasificación de las nuevas empresas</i> .....	164
Figura 59	<i>Resultados del Método para la determinación de Insolvencia Financiera Machine Learning</i> .....	165
Figura 60	<i>Grado de pertenencia de las empresas clasificadas incorrectamente</i> .....	166

### **Resumen**

El presente trabajo de investigación tiene como objetivo proponer un método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada. Se lleva a cabo la aplicación y validación del método en el sector económico de la construcción en Colombia, debido al alto porcentaje de empresas acogidas a procesos de insolvencia para los años 2018 y 2019. Las variables de entrada para el método corresponden con indicadores o ratios financieros de cada una de las empresas estudiadas, de acuerdo con su uso y aplicación tanto al contexto colombiano como al sector económico específico. Al final, se validan y contrastan los resultados de la aplicación del método a partir de técnicas propias del machine learning junto con los resultados emitidos por la superintendencia de sociedades de Colombia y su reporte sobre el estado de las empresas en estudio. Con base en lo anterior, se establecen las ventajas del método propuesto.

*Palabras Clave:* clasificación, insolvencia, solvencia, ratios, machine learning, algoritmos.

**Abstract**

The objective of this research is to propose a method for the determination of financial insolvency from supervised and unsupervised fuzzy classification algorithms. The application and validation of the method is carried out in the economic sector of construction in Colombia, due to the high percentage of companies in bankruptcy for the years 2017 and 2018. The input variables of the method correspond to indicators or ratios financial statements of each one of the companies studied, according to its use and application to both the Colombian context and the specific economic sector. In the end, the results of the application of the method are validated and contrasted using machine learning techniques together with the results issued by the Colombian Superintendency of Companies and its report on the status of the companies under study. Based on the above, the advantages of the proposed method are established.

*Keywords:* classification, insolvency, solvency, ratios, machine learning, algorithms.

## Introducción

La insolvencia financiera y el fracaso empresarial han sido motivo de análisis y discusión en las últimas décadas debido al interés particular por pronosticar o predecir síntomas negativos en las empresas, tendencias negativas en términos financieros o resultados económico que a futuro puedan ocasionar o tener incidencia en el fracaso financiero. En este sentido, se comprende el fracaso y/o fallo empresarial como el instante en el cual una empresa entra en un estado de insolvencia financiera, es decir, cuando no posee recursos económicos en el corto plazo que le permita responder con las obligaciones financieras contraídas con acreedores y terceros.

Bajo el anterior marco de referencia, uno de los elementos más relevante en el estudio de la insolvencia financiera corresponde con la búsqueda e identificación de herramientas que permitan de forma anticipada, detectar riesgos que puedan ocasionar fallos económicos, tendencias negativas en términos financieros o, posibles problemas en la toma de decisiones que puedan llevar a una empresa a declararse o ser declarada como insolvente Charpentier, (2014).

Sin embargo, establecer herramientas acertadas, precisas y que se ajusten a las condiciones de cada sector, cada país e incluso cada periodo de tiempo determinado, constituye un constante desafío. Por un lado, el mercado actual presenta amplias y variadas fluctuaciones que hacen del contexto económico, difícilmente predecible. Por otro lado, las políticas y condiciones de cada sector económico se tornan cambiantes y con características muy particulares. De igual forma, en concordancia con el informe emitido por la superintendencia de sociedades en la emisión de la revista No. 4 de 2012, en Colombia hay

una gran cantidad de empresarios que no cuentan con políticas debidamente establecidas y que permitan detectar de forma temprana y, por lo tanto, prevenir o corregir un posible estado de insolvencia Superintendencia de Sociedades, (2012).

Se puede decir que actualmente, no existe un consenso general sobre los elementos, herramientas y procesos que permiten establecer un diagnóstico de la insolvencia en las organizaciones. En este sentido, la divulgación y aplicación de múltiples enfoques, procedimientos e indicadores para el estudio de la insolvencia, develan la complejidad y dificultad en la determinación e incluso, el pronóstico de la insolvencia financiera.

En términos generales, se suele emplear un conjunto de elementos y herramientas provenientes del análisis financiero, como lo son los ratios o razones financieras y los cuales son extraídos tanto de la información proveniente de los estados financieros, como de la información propia de la macroeconómica del sector. Dicha información se ha combinado con dos grandes perspectivas; por un lado, se encuentra las herramientas propias de la estadística y, por otro lado, se ha desarrollado una vertiente con amplio auge y desarrollo que emplea elementos y técnicas de la inteligencia artificial (Min & Jeong, 2009).

En cuanto al enfoque estadístico, han prevalecido el uso del análisis discriminante múltiple para clasificar las empresas en estados como el solvente, el insolvente y un estado intermedio o gris. Dicho enfoque, se ha desarrollado a partir de relaciones lineales entre las variables que mejor discriminan a las empresas en los estados o subgrupos dados y, en general, requieren que las variables empleadas sigan una distribución normal multivariante. En la mayoría de los trabajos encontrados al respecto y que serán expuestos en la siguiente sección, se evidencia que las variables más empleadas en el desarrollo de métodos, han sido

los ratios o razones financieras extraídas de los estados financieros de las empresas y en general suelen funcionar bastante bien en determinados sectores o incluso, intervalos de tiempo específicos.

Por otro lado, diagnosticar o predecir un estado de insolvencia o quiebra puede formularse como un problema de clasificación supervisado o no supervisado, en el contexto del aprendizaje de máquinas (machine learning), según la información disponible (Chen et al., 2001). Es decir, como un problema en el cual se buscan clasificar objetos en clases o subgrupos predefinidos o no, a partir de un conjunto de datos previos. En este sentido, los procedimientos o algoritmos propios del aprendizaje de máquinas no requieren de ciertos supuestos sobre el conjunto de datos, en contraste con algunas técnicas estadísticas.

De acuerdo con lo anterior, se evidencia que el problema de la determinación o pronóstico de la insolvencia financiera, implica establecer un conjunto de etiquetas o subgrupos en los cuales serán ubicadas las empresas de acuerdo con el comportamiento de sus variables. En este sentido, es frecuente considerar subgrupos o etiquetas diferenciadas, en las cuales una empresa no puede pertenecer a dos subgrupos a la vez (por ejemplo, solvente e insolvente). Sin embargo, se ha observado que en ocasiones abordar dicho problema, puede dejar de lado un estudio más profundo de las empresas en particular o, del sector económico en general. Es decir, la categorización de empresas en solvente e insolvente no es una cuestión que siempre tenga sus límites perfectamente definidos. Puede presentarse la situación, y en general es bastante frecuente, que existan empresas que no necesariamente se encuentran de forma exclusiva, en un estado de insolvencia o solvencia, sino que sus indicadores muestran en un tránsito gradual a alguno de los dos estados.

Por lo tanto, debido a la heterogeneidad de las empresas de cada sector y de cada país, sus políticas de financiamiento y su posicionamiento en el mercado, es pertinente considerar un proceso de clasificación borrosa, en el cual, la pertenencia a dos clases, etiquetas o subgrupos relevantes (solvente e insolvente) puede ser gradual y no excluyente. Es decir, considerar el problema desde una perspectiva borrosa, implica que la clasificación se puede establecer considerando diferentes niveles o grados de pertenencia de las empresas a los diferentes estados de insolvencia (o solvencia). Lo anterior, supone una mejora en la caracterización de un sector específico de la economía y, un estudio más detallado de las mismas empresas.

De acuerdo con la perspectiva anterior, se hace relevante y de amplio interés, diseñar un método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada que permita establecer de forma gradual, la clasificación de una empresa en solvente o insolvente e incluso en otros estados intermedios. En este sentido, se acude a las variables propias de cada empresa correspondientes a los ratios financieros y a los elementos proporcionados por la lógica borrosa como herramienta que permite establecer procesos graduales o determinados a través de grados. Vale la pena recordar que, desde la década de los setenta se ha demostrado la utilidad de los indicadores o ratios financieros para detectar la capacidad económica y financiera que poseen las empresas y a través de ellos pronosticar su éxito o fracaso.

Para el diseño y evaluación del método, se ha seleccionado el sector económico de la construcción durante los años 2017 y 2018. Dicho sector, presenta particular interés por ser aquel que, durante el periodo estudiado, ha presentado menor crecimiento y a su vez, ha

reportado un alto número de empresas que se han acogido a procesos de insolvencia bajo la ley 1116 de 2006 en Colombia.

Por lo tanto, en aras de proponer un método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada para las empresas del sector de la construcción pertenecientes al grupo 1 de NIIF plenas individuales, se hace necesario partir del estudio y análisis de los diferentes modelos de insolvencia que se han aplicado hasta el momento y que de alguna manera son marco de referencia para el trabajo que se pretende desarrollar.

Un estudio profundo de los modelos propuestos en la literatura permite establecer la necesidad de un método que integre diversas perspectivas y mayor flexibilidad en los supuestos sobre las variables. Por un lado, como se mencionó anteriormente, cuando se acude a los algoritmos propios del machine learning no se imponen restricciones al comportamiento de las variables como la normalidad o la similitud de las matrices de covarianza. De igual forma, cuando se acuden a tales algoritmos, se puede realizar una combinación de los mismos de forma que se pueden trabajar dos fases. Primero una fase no supervisada donde se determinan la cantidad de grupos a clasificar y posteriormente la clasificación de nuevos elementos.

De igual forma, al emplear la lógica borrosa, la clasificación tanto supervisada como no supervisada permite la clasificación de una empresa en dos o más categorías a la vez. Es decir, bajo ciertas condiciones, algunas empresas pueden presentar rasgos característicos de solvencia e insolvencia de forma simultánea, sin que esto implique que se encuentra exclusivamente en uno de los dos estados. Lo anterior permite en general, detectar posibles

tendencias en el paso de un estado a otro y mejor aún, aquellas empresas cuyas variables no permitan definir la clasificación a un grupo de forma exclusiva, podrán ser caracterizadas en términos de grados, verificando que tanto se encuentran en varios estados.

En conclusión, con el diseño y validación de un método que permita determinar la insolvencia financiera de empresas de un sector específico empleando el machine learning y la lógica borrosa, se busca flexibilizar los supuestos de los modelos estadísticos tradicionales, los cuales en ocasiones no son posibles satisfacer y, considerar el problema de la insolvencia como un proceso gradual y sin límites claramente definidos; no siempre las empresas son completamente solventes o completamente insolventes. Por lo tanto, la pregunta que ha motivado el planteamiento de un método con las anteriores características, se enuncia como: ¿Cuáles son las ventajas de emplear técnicas de machine learning y lógica borrosa en la determinación de la insolvencia financiera de empresas en el sector de la construcción en Colombia?

La anterior problemática, cobra validez en Colombia en tanto, diversos investigadores se han dedicado a probar modelos ya existentes en el contexto del país. Sin embargo, tales modelos, como se enunció anteriormente, presentan amplias variaciones en sus resultados dependiendo de la época o intervalo de tiempo empleado y de los sectores abordados. Se destacan dentro de los estudios consultados los siguientes: *Modelo de Predicción de Quiebras de las Empresas Colombianas en el año 2002* (Rosillo, 2002), este modelo se enfoca en el análisis eficiente de ratios financieros que detectan situaciones de peligro en las empresas y basado en el trabajo de Edward Altman en 1968 a partir del análisis discriminante múltiple (MDA).

*Modelo Determinantes de Fragilidad en las empresas Colombianas* (Martínez, 2003) diseñado para identificar los determinantes de la insolvencia en el año 2001, utilizando la técnica probabilística y estadística de regresión PROBIT.

*Modelos de predicción de la fragilidad empresarial aplicación al caso colombiano para el año 2011* (Pérez et al., 2013), permitió medir de qué manera una empresa entra en quiebra o insolvencia financiera usando los modelos probabilísticos de selección discreta PROBIT y LOGIT.

Considerando tanto los estudios realizados a nivel internacional como nacional y teniendo en cuenta las herramientas de interés a usar, el machine learning y la lógica borrosa, el presente trabajo monográfico establece como objetivo principal definir una propuesta de método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada para empresas del sector de la construcción en Colombia. Para alcanzar dicho objetivo, se establece en primera medida un amplio estado del arte sobre los métodos existentes para determinar la insolvencia financiera mediante ratios financieros. A partir de dicha revisión, se comparan las técnicas o modelos de predicción de insolvencia existentes determinando a su vez, los ratios más relevantes y usados en la literatura referente. Finalmente, se realiza un cruce de información entre las técnicas y ratios que permita integrar un método aplicable al sector económico específico.

Por lo tanto, el presente trabajo se estructura bajo los siguientes apartados. En la sección uno se hace referencia al marco conceptual, abordando aquellos términos que son relevantes en la propuesta y que permiten su comprensión. Posteriormente, en la sección dos,

se trabaja el marco normativo colombiano sobre el cual el estado determina el instante cuando una empresa entra en un estado de insolvencia y sus características.

En la sección tres, se abordan los diferentes modelos estudiados y que dan paso al planteamiento del problema y la justificación del uso de nuevas herramientas. Dicha sección, constituye una gran parte del trabajo en tanto, el análisis de todos los modelos expuestos, permiten corroborar, por un lado, los ratios más efectivos y empleados en la literatura y a su vez, la inexistencia de una propuesta como la desarrollada en el trabajo. Por otro lado, permite establecer cuáles son las mejoras que de forma sucesiva han establecido los autores para ganar precisión en la determinación y pronóstico de la insolvencia. En la misma sección, se aborda los elementos fundamentales y empleados concernientes a los algoritmos de clasificación supervisada y no supervisada borrosa.

En la sección cuatro, se presenta el desarrollo metodológico que permitió el planteamiento del método y se describe su funcionamiento. De igual forma, mediante algunas técnicas seleccionadas, se validan los resultados. Finalmente, en la sección cinco se presentan los resultados y conclusiones de la implementación en el sector específico.

## 1. Marco conceptual

En la presente sección se presentan y explican los conceptos más relevantes dentro del trabajo de investigación.

El diccionario de la Real Academia Española (RAE por sus siglas) define la palabra insolvencia como la falta de solvencia, incapacidad de pagar una deuda, en el campo empresarial la insolvencia es la imposibilidad que tiene una empresa para pagar una obligación de carácter financiero contraída con terceros por ejemplo un banco (RAE, 2020). En otras palabras, es una situación financiera de un negocio que le impide cumplir con sus obligaciones cuando ellas son exigibles o mientras maduran en el tiempo, la insolvencia financiera corresponde con el instante en el cual los pasivos totales de una empresa exceden el valor total de los activos, siendo una de las causas más frecuentes del fracaso empresarial.

(Beaver, 1966) define el fracaso empresarial como “la incapacidad de una empresa para solventar sus obligaciones financieras a medida que maduran” (p. 71), operacionalmente una empresa fracasa en los siguientes escenarios; insolvencia, quiebra, incumplimiento de bonos, una cuenta bancaria sobregirada o fallos en el pago de un dividendo de acciones preferentes. Son “cuatro términos los que se encuentran comúnmente en la literatura referentes a fallos empresariales: insolvencia, fracaso, incumplimiento y bancarrota”.

El fracaso es un efecto adverso, se da cuando el porcentaje de la utilidad del monto invertido, es menor al retorno de la inversión, de igual manera han sido usados diferentes criterios económicos, como los ingresos escasos para cobijar costos. La insolvencia ocurre cuando una organización no está en capacidad de responder las deudas a corto plazo,

traducido en una falta de liquidez, es así como se da origen al incumplimiento de obligaciones generándose de esta manera acciones legales para el cobro efectivo; por otro lado la bancarrota empresarial se refiere a que una empresa no puede cumplir con sus acreencias, y que los signos de esto se dan antes de que la empresa pueda continuar con sus operaciones, o que esta sea intervenida (Altman & Hotchkiss, 2011).

Por otro lado “las causas del fracaso de un negocio, y consecuentemente de una bancarrota, pueden dividirse en: económicas, financieras, negligencia gerencial, fraude, desastres naturales, entre otros” (Leaño, 2004, p. 30). Para la presente investigación, es importante aclarar que el foco de estudio está en el fracaso empresarial relacionado con insolvencia financiera principalmente.

Los ratios también conocidos como razones o indicadores financieros, son usados para el análisis de la eficiencia financiera de una empresa en un periodo determinado, para esto por lo general, se asocian mediante una división o cociente los elementos y/o subelementos pertenecientes a los estados financieros, para lo cual los ratios más empleados constan de su correspondiente interpretación. Su principal enfoque está orientado al otorgamiento de créditos por parte de entidades financieras, sin embargo, también son útiles para la toma de decisiones, en el análisis y planeación financiera de una empresa y como predictores de fallos, síntomas o efectos negativos asociados al fracaso empresarial (Beaver, 1966).

Por otro lado, una de las técnicas usadas en la minería y análisis de datos corresponde con el aprendizaje de máquina (machine learning), lo cual hacen referencia a una serie de algoritmos o secuencias de instrucciones ordenadas que hacen que un dispositivo o una aplicación sean artificialmente inteligentes, es decir, estos dispositivos se entrenan para

realizar algunas tareas propias de los seres humanos, sobre todo, en el ámbito de la toma de decisiones. (Banco Bilbao Vizcaya Argentaria [BBVA], 2019).

La base del aprendizaje automático son los algoritmos que se pueden entrenar, ya que al ser una secuencia ordenada y lógica de pasos finitos bajo los cuales se comparan los elementos de entrada, se facilitan procesos enfocados a cumplir objetivos y resolver problemas difíciles que evitan el sesgo humano. Los algoritmos de machine learning no requieren un diseño detallado explícito o restricciones sobre las variables de entrada. Por su parte, entre más grande sea el conjunto de datos, más precisos serán los resultados. (Rebala et al., 2019).

Para el machine learning se destacan dos tipos de aprendizaje que pueden ser simulados por las computadoras, aunque existen más tipos de aprendizaje, los más ampliamente difundidos y tratados son: el aprendizaje no supervisado y el aprendizaje supervisado. El aprendizaje no supervisado, parte de una muestra de datos sin clasificar o etiquetar y su propósito es encontrar los subgrupos o clústeres con los elementos más similares entre sí y más disimilares entre los grupos. Se distingue del aprendizaje supervisado por el hecho de que no hay un conocimiento a priori de la muestra de datos, es decir, el aprendizaje supervisado parte de una clasificación previa o a priori de los datos, se emplean datos ya clasificados o divididos en subgrupos, sobre los cuales el algoritmo *aprende* el comportamiento de las variables. Así, el aprendizaje no supervisado típicamente trata los datos de entrada como un grupo de variables aleatorias, bajo el cual se busca establecer un modelo de distribución que permita de alguna manera su agrupación o clasificación (García & Gómez, 2006).

Un problema al utilizar el aprendizaje no supervisado corresponde a la existencia de datos poco precisos, bajo los cuales se hace complejo identificar similitudes, tendencias y patrones dificultando el agrupamiento de los datos. Por lo tanto, se suele recurrir a herramientas como la lógica difusa debido a su flexibilidad y tratamiento gradual en los valores de las variables de entrada. (Rojas et al., 2009).

## **2. Marco jurídico**

En la presente sección se abordan las generalidades de la ley 1116 de 2006, bajo la cual el estado colombiano define la insolvencia empresarial, los procesos que se deben seguir cuando una empresa se encuentra en dicho estado y, las diferencias en procedimientos que a nivel local se establecen para las empresas con dificultades financieras.

### **2.1. Ley 1116 insolvencia**

La predicción de insolvencia en las empresas colombianas es un tema que no ha tenido gran importancia y requiere en particular, un análisis para poder detectar cuáles empresas tiene la oportunidad de recuperarse. En este sentido, con la Ley 1116 de 2006, se da uno de los primeros pasos para abordar esta problemática gracias a que a partir de dicha ley, nace el proceso de reorganización empresarial, el cual pretende a través de acuerdos con la gerencia, “preservar empresas viables y normalizar sus relaciones comerciales y crediticias mediante la reestructuración operacional y administrativa de los activos o pasivos” (L. 1116, art. 1, 2006).

El marco legal colombiano establece que:

Estarán sometidas al régimen de insolvencia las personas naturales comerciantes y las jurídicas no excluidas de la aplicación del mismo, que realicen negocios permanentes en el territorio nacional, de carácter privado o mixto. Así mismo, estarán sometidos al régimen de insolvencia las sucursales de sociedades extranjeras y los patrimonios autónomos afectos a la realización de actividades empresariales. El Gobierno Nacional establecerá los requisitos de admisión de dichos patrimonios autónomos al trámite de insolvencia a que se refiere la presente ley (L. 1116, art. 2, 2006).

La ley 1116 de 2006 determina las condiciones que debe cumplir las empresas en situación de insolvencia y/o bancarrota para poder evitar la declaración legal de quiebra, dando la posibilidad de realizar acuerdos entre los acreedores y la empresa. Por lo tanto, se tratar de mejorar la situación financiera y evitar posibles cierres y liquidaciones; la reorganización empresarial o judicial es una figura establecida en el régimen de insolvencia de dicha ley.

El régimen de insolvencia económica, es una herramienta que tiene como fin brindar una ayuda a través de mecanismos especializados a aquellas empresas que se encuentran en situaciones económicas críticas y que ven un deterioro en el flujo de efectivo. En Colombia, el estado brinda soluciones y apoyo, pero se hace necesaria la colaboración de los empresarios, pues suele suceder que éstos dejan de interesarse por el pago de gastos de administración generando así una liquidación por incumplimiento de los acuerdos contraídos con terceros.

Dentro de los procedimientos concursales en el régimen de insolvencia empresarial se encuentran el *procedimiento de reorganización*, este procedimiento aplica para empresas que se consideren viables y se inicia con la voluntad del deudor por cumplir las obligaciones con los acreedores. El deudor se compromete con todos los acreedores en el pago de las obligaciones morosas y se establecen formas de financiación de las deudas mientras ordena sus finanzas. Para tal fin, los acreedores establecen por votación la aceptación o no del deudor al proceso de reorganización

El *procedimiento de liquidación judicial* se inicia por incumplimiento del acuerdo de reorganización, de un acuerdo de reestructuración o por solicitud del deudor por una de las causales de liquidación inmediata previstas en la ley, conlleva a la venta de los activos del deudor o de la empresa en marcha. De no ser posible la venta, se realizará la adjudicación de bienes del deudor y en todo caso conlleva el pago de las acreencias hasta donde sea posible con la consecuencial extinción de la persona jurídica y el descargue de obligaciones para la persona natural comerciante

Por otro lado, se encuentra el *procedimiento de liquidación por adjudicación*, este hace referencia a que el patrimonio del deudor se transfieren a los acreedores incluso por el valor total de las obligaciones del deudor y conlleva a la pérdida de la persona jurídica o la liberación de deberes de la persona natural comerciante. A raíz de la emisión del Decreto ley 560 de abril de 2020 “Por el cual se adoptan medidas transitorias especiales en materia de procesos de insolvencia, en el marco del Estado de Emergencia, Social y Ecológica”, este procedimiento queda suspendido por dos años e introduce:

Dos nuevas herramientas para las reorganizaciones empresariales una de carácter extra-judicial, que permite la solución pronta y adecuada de la crisis de la empresa y la atención de diferentes sectores de la economía, estos involucran las negociaciones de emergencia de acuerdos de reorganización y el procedimiento de recuperación empresarial en las cámaras de comercio (Cvalora, 2020, párr. 4).

La otra de carácter judicial que permite que las empresas con deudas morosas y problemas financieros puedan continuar su operación, mientras ordena sus finanzas adquiriendo para dicho fin facilidades de pago adeudas por medio de un acuerdo alcanzado con los acreedores, a diferencia “del proceso de liquidación judicial que persigue la liquidación pronta y ordenada, a través del aprovechamiento del patrimonio del deudor” (L. 1116, 2006).

### 3. Marco teórico

En la presente sección se establece un recorrido histórico sobre los distintos modelos de predicción de insolvencia financiera en su mayoría, a nivel internacional, estableciendo sus dificultades, mejoras y avances. A partir de dicho recorrido, se logra evidenciar que los algoritmos propuestos por el *machine learning*, permiten en gran medida el tratamiento de un mayor número de variables y no exigen el cumplimiento de supuestos o comportamientos sobre las mismas. Sin embargo, debido a las marcadas diferencias entre empresas, sectores y países, se evidencia que un método que incluya elementos de la lógica borrosa, permite una mayor flexibilidad e interpretabilidad en la determinación de la insolvencia financiera, es decir, un estado de insolvencia financiera no siempre tiene sus límites claramente definidos y es imprescindible determinar los posibles cambios de un estado a otro.

James Horrigan fue uno de los primeros investigadores en realizar un estudio referente al origen e historia del análisis de los ratios financieros, para él “las primeras causas del análisis de los estados financieros se remontan a las últimas etapas del impulso de Estados Unidos hacia la madurez industrial en la última mitad del siglo XIX” (Horrigan, 1968, p.284), esto debido a que los emprendedores capitalistas dueños de empresas industriales delegaron la gestión administrativa y financiera a profesionales, sin embargo para esta fecha el análisis financiero aun no planteaba “los factores comunes de los procesos de fallo ya que el fracaso se entendía como un evento súbito causado en principio directamente por las particularidades de cada empresa, no generalizables a otras organizaciones” (Gordon, 1971, p.3). El análisis enfocado a la prevención y/o predicción prematura de la insolvencia financiera toma

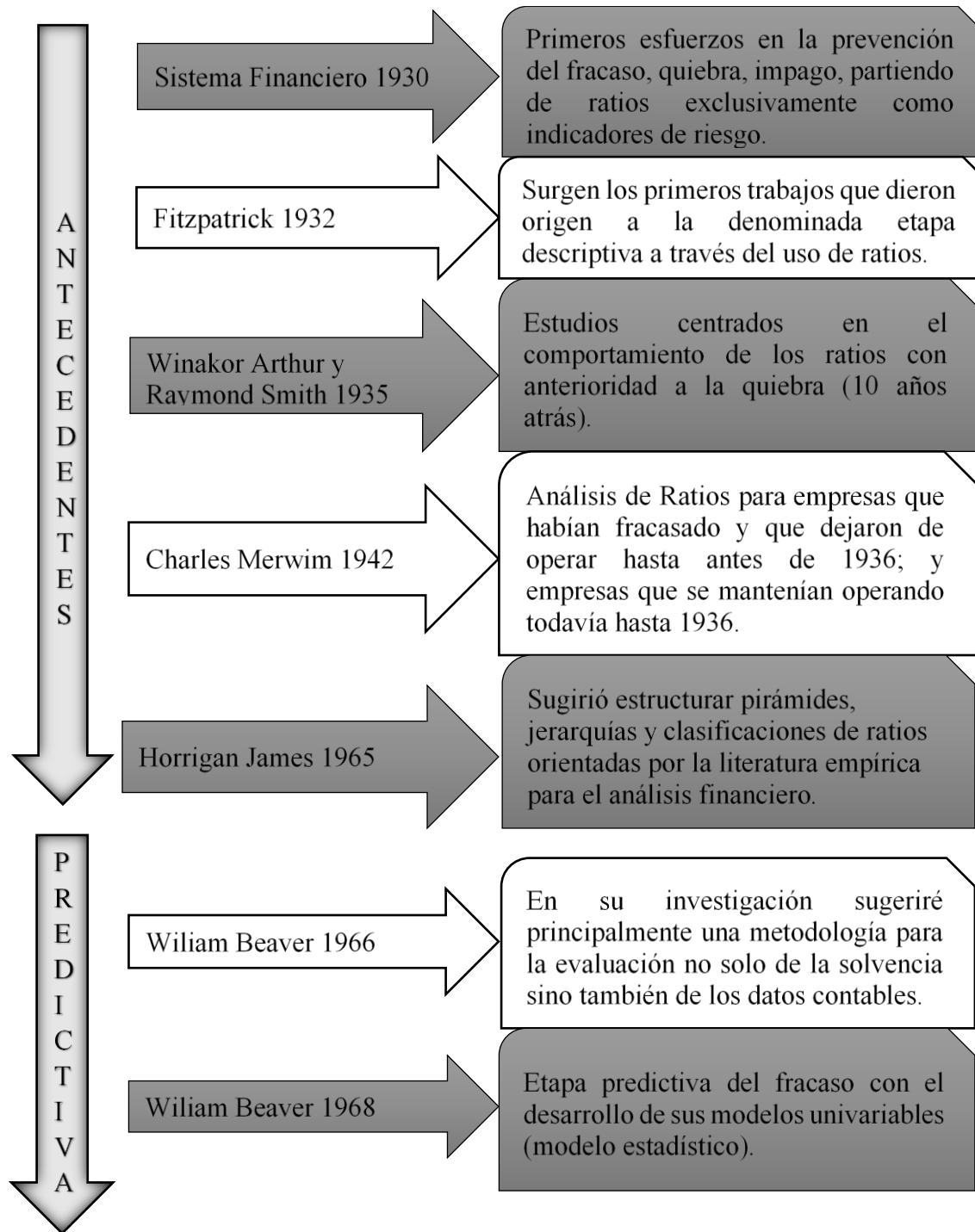
relevancia en la investigación en la década de 1960 sin embargo, hay algunos antecedentes de aproximadamente 30 años antes.

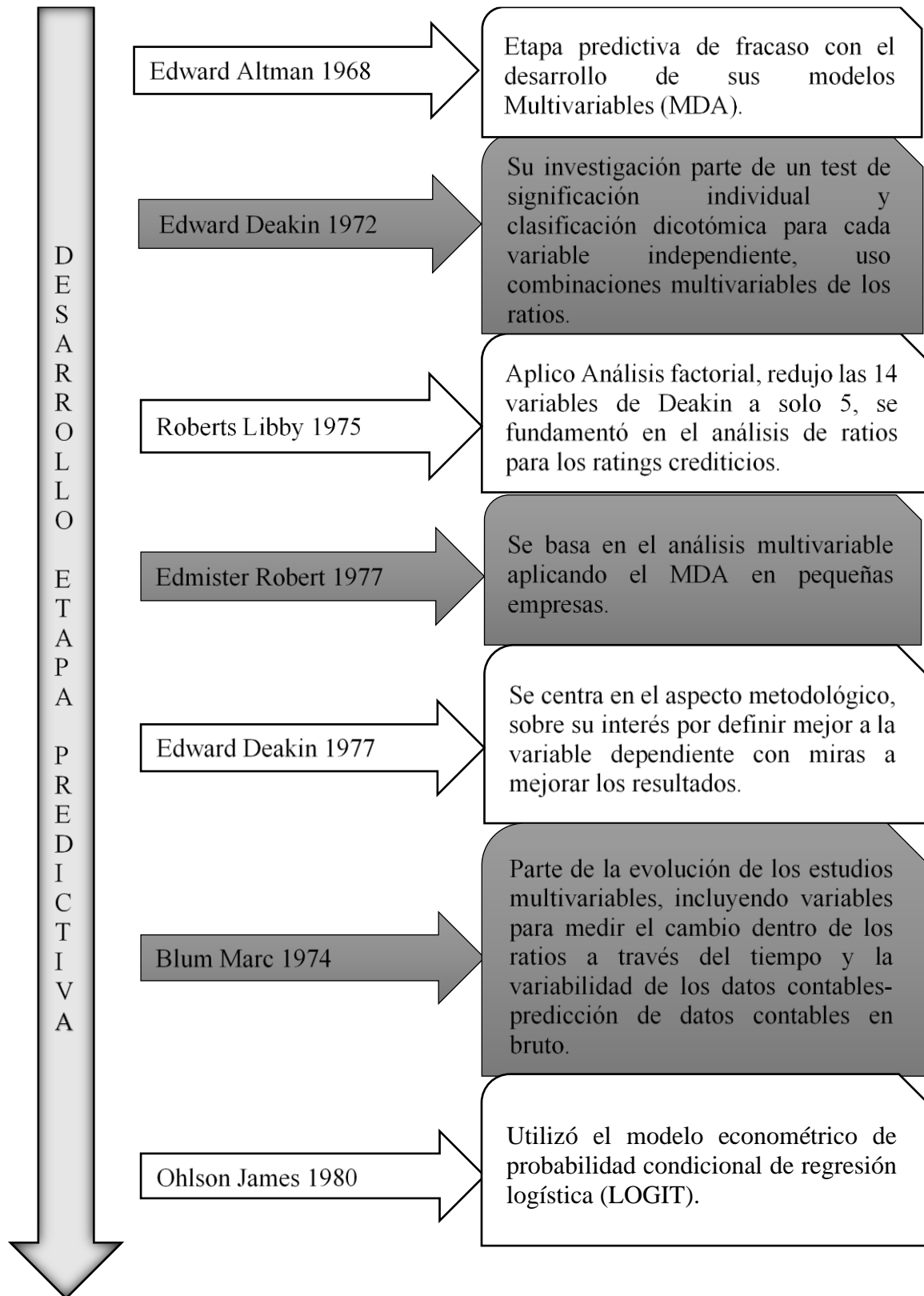
Horrigan, (1968) en su trabajo *A short history of financial ratio analysis*, agrupa algunos antecedentes, autores e investigaciones referentes al análisis de ratios financieros los cuales se presentarán a continuación: Uno de los primeros trabajos fue el de Fitz Patrick quien publicó para el año de 1932, un estudio de 19 pares de empresas categorizadas en fallidas y no fallidas. Sus resultados indicaron que hubo diferencias persistentes en las proporciones durante al menos 3 años antes del fracaso.

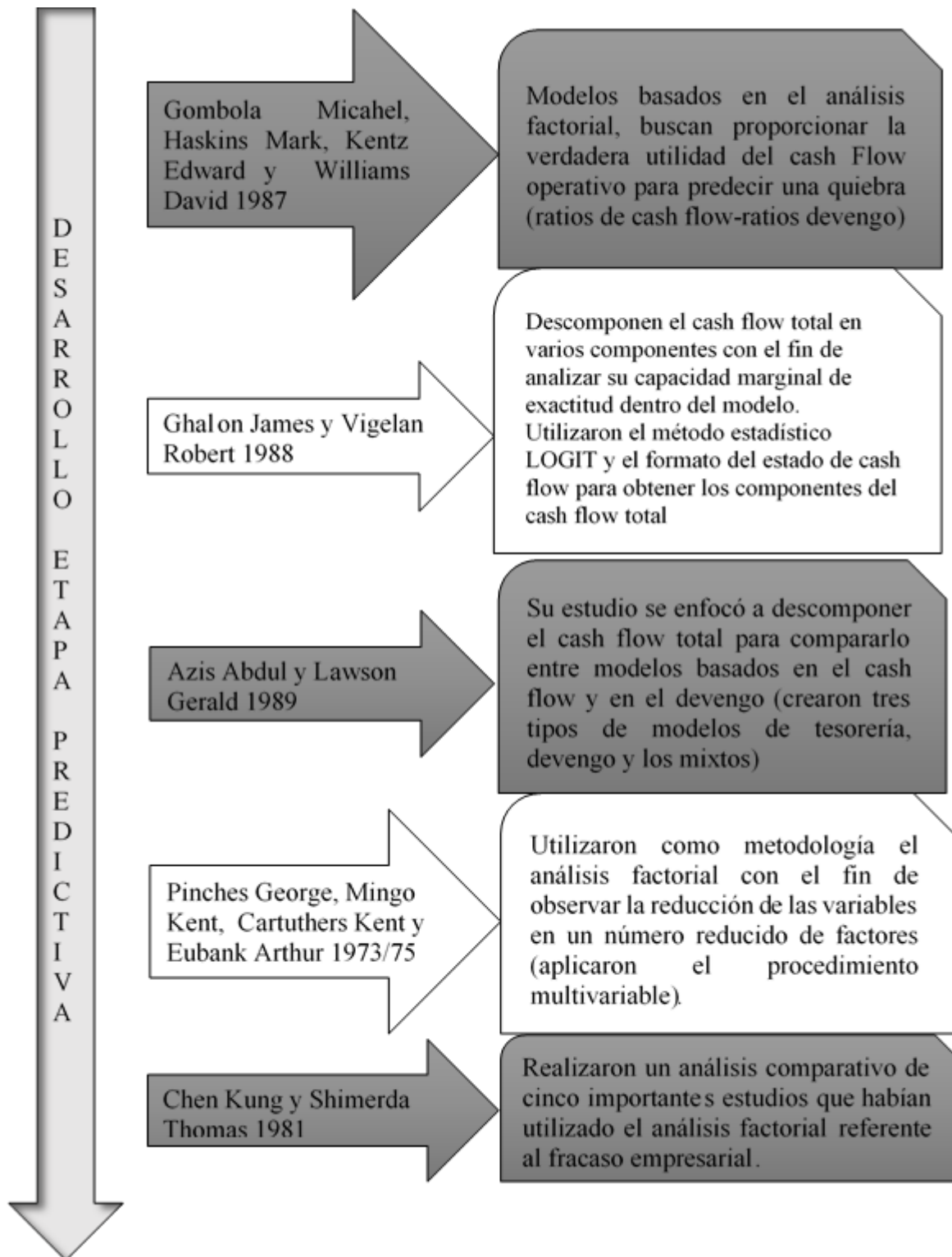
Posteriormente Winakor y Smith en 1935 investigaron la tendencia de las medias de 21 ratios de empresas fallidas durante 10 años antes del fracaso, se dieron cuenta que el deterioro de las tasas en los valores medios aumentaba a medida que se acercaba el fracaso, concluyeron que el ratio capital de trabajo neto sobre activos totales era el indicador más preciso y constante de fracaso y Merwin en 1942 analizó la tendencia de los ratios y comparó las proporciones medias de las empresas continuas con las de las empresas descontinuadas durante un periodo de años entre 1926 y 1936, concluyó que los ratios capital de trabajo neto sobre activos totales, patrimonio Neto sobre deuda y razón circulante o corriente eran predictores muy sensibles de discontinuación como se cita en Horrigan, (1968).

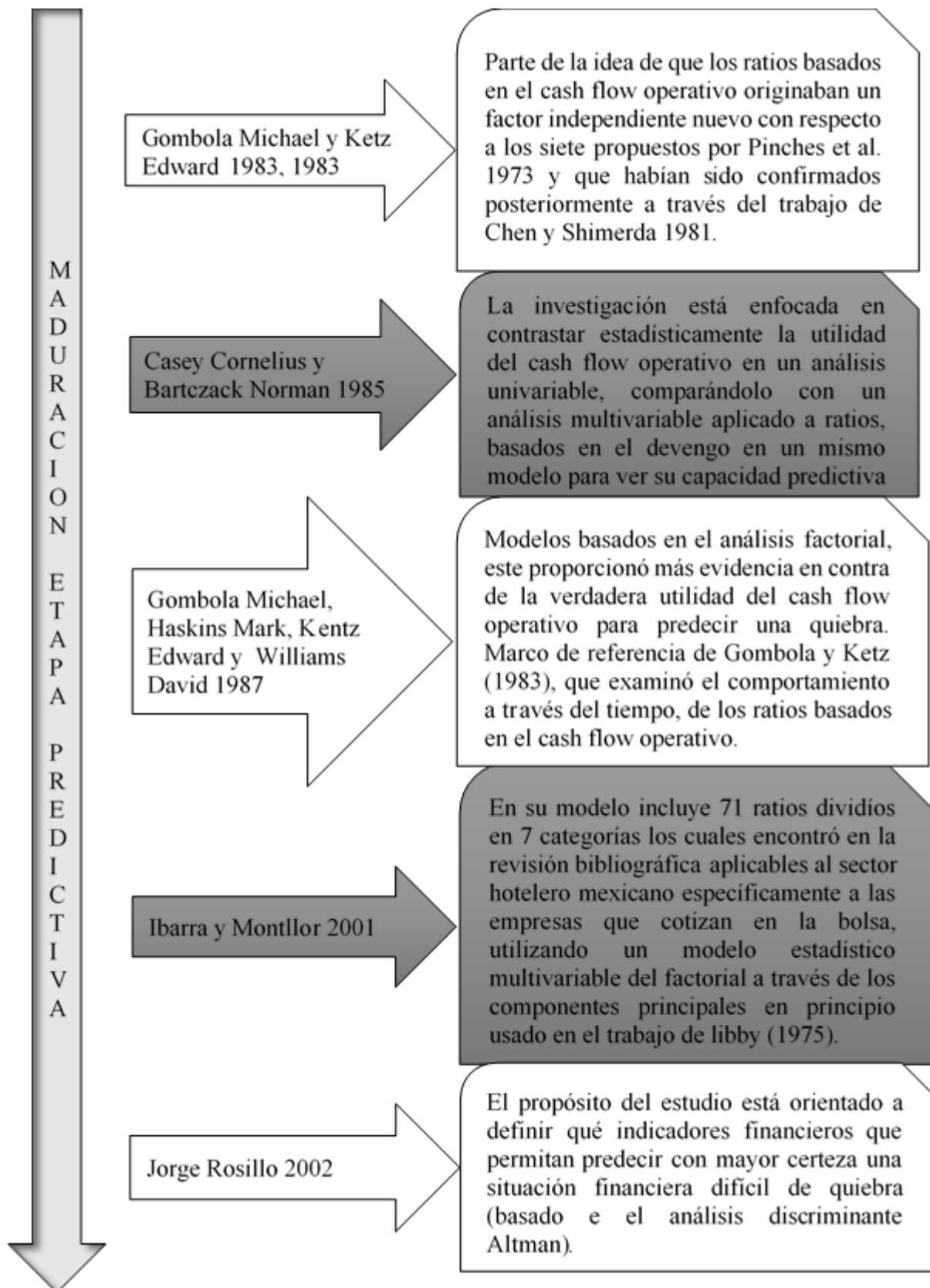
### **3.1. Línea de Tiempo**

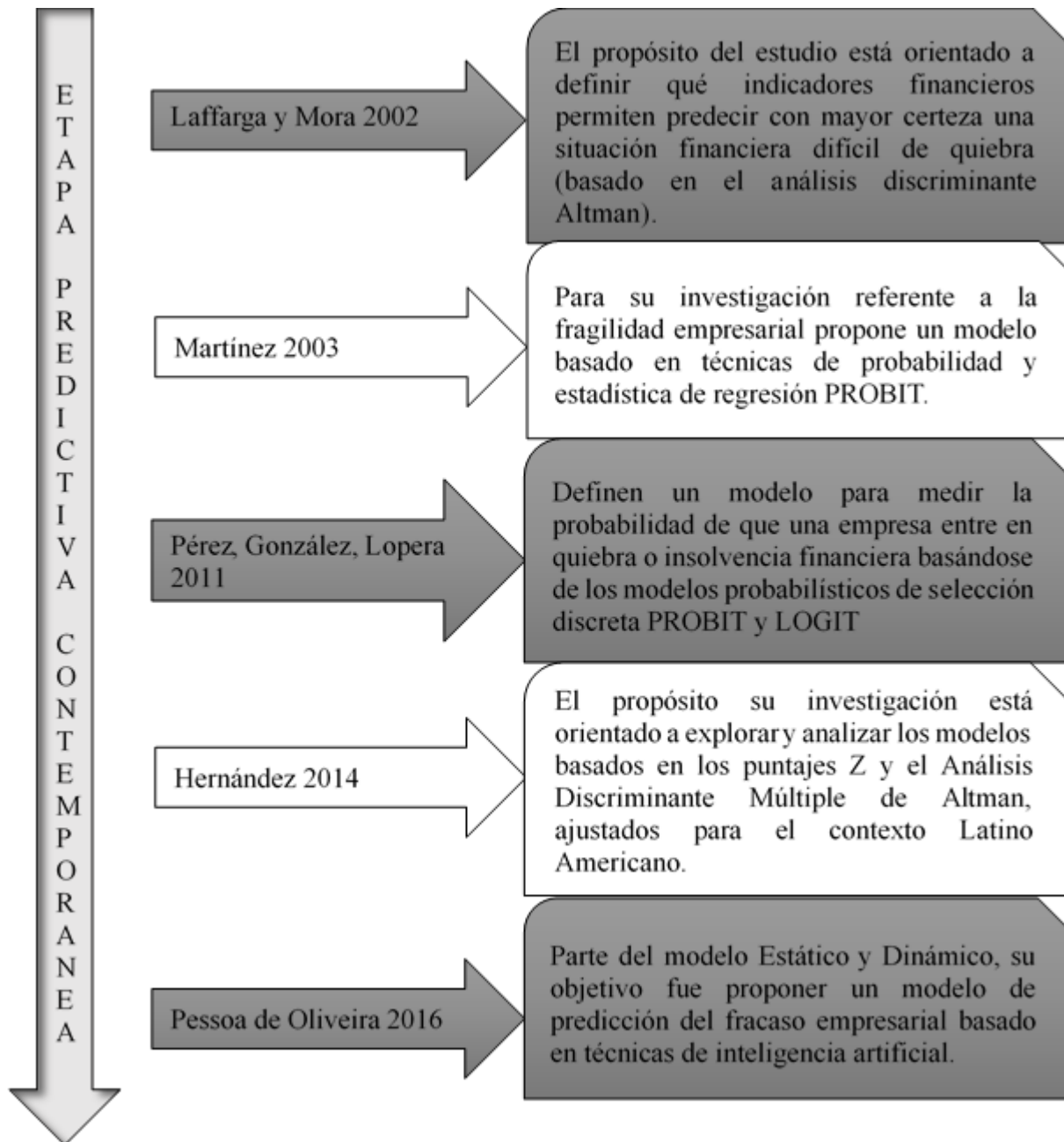
A continuación, se presenta una línea de tiempo cronológica referente a los modelos de predicción de fracaso empresarial desde sus antecedentes en el año 1930 hasta el 2016, adicionalmente se presenta una pequeña descripción referente a los modelos. Es importante aclarar que existen un sin número de modelos relacionados, pero para la investigación fueron considerados algunos de los más tradicionales o conocidos.

**Figura 1***Línea de tiempo Modelos de Insolvencia y/o quiebra*









Elaboración Propia

### 3.2. Algunos Modelos de Predicción de Insolvencia y/o Quiebra

#### 3.2.1. Modelos Univariables

El análisis univariable estudia el comportamiento de una variable de forma individual, este análisis al igual que otras técnicas que se abordaran más adelante es utilizado en los modelos para revisar si se pueden relacionar los valores de una variable con los valores de otra mediante una función de la siguiente manera.

$$Y = F(X)$$

En las funciones existen 2 tipos de variables, por un lado, los valores de (X) representan las variables independientes y por el otro los valores de (Y) representan la variable dependiente, como se puede observar en la función los valores de (Y) guardan una relación de dependencia de los valores de (X). En los modelos de predicción de insolvencia y/o quiebra el análisis univariable es utilizado para estudiar el comportamiento individual de los ratios financieros como variables independientes en función de la solvencia o liquidez como variable dependiente a fin de establecer una herramienta que permita predecir prematuramente el fracaso empresarial por insolvencia financiera o quiebra.

#### ***Modelo de Beaver (1966)***

Uno de los primeros modelos de predicción de insolvencia financiera univariable fue desarrollado por Beaver, (1966) en su estudio titulado *Financial Ratios as Predictors of Failure*, parte del paradigma de la utilidad de la información contable y los ratios como

predictores del fracaso empresarial, utilizó métodos estadísticos para descomponer los ratios financieros.

Para su investigación seleccionó una muestra de empresas fallidas que se emparejaron con empresas no fallidas entendidas como Empresas Sanas lo que quiere decir que no tienen problemas de insolvencia o quiebra y gozan de buena situación financiera, agrupadas según su industria y tamaño de los activos, esta información fue validada si las parejas de firmas estaban en la agencia de calificación de riesgo Moody's, para el análisis y evaluación de la información financiera y ratios utilizó las siguientes tres técnicas: La comparación de medias, pruebas de clasificación dicotómica y un análisis de probabilidad de los ratios.

La información obtenida para la muestra de empresas recopila estados financieros de cinco años antes del fracaso desde 1954 hasta 1964 para las 79 firmas fallidas que tuvieron problemas de insolvencia y cayeron en quiebra, la revisión bibliográfica de su investigación aporta para la selección de la variable independiente 30 ratios divididos en 6 grupos como se puede observar en la Figura 2 es importante destacar que la escogencia de los ratios se hizo según la frecuencia o popularidad con la que aparecían en la literatura y el desempeño superior de los mismos en estudios previos.

**Figura 2***Lista de Razones Probadas por Beaver*

<p><b>Grupo 1 - Flujo de Caja – Liquidez</b></p> $\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$ $\frac{\text{Flujo de Caja}}{\text{Ventas}}$ $\frac{\text{Flujo de Caja}}{\text{Activos Totales}}$ $\frac{\text{Flujo de Caja}}{\text{Patrimonio Neto}}$ <p><b>Grupo 2 - Utilidad</b></p> $\frac{\text{Utilidad Neta}}{\text{Ventas}}$ $\frac{\text{Utilidad Neta}}{\text{Activos Totales}}$ $\frac{\text{Utilidad Neta}}{\text{Patrimonio Neto}}$ $\frac{\text{Utilidad Neta}}{\text{Pasivo Total}}$ <p><b>Grupo 3 - Pasivo / Activo – Cobertura</b></p> $\frac{\text{Pasivos Corrientes}}{\text{Activos Totales}}$ $\frac{\text{Pasivos No Corrientes}}{\text{Activos Totales}}$ $\frac{\text{Pasivo Total}}{\text{Activos Totales}}$ $\frac{\text{Pasivo Total} + \text{Acciones Preferentes}}{\text{Activos Totales}}$ <p><b>Grupo 4 - Activos Líquidos / Activos</b></p> $\frac{\text{Efectivo}}{\text{Activos Totales}}$ $\frac{\text{Activos Disponible}}{\text{Activos Totales}}$	<p><b>Grupo 5 - Activos Líquidos / Deuda Corriente</b></p> $\frac{\text{Activo}}{\text{Pasivos Corrientes}}$ $\frac{\text{Activo Disponible}}{\text{Pasivos Corrientes}}$ $\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$ <p><b>Grupo 6 Rotación</b></p> $\frac{\text{Efectivo}}{\text{Ventas}}$ $\frac{\text{Cuentas por Cobrar}}{\text{Ventas}}$ $\frac{\text{Inventario}}{\text{Ventas}}$ $\frac{\text{Activo Disponible}}{\text{Ventas}}$ $\frac{\text{Activos Corrientes}}{\text{Ventas}}$ $\frac{\text{Capital de Trabajo Neto}}{\text{Ventas}}$ $\frac{\text{Patrimonio Neto}}{\text{Ventas}}$ $\frac{\text{Activos Totales}}{\text{Ventas}}$ $\frac{\text{Efectivo}}{\text{Necesidades Operativas de fondos}}$ $\frac{\text{Activos defensivos}}{\text{Necesidades Operativas de Fondos}}$ $\frac{\text{Activos Defensivos} - \text{Pasivos corrientes}}{\text{Necesidades Operativas de Fondos}}$
---	--

Adaptada de: “Financial Ratios as Predictors of Failure”, por W. Beaver, 1966.

<https://doi.org/10.2307/2490171>

En el modelo Beaver, (1996) se calcularon los valores medios o media de las razones para las empresas fallidas y para las empresas no fallidas cinco años antes del fracaso. La comparación de los valores medios se denominó análisis de perfil. La técnica de análisis de perfil no es una prueba predictiva, se usa simplemente para delinear las relaciones generales entre las empresas fallidas y no fallidas. -Se utilizó también la técnica de clasificación dicotómica para agrupar las empresas entre fallidas y no fallidas al considerar solo esos dos posibles estados para las empresas siendo esta una prueba predictiva, es decir permite predecir el estado de fracaso de una empresa, empleando únicamente el análisis de los ratios financieros, aunque no proporcionó tanta información como el análisis de probabilidad de los ratios.

El análisis de probabilidad de ratios reduce a seis (6) los treinta (30) recopilados, seleccionados en función del menor porcentaje de error de clasificación para su grupo durante el período de cinco años, el ratio con mejores resultados en conjunto fue la relación cash flow o flujo de caja dividido en el pasivo total, con un error en la clasificación del 13% el último año previo al fracaso, en la siguiente Figura se puede observar los 6 ratios ordenados del menor al mayor error de clasificación.

**Figura 3***Lista Reducida de Razones Modelo de Beaver*

<b>Grupo 1 - Flujo de Caja – Liquidez</b>	<b>Grupo 4 - Activos Líquidos / Activos</b>
$\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$	$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$
<b>Grupo 2 - Utilidad</b>	<b>Grupo 5 - Activos Líquidos / Deuda Corriente</b>
$\frac{\text{Utilidad Neta}}{\text{Activos Totales}}$	$\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$
<b>Grupo 3 - Pasivo / Activo – Cobertura</b>	<b>Grupo 6 Rotación</b>
$\frac{\text{Pasivo Total}}{\text{Activos Totales}}$	$\frac{\text{Activos Defensivos – Pasivos Corrientes}}{\text{Necesidades Operativas de Fondos}}$

*Nota:* Flujo de Caja es igual a la Utilidad Neta + Amortizaciones, Activo defensivo igual al Activo Corriente – Inventario, Necesidades Operativas es igual a los Gastos de Explotación – Amortizaciones y Provisiones. Adaptado de: “Financial Ratios as Predictors of Failure”, por W. Beaver, 1966.  
<https://doi.org/10.2307/2490171>

El análisis examina la capacidad de pronóstico de los ratios, uno a la vez, aunque los resultados no fueron muy alentadores en términos del mejor predictor para el fracaso, la finalidad del modelo fue evaluar la capacidad predictiva de los ratios, el deterioro de los ratios durante el periodo de estudio y que la razón cash flow sobre activo total puede dar signos estadísticos altos de causa de insolvencia y quiebra Charpentier, (2014), el ratio cash flow entre activo total substituyó al ratio cash flow entre pasivo total para el análisis de pares.

Uno de los principales problemas del modelo, es que las empresas en situación de insolvencia financiera o quiebra pueden intencionalmente disfrazar sus estados financieros efecto window dressing afectando los ratios para así no reflejar la situación financiera real de la empresa. Otro problema son los errores de tipo I clasificar una empresa fallida como no

fallida y errores Tipo II clasificar una empresa no fallida como fallida, ya que son probabilidades de error condicionadas al estado real de la empresa, es importante determinar y tener en cuenta la probabilidad de ocurrencia para cada caso.

### **3.2.2. Análisis Discriminante**

El análisis de discriminantes es una técnica estadística desarrollada por Fisher en 1936, se usa para clasificar distintos individuos dentro de una muestra o población, es muy usada en los modelos de predicción de insolvencia y quiebra financiera, “su objetivo es especificar la relación en términos de una función que idealmente separa cada elemento dentro del conjunto de datos” (Peña et al., 2009, p. 4). Es una técnica apropiada para distinguir cuales variables independientes o predictivas permiten diferenciar individuos o grupos dentro de una muestra y cuántas de esas variables son necesarias para conseguir la mejor clasificación posible. Antes de utilizar el análisis discriminante.

Se debe comprobar el cumplimiento de dos supuestos básicos de aplicación. El primero es que las variables independientes deben seguir una distribución normal multivariante y el segundo es que las matrices de covarianzas deben ser iguales en todos los grupos sin embargo, el análisis discriminante no se ve en gran medida afectado si alguno de los anteriores supuestos no se cumple (Torrado Fonseca & Berlanga Silvente, 2013, p. 153).

***Modelo Altman (1968)***

Como consecuencia de los estudios de Beaver, varios investigadores en el tema de insolvencia financiera se enfocaron en la aplicación de “modelos multivariantes” con el propósito de llegar a comprobar con más exactitud qué empresas conducían a la quiebra y cuáles no. Uno de ellos fueron los estudios realizados por Edward Altman en su artículo denominado *Razones Financieras, Análisis Discriminante y Predicción de Bancarrota Corporativa* el objetivo de este artículo fue evaluar la calidad del análisis de razones como técnica analítica, a partir de la predicción de quiebra empresarial. Especialmente el trabajo se enfocó en investigar un acumulado de razones financieras y económicas en un contexto de predicción de quiebra en el que se emplea una técnica estadística de análisis discriminante múltiple (MDA). Los datos que fueron objeto de estudio se limitaron a empresas manufactureras.

Altman, (1968) centra su análisis de ratios tradicionales, y a partir del análisis de varios estudios realizados en la década de 1930 y algunos otros posteriores concluye que las empresas en quiebra muestran medidas de razón significativamente contraria a las de las organizaciones en operaciones continuas, además identifico que estudios de la época centraron su atención en una lista de ratios financieros individualmente para empresas en quiebra y una muestra emparejada de empresas que no lo hicieron; la evidencia observada durante 5 años antes de la falla se cita como concluyente que el análisis de indicadores o razones puede ser útil en la predicción de fracasos.

Por otro lado, centró su atención en el MDA, como la técnica estadística más adecuada y utilizada para clasificar información en una de varias agrupaciones a priori que dependen de

las características individuales de la información; entre otras palabras señala que esta técnica es utilizada para clasificar o hacer pronósticos en problemas donde la variable dependiente aparece en forma cualitativa. Indica que lo primero que se debe hacer es, establecer claramente clasificaciones de grupos (en quiebra/no quiebra). A partir de esta información el MDA tiene una gran ventaja y es que, considera un perfil completo de características comunes que discriminen entre los diferentes grupos analizados, el MDA tiene también la capacidad de analizar todo el perfil variable del objeto en estudio simultáneamente en lugar de examinar secuencialmente sus características individuales.

Bajo estas cualidades descriptivas el autor indica que, se selecciona la técnica del MDA como la más apropiada para el estudio de quiebra, bajo una técnica de análisis de regresión múltiple cuidadosamente diseñada e interpretada. Altman, (1968) seleccionó una muestra inicial de 66 empresas, estas empresas fueron seleccionadas de acuerdo con dos criterios: el tipo de industria y el tamaño de los activos. Adicionalmente estas empresas fueron clasificadas en dos grupos: el primer grupo estaba clasificado por las empresas del sector industrial público americano con un rango de activos entre 700 hasta 25,9 millones de dólares las cuales entraron en bancarrota entre 1946 y 1965, el segundo grupo clasificado por empresas manufactureras americanas con un rango de activos entre 1 y 25 millones de dólares, las cuales estaban solventes.

El trabajo consistió en tomar cada empresa fracasada y emparejarla con otra empresa no fracasada, teniendo en cuenta como aspectos importantes la uniformidad entre su tamaño y sector, pues se buscaba evitar efectos de distorsión de ambos factores en los resultados.

Dentro del proceso de selección de las variables independientes para el estudio tomo un grupo

de 22 ratios los cuales fueron aplicados para ambas submuestras. La selección de ratios se basó fundamentalmente en tres criterios: por un lado la popularidad dentro de la literatura, por su significativa potencial para el estudio y por la forma innovadora en que se mostraron los ratios en el análisis (Altman, 1968).

Otros aspectos importantes para el autor en la predicción de quiebras corporativas fueron: el análisis de la significancia estadística de varias combinaciones de ratios, el análisis de reciprocidades entre ratios, el análisis de la precisión de predicción de varias composiciones de ratios, y, el juicio del análisis sobre los resultados adquiridos. La función discriminante final utilizada fue como se muestra en la Figura 4.

Finalmente, de los 22 ratios se seleccionaron cinco, determinados por factores de rentabilidad, apalancamiento, liquidez, solvencia y actividad como se muestra en la siguiente Figura.

**Figura 4**

*Ratios usados por Altman (1968)*

<b><math>Z = 0.12 X1 + 0.014 X2 + 0.033 X3 + 0.006 X4 + 0.999 X5</math></b>	
<p><b>Grupo 1 – Indicador de Rentabilidad</b></p> $X3 = \frac{\text{Utilidad Antes de Intereses e Impuestos}}{\text{Activos Totales}}$ <p><b>Grupo 2 – Indicador de Actividad</b></p> $X5 = \frac{\text{Ventas}}{\text{Activos Totales}}$ <p><b>Grupo 3 – Indicador Endeudamiento</b></p> $X4 = \frac{\text{Patrimonio}}{\text{Valor Contable Deuda Total}}$	<p><b>Grupo 4 – Indicar de Rendimiento</b></p> $X2 = \frac{\text{Ganancias Retenidas}}{\text{Activos Totales}}$ <p><b>Grupo 5 – Indicador de Liquidez</b></p> $X1 = \frac{\text{Capital de Trabajo}}{\text{Activos Toatales}}$ <p style="text-align: center;"><b>Z = Indice General</b></p>

Adaptado de “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, por E. Altman, 1968. <https://doi.org/10.1111/jofi.12742>

Los resultados obtenidos del MDA, demostró ser ampliamente preciso para predecir correctamente la quiebra en el 94% de la muestra inicial y del 95% de todas las empresas de los grupos en quiebra y no quiebra. Reflejó un error tipo I del 6% y un error tipo II del 3%. La investigación demostró que el modelo puede pronosticar la quiebra con exactitud hasta un año antes de la falla, aunque esta exactitud puede disminuir su porcentaje después del segundo año.

Por otro lado concluye que, de acuerdo con los análisis realizados los ratios de rentabilidad, liquidez y solvencia predominan como los indicadores más relevantes. Estos resultados se dieron “debido a la inclusión de los cinco ratios en la función discriminante con los que Altman construyó la puntuación Z-Score, considerada por varios estudios como uno de los mejores modelos teóricos de predicción de quiebras” (Altman, 1968, p. 609 ).

Finalmente sugiere que las razones deben ser analizadas dentro de un esquema multivariado las cuales brindan mayor importancia estadística, ya que la técnica común de comparaciones secuenciales de razones es poco sofisticada en cuanto a su presentación.

### ***Modelo de Edmister Robert (1972)***

El propósito de la investigación realizada por Robert O Edmister fue probar la utilidad de diferentes ratios financieros desde el análisis discriminante multivariable para predecir el fracaso empresarial específicamente en las pequeñas empresas ya que en estudios anteriores como los de Beaver 1966, Altman 1968 y Deakin 1972 se enfocaron únicamente en el estudio de empresas de tamaño mediano y grande.

Partiendo de los estudios anteriores basados en los fundamentos de Beaver, (1966), con respecto a la importancia estadística para predecir una quiebra, los primeros estudios indicaban que, se podían efectuar combinaciones con pocos ratios con el propósito de construir una función discriminante de considerable grado de exactitud, también consideró que algunos ratios son mejores predictores que otros, y que no siempre los mismos ratios son los mejores predictores en todas las investigaciones (Ibarra, 2001, p. 64).

La técnica estadística de la investigación es la de clasificar empresas individuales ya sea en las del grupo de empresas de éxitos o en el grupo de empresas de fracaso. Los datos para probar las hipótesis fueron tomados de la *Small Business Administration* (SBA) y *Robert Morris Associates* (RMA), con un periodo de tiempo de tres años previos a la quiebra y como unidad temporal el período 1954-1969.

Para el estudio se tuvo en cuenta características relevantes de la información para probar las hipótesis las cuales serán consideradas como parte de los resultados obtenidos, teniendo en cuenta la información recopilada como: 1). Se encuentra disponibles tres estados anuales consecutivos antes de la fecha en que se otorgó el préstamo, 2). Los estados financieros reportan operaciones correspondientes a los años 1958-1965, 3). Los extractos reportan una cantidad distinta de cero para los activos corrientes y las ventas netas, 4). Existen declaraciones compuestas de RMA correspondientes, 5). Existen declaraciones compuestas de la SBA correspondientes, 6). La muestra de prueba contiene un número igual de pérdidas y no pérdidas Edmister, (1972).

Los datos recolectados correspondían a una muestra de 42 empresas clasificadas en dos grupos: el primer grupo de 21 empresas de tamaño pequeño que casualmente habían quebrado y el segundo grupo de 21 empresas de tamaño pequeñas que no estaban en quiebra.

Un aspecto importante que evaluó el autor es que, la información de estas empresas tenía la característica de ser poco confiable dada la falta de obligación para llevar a cabo auditorías a los estados financieros Edmister, (1972). Por este motivo optó por no entrar en el análisis de temas metodológicos, ilustraciones significativas sobre el fracaso como variable dependiente como tampoco en la selección de la muestra, sin embargo advirtió que ante la variable y el tiempo sumado la inestabilidad de algunos parámetros era importante reflejar ambos fenómenos.

Edmister, (1972) para desarrollar su investigación acudió a cuatro hipótesis que fueron las siguientes:

- **La primera hipótesis** establecía que el nivel del ratio es un predictor de quiebra en las pequeñas empresas, sin importar el lugar donde se ubique la firma (sector, país entre otros)
- **La segunda hipótesis** referida al espacio de tiempo (tres años) de cada ratio el cual representa el poder predictivo de fracaso, esta tendencia se definía estadísticamente como la relación significativa de la variable y tiempo.
- **La tercera hipótesis** establecía que el promedio de tiempo de tres años de una razón es un predictor más efectivo con respecto a un solo promedio anual.
- **La cuarta hipótesis** establecía que la combinación entre la tendencia relativa y el nivel relativo de la industria, son representaciones explícitas de la naturaleza

condicional de las razones reconocidas desde hace mucho tiempo por analistas de ratios.

Para el estudio se examinó 19 razones populares y cinco métodos de análisis predominantes. Las razones seleccionadas para este estudio se muestran en la Figura 5, las cuales han sido defendidas por teóricos, o han sido buenos predictores de fracaso empresarial en investigaciones experimentales anteriores.

**Figura 5**

*Razones seleccionadas para el modelo de Edmister.*

No Ratios	Componente del Ratio
1	Indicador de Liquidez
2	$\frac{\text{Capital de Trabajo}}{\text{Pasivo Total}}$
3	$\frac{\text{Inventarios}}{\text{Capital de Trabajo Neto}}$
4	$\frac{\text{Capital de Trabajo Neto}}{\text{Activo Total}}$
5	$\frac{\text{Activos Corrientes}}{\text{Pasivo Total}}$
6	$\frac{\text{Deuda Total}}{\text{Patrimonio Total}}$
7	$\frac{\text{Activos Fijos}}{\text{Patrimonio}}$
8	$\frac{\text{Flujo de Caja}}{\text{Pasivos Corrientes}}$

9	$\frac{\text{Pasivos Corrientes}}{\text{Patrimonio}}$
10	$\frac{\text{Patrimonio} + \text{Pasivos no Corrientes}}{\text{Activos Fijos}}$
11	$\frac{\text{Inventarios}}{\text{Ventas}}$
12	$\frac{\text{Activos Fijos}}{\text{Ventas}}$
13	$\frac{\text{Activo Total}}{\text{Ventas}}$
14	$\frac{\text{Capital de Trabajo Neto}}{\text{Ventas}}$
15	$\frac{\text{Patrimonio}}{\text{Ventas}}$
16	$\frac{\text{Utilidad antes de impuestos e intereses}}{\text{Ventas}}$
17	$\frac{\text{Utilidad antes de impuestos e intereses}}{\text{Activo Total}}$
18	$\frac{\text{Utilidad antes de impuestos e intereses}}{\text{Patrimonio}}$
19	$\frac{\text{Utilidad antes de impuestos e intereses}}{\text{Pasivo Total}}$

*Nota:* El Capital de Trabajo es igual al Activo Corriente – Pasivo Corriente, Capital de Trabajo Neto es igual al Activo Total – Pasivo Total, Flujo de Caja es igual a el Beneficio Neto + Amortizaciones, Adaptado de: “An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction”, R. Edmister, 1972. <https://doi.org/10.2307/2329929>.

Posteriormente, utilizó los ratios más representativos de estudios previos, como también consideró el valor del ratio en relación al porcentaje de estándares, convirtió los ratios a variables dicotómicas para compararlos con el valor individual del ratio Edmister, (1972).

**Figura 6***Resumen de Ratios usados por Edmister*

<b>Nombre del Ratio</b>
Indicador de Liquidez
Indicador de Circulante
$\frac{\text{Ratio de Inventarios}}{\text{Capital de Trabajo Neto}}$
$\frac{\text{Capital de Trabajo Neto}}{\text{Activo Corriente}}$
$\frac{\text{Activo Corriente}}{\text{Pasivo Corriente}}$
$\frac{\text{Flujo de Caja}}{\text{Pasivos Corrientes}}$
Indicador de Retorno de Capital o Intervalo sin crédito
$\frac{\text{Capital de Trabajo}}{\text{Pasivos Corrientes}}$
$\frac{\text{Activo Corriente} - \text{Existencias} - \text{Pasivo Corriente}}{\text{Gasto de Explotación} - \text{Amortización y Provisiones}}$

*Nota:* El ratio Intervalo sin crédito mide el volumen de ventas o facturación de la empresa y es un indicador del número de días que los Activos Defensivos (Disponible + Realizable) son capaces de hacer frente a las salidas de tesorería provocadas por los gastos de explotación sin ningún tipo de ayuda adicional. Adaptado de: “An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction”, R. Edmister, 1972. <https://doi.org/10.2307/2329929>.

Con respecto a los métodos seleccionados para el estudio de las variables independientes, detecto que, la capacidad predictiva de los ratios tenía que ver con la selección de los mismos, los métodos de análisis, así como también de las razones (Edmister, 1972, p. 11).

Planteó una técnica que se fundamentó en fraccionar la razón entre su respectivo promedio *bursátil de mercado*.

Según el análisis e interpretación de la información, indica que, los resultados más relevantes se lograron al elegir: “Un ratio para cada una de las diferentes características, con base en anteriores investigaciones había confirmado que un limitado número de ratios clasificados adecuadamente eran más útiles en comparación con el uso de un gran número de ratios” (Edmister, 1972, p. 11)

En la Figura 7 se muestra la función lineal discriminante con siete variables independientes.

**Figura 7**

*Función Lineal de Edmister*

$Z = 0.951 - 0.423X1 - 0.293X2 - 0.482X3 + 0.277X4 - 0.452X5 - 0.352X6 - 0.924X7$ <p style="text-align: center;"> <small>(-4.24**)    (-2.82**)    (-4.51**)    (2.61*)    (-2.60*)    (-1.68)    (-7.11*)</small> </p>	
<p><b>Z = Índice General</b></p> <p><math>X1 = \frac{\text{Flujo de Fondos Anuales}}{\text{Pasivo Corriente}}</math></p> <p><math>X2 = \frac{\text{Capital social}}{\text{Ventas}}</math></p> <p><math>X3 = \frac{\text{Capital de Trabajo Neto}}{\text{Ventas(promedio industrial)}}</math></p> <p><math>X4 = \frac{\text{Pasivo corriente}}{\text{Capital Social(Promedio Industrial)}}</math></p> <p><math>X5 = \frac{\text{Inventario}}{\text{Ventas(Promedio Industrial *)}}</math></p> <p><math>X6 = \frac{\text{Habilidad del ratio}}{\text{Ventas (Promedio del ratio **)}}</math></p> <p><math>X7 = \frac{\text{Habilidad del Ratio}}{\text{Promedio Industrial}}</math></p>	<p><b>* Variable de Interacción:</b> El ratio debe ser menor que el cuartil más bajo de la industria y exhibir un alza de tres años</p> <p><b>** Variable de Interacción:</b> El ratio debe ser menor que el cuartil más bajo de la industria y exhibir una baja de tres años</p> <p><b>Coefficiente de determinación</b> R<sup>2</sup> = 0,74</p> <p>F - ratio = 14,02</p> <p><b>Niveles de Significancia</b></p> <p>* = 0.05</p> <p>** = 0.01</p>

Adaptado de: “An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction”, R. Edmister, 1972. <https://doi.org/10.2307/2329929>.

Con respecto a la variable dependiente y a las variables independientes finales, el investigador las parametrizo de la siguiente manera:

- $Z = 1$  para un negocio exitoso (sin pérdidas) y  $Z = 0$  para un negocio fallido (sin pérdidas).
- $X1 = 1$  si el índice anual de flujo de fondos sobre pasivo corriente es menor que 0.05; de lo contrario,  $X1 = 0$ . Como podría esperarse, un flujo de fondos extremadamente bajo en relación con los compromisos a corto plazo es un predictor de fallas. La administración de pequeños negocios no ofrece ganancias después de impuestos y el flujo de fondos se define como la ganancia neta antes de impuestos más la depreciación.
- $X2 = 1$  si la relación capital sobre ventas es menor que 0.07; de lo contrario,  $X2 = 0$ . También es más probable que fracasen las empresas con una base de capital pequeña para sus ventas, independientemente de su clasificación industrial.
- $X3 = 1$  si el índice de capital de trabajo neto sobre ventas dividido por su respectivo índice de Asociados de Robert Morris - RMA es menor que -0.02; de lo contrario,  $X3 = 0$ . Una rotación de capital de trabajo relativamente alta (a RMA) presagia un fracaso.
- $X4 = 1$  si el pasivo corriente sobre patrimonio dividido por el respectivo índice para la administración de pequeños negocios promedia menos de 0.48; de lo contrario,  $X4 = 0$ . Confirmando creencias arraigadas, una relación deuda sobre capital baja en relación con la industria reduce la posibilidad de fracaso.

- $X5 = 1$  si la relación inventario sobre ventas muestra una tendencia alcista y aún es menor que 0.04; de lo contrario,  $X5 = 0$ . La inversa de esta relación es la rotación de inventario. Un prestatario con una rotación de inventario muy alta pero decreciente en relación con el promedio de la industria RMA es un fracaso potencial.  $X5 = 1$  si la tendencia del ratio rápido sobre RMA es a la baja y su nivel justo antes del préstamo 6 es inferior a 0,34; de lo contrario,
- $X6 = 0$ . Sólo dos empresas muestran esta característica de una relación rápida relativa de RMA baja y decreciente, pero ambas fracasaron.  $X7 = 1$  si el índice rápido del prestatario dividido por el índice rápido RMA muestra una tendencia alcista; de lo contrario,
- $X7 = 0$ . El signo negativo delante del coeficiente de esta variable indica que una relación rápida relativa RMA en aumento presagia un fracaso. Una posible explicación de este sorprendente resultado es que los acreedores se niegan a otorgar crédito o que el crédito no es necesario debido a las malas perspectivas comerciales.

Como consecuencia de la aplicación de las variables dependientes e independientes el balance que se obtuvo del Z-Scores se puede resumir de la siguiente de la siguiente manera: Z-Scores por debajo de 0.47 se obtuvo sólo para empresas en quiebra. Z-Scores sobre 0.53 fueron obtenidas sólo para empresas sanas. Se identifica la zona gris, con gran similitud a la zona de ignorancia catalogada por Altman la cual fue definida entre 0.47 y 0.53. La zona gris mostró un punto de corte de 0.52; reflejando esta la mejor precisión de clasificación global del 93%.

El análisis discriminante múltiple MDA permitió la selección adecuada para un conjunto óptimo de ratios y métodos. De igual manera permitió determinar las ponderaciones óptimas de cada uno de los coeficientes dentro de la función lineal.

Los resultados indicaron que la efectividad de los ratios individuales fue efectiva para pronosticar quiebras. Para Edmister, (1972) el MDA permitió la selección de la mitad de un medio para un conjunto óptimo de ratios y métodos, como también facilitó la asignación de las ponderaciones de cada uno de los factores dentro de la función lineal.

Con el paso a paso utilizado en el MDA concluye que los métodos de análisis fueron útiles para: observar la tendencia del comportamiento de las empresas en un periodo de tres años, tendencia de los ratios y de sus resultados más recientes, capacidad para calcular en promedio el comportamiento de tres años antes de la falla, posibilidad de fraccionar los ratios entre sus correspondientes promedios bursátil de mercado los valores de los ratios entre los valores medios del sector industrial

La función discriminante mostró una alta capacidad de clasificación, alcanzando un porcentaje del 93% de precisión. Un Z-Score menor a 0.520 indica predicción de quiebra; por el contrario, Z-Score mayor de 0.520 indican predicción de no quiebra. Edmister, (1972)

En la Figura 8, se muestra algunos indicadores que fueron relevantes en la investigación.

**Figura 8***Valores de los indicadores Relevantes - Edmister*

Puntuaciones para Z	Exactitud Predictiva	
	Quiebra	No Quiebra
Arriba de 0.469	80%	100%
de 0.470 a 0.519	25%	95%
de 0.520 a 0.529	90%	95%
0.530 en adelante	100%	86%

Adaptado de: "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction", R. Edmister, 1972. <https://doi.org/10.2307/2329929>.

Edmister concluyó basándose en los descubrimientos de Altman que un grupo pequeño de ratios tienen mejor precisión que cualquier ratio único y que estandarizando los ratios, la estandarización de los ratios y la conversión de variables continuas a suman significancia al modelo.

### ***Modelo Deakin (1972, 1976)***

En su investigación titulada *A discriminant analysis of predictors of business failure*, Deakin (1972) comparó el modelo de Beaver de 1966 con el de Altman de 1968 concluyendo que, pese a que el modelo univariable de Beaver obtuvo mejores resultados predictivos, el modelo multivariable de Altman tenía más atracción intuitiva.

**Tabla 1**

*Tasas de Error de los modelos de predicción de insolvencia de Altman y Beaver 1968 Modelo de Deakin*

<b>Años antes del Fracaso</b>	<b>Beaver - Cash Flow / Total Debt</b>	<b>Altman - Función Discriminante</b>
1	13	5
2	21	28
3	23	52
4	24	71
5	22	64

Adaptada de: "A Discriminant Analysis of Predictors of Business Failure", por: E. Deakin, 1972. Journal of Accounting Research, Vol. (10), p.1. <https://doi.org/10.2307/2490225>

El propósito de Deakin fue proponer un modelo alternativo al de Beaver y Altman para predecir el fracaso y mejorar los errores de clasificación tipo I y tipo II un año antes del fracaso - quiebra, para tal fin combinó ambas investigaciones. Inicialmente replicó el análisis de clasificación dicotómica de Beaver teniendo en cuenta la muestra de 79 empresas fallidas emparejadas con las 79 empresas No Fallidas entre los años de 1954 a 1964 y 14 de los ratios financieros usados en su estudio, luego introdujo una segunda muestra de 32 empresas Fallidas emparejadas con 32 No Fallidas entre los años 1964 a 1970 también teniendo en cuenta las 14 razones utilizadas previamente por Beaver en su estudio resumidas la Figura 9.

**Figura 9***Razones Financieras usadas en el Modelo de Deakin 1972*

<b>Grupo de Activos no Líquidos</b>	<b>Grupo de Activos Líquidos entre Pasivos Corriente</b>
$\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$	$\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$
$\frac{\text{Utilidad Neta}}{\text{Activos Totales}}$	$\frac{\text{Activo Disponible}}{\text{Pasivos Corrientes}}$
$\frac{\text{Pasivo Total}}{\text{Activos Totales}}$	$\frac{\text{Efectivo}}{\text{Pasivos Corrientes}}$
<b>Grupo de Activos Líquidos entre Activos Totales</b>	<b>Grupo del Retorno de los Activos Líquidos</b>
$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$	$\frac{\text{Activos Corrientes}}{\text{Ventas}}$
$\frac{\text{Activo Disponible}}{\text{Activos Totales}}$	$\frac{\text{Activo Disponible}}{\text{Ventas}}$
$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$	$\frac{\text{Capital de Trabajo Neto}}{\text{Ventas}}$
$\frac{\text{Efectivo}}{\text{Activos Totales}}$	$\frac{\text{Efectivo}}{\text{Ventas}}$

Adaptada de: "A Discriminant Analysis of Predictors of Business Failure", por: E. Deakin, 1972. Journal of Accounting Research, Vol. (10), p.1. <https://doi.org/10.2307/2490225>.

Luego de que Deakin confrontó sus resultados con los del estudio de Beaver analizó el coeficiente de correlación de Spearman ( $r_8$ ) el cual puede usarse para indicar el orden del poder predictivo de las proporciones en los dos estudios, realizó una comparación de las medias de 13 elementos de los estados financieros que se usaron en el cálculo de los ratios financieros con la esperanza de entender por qué este coeficiente de correlación es menor el tercer año. Posteriormente utilizó el análisis discriminante de Altman para buscar la

combinación lineal de las 14 razones financieras usados por Beaver que mejor discriminen entre los grupos que se están clasificando.

Los resultados luego del análisis indican que los errores de clasificación errónea promediaron 3% para el primer año, 4,5% el segundo año y 4,5% el tercer año, en comparación con los resultados de clasificación que se muestran en la Tabla 2, el modelo de Deakin arrojó resultados mejores que el modelo de Beaver y Altman un solo año, sin embargo, las tasas de error aumentan notablemente después del tercer año.

La razón aparente que observó Deakin es que las empresas fallidas tendieron a expandirse rápidamente después del tercer año antes del fracaso, al revisar la estructura de capital parece que la expansión fue financiada por un aumento de la deuda y acciones preferentes en lugar de acciones comunes o ganancias retenidas, los fondos recaudados se invirtieron en plantas y equipos en lugar de en activos líquidos en consecuencia estas empresas no pudieron generar las ventas y el ingreso neto para respaldar su deuda, si bien estas tasas de error son probablemente demasiado altas para la toma de decisiones, aún son más bajas que las tasas del modelo de Altman y Beaver. Finalmente, Deakin realizó una prueba de pertenencia a nueva muestra de las funciones derivadas previamente en los grupos de empresas fallidas y no fallidas que sabía que no habían fallado en los próximos cinco años o que habían fallado entre uno y cinco años en el futuro resumidos en la Figura 10.

**Figura 10**

*Tasas de error usando funciones discriminantes en datos mixtos Modelo de Deakin.*

Año anterior a la falla de la que se derivó la función	Porcentaje mal clasificado		Error Total
	Fallida	No Fallida	
1	14	13	13
2	10	9	10
3	19	17	18

Adaptada de: "A Discriminant Analysis of Predictors of Business Failure", por E. Deakin, 1972. Journal of Accounting Research, Vol. (10), p.1. <https://doi.org/10.2307/2490225>.

Deakin concluyó que la aplicación de técnicas estadísticas como el análisis discriminante, se puede utilizar para predecir el fracaso empresarial a partir de datos contables de hasta tres años de anticipación con una precisión bastante alta, no obstante, estas probabilidades de pertenencia a un grupo solo deben ser usadas como evidencia adicional referente al fracaso más no como una prueba concluyente del mismo.

En sus posteriores investigaciones su propósito fue definir mejor a la variable dependiente con el fin de mejorar los resultados de su modelo, tomando 11 de los 14 ratios utilizados por Beaver, los ratios faltantes no se consideraron por la falta de información, como recomendación indica que los Ratios de contabilidad financiera podrían distribuirse normalmente dentro de un sector específico de las industrias, en la Figura 11 se ilustran las 11 razones usadas.

**Figura 11**

*Razones Financieras usadas en el Modelo de Deakin 1976*

<b>Grupo de Activos no Líquidos</b>	<b>Grupo de Activos Líquidos entre Pasivos Corriente</b>
$\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$	$\frac{\text{Activos Corrientes}}{\text{Pasivos corrientes}}$
$\frac{\text{Utilidad Neta}}{\text{Activos Totales}}$	$\frac{\text{Activo Disponible}}{\text{Pasivos Corriente}}$
$\frac{\text{Pasivo Total}}{\text{activos Totales}}$	<b>Grupo del Retorno de los Activos Líquidos</b>
<b>Grupo de Activos Líquidos entre Activos Totales</b>	$\frac{\text{Activos Corrientes}}{\text{Ventas}}$
$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$	$\frac{\text{Activo Disponible}}{\text{Ventas}}$
$\frac{\text{Activo Disponible}}{\text{activos Totales}}$	$\frac{\text{Capital de Trabajo Neto}}{\text{Ventas}}$
$\frac{\text{Efectivo}}{\text{Activos Totales}}$	

Adaptada de: “A Discriminant Analysis of Predictors of Business Failure”, por E. Deakin, 1972. Journal of Accounting Research, Vol. (10), p.1. <https://doi.org/10.2307/2490225>.

### ***Modelo Blum (1978)***

Con el objetivo de evitar los monopolios, Blum desarrollo en 1969 un modelo en su tesis doctoral llamado *The falling Company Doctrine*, siendo una herramienta que permitiera a las empresas no concluir fusiones ilegales, al identificar empresas con gran probabilidad de fracaso empresarial que se quieran unir con otras que desconocen dicha condición, con la ayuda la firma de abogados Gorden, Feinblatt y Asociados de Baltimore Blum profundizó más en el tema y desarrolló dos artículos.

Introdujo posteriormente resultados de un análisis discriminante en su primer modelo de 1974 dando lugar a la creación de un modelo llamado *The Failing Company Model*, en donde explica que este modelo sirve como una defensa antimonopolio que mejora la fluidez

del capital de mercado, siendo un factor para minimizar las pérdidas de inversión, estimular la competencia y ayudar a las nuevas empresas a entrar al mercado, tomando como referencia en el modelo dicotómico de predicción de insolvencia de Beaver, en su segundo modelo incluyo un análisis de ratios o razones de probabilidad. Según Blum “el empleo de un análisis discriminante es un método que se direcciona a obtener una clasificación de individuos de diversas clases de forma a priori y pertenecientes a un grupo de variables cuantitativas y cualitativas”, (como se cita en Walteros, 2018, p. 77).

Blum definió al fracaso como el momento en que la organización empresarial inicia un procedimiento con el motivo de declarar la quiebra o cuando se da un acuerdo explícito con los acreedores, los cuales permiten una reducción sobre sus pasivos, considerando concluir que las empresas fracasadas eran aquellas que no podían pagar sus deudas es decir que no podían cumplir las obligaciones financieras con sus acreedores en el corto plazo.

En su modelo utilizó inicialmente 12 razones que posteriormente redujo a solo 10 razones en donde incluía variables para medir el cambio de dichas razones en el tiempo para esto se valió de información financiera como el balance, los estados de resultados y precios de mercado, en los resultados de las pruebas de validación de la muestra pudo obtener una precisión de hasta el 94% para la predicción de fallas comerciales.

Figura 12

Razones usadas en el Modelo de la Empresa Fallida Blum

I. Liquidez	A. Liquidez a Corto Plazo	Flujo:	1. $\frac{(\text{Efectivo} + \text{Cuentas por Cobrar} + \text{Inversiones} + \frac{\text{Ventas anuales}}{12})}{\text{Costo de Mercancía Vendida} - \text{Depreciación} + \text{Gastos Administrativos} + \text{Intereses}}$
		Posición:	2. $\frac{\text{Activo Corriente} - \text{Inventarios} - \text{Pasivos Corrientes}}{\text{Inventarios}}$
	B. Liquidez a Largo Plazo	Flujo:	3. $\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$
		Posición:	4. $\frac{\text{Patrimonio Neto A valor Razonable de Mercado}}{\text{Pasivo Total}}$
			5. $\frac{\text{Patrimonio Neto a Valor en Libros}}{\text{Pasivo Total}}$
II. Probabilidad		6. Ratio de Rentabilidad: (el rendimiento se basa en la ganancia o pérdida del accionista y los dividendos en efectivo recibidos)	
III. Variabilidad		7. Desviación estándar de los Ingresos Netos de un Período	
		8. Tendencia de los Ingresos Netos	
		9. Pendiente de los Ingresos Netos	
		10, 11, 12. Desviación Estándar Ratios, Tendencia Ratios, Pendiente de los ratios; Las variables 10, 11 y 12 solo fueron usadas el primer y segundo año antes del fracaso de las empresas	

Adaptado de "The Failing Company Model", por M. Blum, 1974.

<http://lawdigitalcommons.bc.edu/bclr/vol16/iss1/2>

Su muestra se compuso de 230 empresas industriales de las cuales 115 empresas en quiebra durante el periodo de 1954 a 1968 y cuyos pasivos superan 1 millón de dólares y otras 115 empresas sanas es decir que no presentaban problemas financieros como insolvencia para su emparejamiento homólogo al trabajo de Beaver, tuvo en cuenta cuatro criterios el primero fue la clasificación industrial de la empresa, el segundo información financiera de las ventas

de por lo menos cuatro años antes de la falla - quiebra, el tercero fue el número de empleados y el cuarto el año fiscal del cierre contable.

Al continuar con sus investigaciones Blum analizó la capacidad de predicción de datos en bruto y/o absolutos en donde observó que obtuvo una mayor precisión en la predicción para los primero años antes de la quiebra, investigando los efectos al considerar periodos o rangos de tiempo, también desarrolló un modelo en donde no tuvo en cuenta ratios financieros sin embargo los resultados del primer año previo a la quiebra eran inferiores a los del modelo anterior determinando así que no era un modelo muy eficiente con respecto a anteriores modelos.

### **3.2.3. Análisis Factorial**

El análisis factorial es una técnica estadística para la reducción de datos, el análisis factorial es usado para disminuir un conjunto de variables observadas por medio de combinaciones lineales y expresarlas en términos de un menor número de variables no observadas llamadas factores. En los modelos de predicción de insolvencia financiera y/o quiebra, el análisis factorial se utiliza para agrupar los ratios financieros que más se correlacionan en categorías relativamente no correlacionadas, el objetivo es reducir un la cantidad o número de ratios intercorrelacionados a una cantidad inferior de factores no correlacionados, que permiten explicar la mayor parte de variabilidad de cada uno de los ratios (Montoya, 2007).

*Modelo de Robert Libby (1975)*

Tomando como ejemplo el estudio de Beaver (1968) y Deakin (1972) acerca de la capacidad predictiva y el comportamiento de la información contable, éste estudio busca evaluar la fuerza y utilidad de los ratios contables en la predicción del fracaso empresarial así como la capacidad de los oficiales de préstamos o banqueros de evaluar dicho comportamiento e información; así pues, está elaborado teniendo en cuenta que ciertos ratios están estrechamente relacionados con el fracaso empresarial y que los coeficientes contables son propicios para realizar la predicción del fracaso.

Para el desarrollo de este estudio se tuvieron en cuenta tres preguntas y/o argumentos que permitirían la correcta evaluación del análisis en la interpretación de ratios, éstos fueron:

- La precisión de la predicción no se desvía de las predicciones aleatorias
- La pertenencia a un subgrupo no afecta en la predicción del logro
- La fiabilidad del test-retest es igual medida en intervalos inmediatos y de una semana

Estos argumentos fueron estudiados uno a uno a través de dos subgrupos, en donde para la elección de los mismos se tuvo en cuenta las muestras de derivación de Deakin y los 14 ratios estudiados por Beaver y Deakin, los cuales como resultado de un proceso de depuración se redujeron a 5, cada uno con su factor rotatorio como se muestra a continuación:

**Figura 13***Razones Utilizadas en el modelo de Libby*

No. Razón	Componentes de la Razón
1	$\frac{\text{Ingresos Netos}}{\text{Activos Totales}}$ (Rentabilidad)
2	$\frac{\text{Activos Corrientes}}{\text{Ventas}}$ (Actividad)
3	$\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$ (Liquidez)
4	$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$ (Saldo de Activos)
5	$\frac{\text{Efectivo (Cash Flow)}}{\text{Pasivos Totales}}$ (Situación de Caja)

Adaptado de: "Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence". R. Libby, 1975, Journal of Accounting Research. Vol. (13), p.1. <https://doi.org/10.2307/2490653>.

Luego entonces, de la muestra de derivación de Deakin se tomaron los 60 casos junto con 10 repetidos para un total de 70 casos, los cuales fueron ordenados en dos cuestionarios diferentes, uno que contaba con los 70 casos completos y sería entregado a aquellos banqueros que lo realizarían en una sola sesión y otro que se dividía en dos, 30 y 40 casos respectivamente y que serían entregados a aquellos banqueros que lo realizarían en dos sesiones.

Una vez que los cuestionarios fueron realizados, se obtuvo que para el primer argumento los resultados dependieron de la correcta utilización de la información suministrada y se concluyó que el logro de la predicción es superior al rendimiento aleatorio;

para el segundo argumento se analizaron las diferencias en la exactitud media, sin embargo estas no fueron significativas y se concluyó que las diferencias individuales son relacionadas con la exactitud de la predicción y finalmente, el tercer argumento se midió a través de la concordancia de cada sujeto entre sus predicciones anteriores y posteriores, concluyendo que las interpretaciones de los datos contables no varían a lo largo del tiempo. “El estudio afirmaba que los ejecutivos de préstamo de las entidades financieras a través del uso de razones contables eran capaces de predecir la quiebra empresarial con un 74% de precisión” (Libby, 1975, p. 8).

Las limitaciones halladas en el estudio realizado tienen que ver con los resultados del mismo, pues el aspecto general de éstos difieren en base a factores como los actores, otras situaciones experimentales, situaciones del mundo real y otras definiciones de fracaso; sin embargo, pese a estas limitaciones fue posible concluir que toda aquella información contable es útil si permite hacer predicciones correctas, teniendo en cuenta además la gran importancia de la capacidad de los usuarios de analizar e interpretar los datos obtenidos; finalmente, este estudio fue propicio para medir el impacto de los ratios contables junto con el análisis de los oficiales de préstamo en tanto a la predicción del fracaso.

Si bien el logro de predicción de los tomadores de decisiones individuales promedió en un 74.4%, el grupo compuesto tuvo un desempeño del 81.7%, solo una predicción menos precisa que el mejor analista individual. En cuanto a las limitaciones, las instrucciones específicas, las presentaciones de datos y las medidas dependientes utilizadas en la situación experimental pueden afectar los resultados del estudio. En la decisión de préstamo real, los

analistas de crédito tienen a su disposición abundante información cuantitativa y cualitativa de varios períodos (Libby, 1975).

***Modelo Pinches, Eubank, Mingo y Caruthers (1975)***

En un primer estudio Pinches, Mingo y Caruthers emplearon el factor de análisis con la finalidad de examinar las interrelaciones entre las razones financieras y su estabilidad en el largo plazo sus resultados indicaron que se pueden determinar agrupaciones significativas y estables en el tiempo incluso cuando las magnitudes financieras estén cambiando. Para su investigación *The hierarchical classification of financial Ratios* del año 1975, el objetivo principal de Pinches, Eubank, Mingo y Caruthers fue examinar la estabilidad en el corto plazo de razones financieras empíricas agrupadas para determinar relaciones jerárquicas entre los grupos e integrar la importancia predictiva individual de las razones financieras identificadas en el estudio.

Para esto tomaron una muestra de 221 empresas industriales obtenida de la base de datos financieros COMPUSTAT con observaciones en los 1966, 1967, 1968, 1969 las cuales tenían la información completa requeridas para el cálculo de 48 razones financieras las cuales se seleccionaron después de una revisión bibliográfica de estudios que involucraran razones financieras.

**Figura 14**

Razones Financieras para Empresas Industriales Modelo de Pinches, Eubank, Mingo y Caruthers

$\frac{\text{Flujo de Efectivo}}{\text{Ventas}}$	$\frac{\text{Activos Rápidos}}{\text{Gastos del Fondo}}$
$\frac{\text{Ingreso Total}}{\text{Ventas}}$	$\frac{\text{Efectivo}}{\text{Gastos del Fondo}}$
$\frac{\text{Ingreso Neto}}{\text{Ventas}}$	$\frac{\text{Flujo de Efectivo}}{\text{Activo Total}}$
$\frac{\text{Pasivo Corriente}}{\text{Planta Neta}}$	$\frac{\text{Flujo de Efectivo}}{\text{Patrimonio}}$
$\frac{\text{Pasivo Corriente}}{\text{Patrimonio}}$	$\frac{\text{Flujo de Efectivo}}{\text{Pasivo Total}}$
$\frac{\text{Deuda}}{\text{Planta Neta}}$	$\frac{\text{Ingreso Total}}{\text{Activo Total}}$
$\frac{\text{Deuda}}{\text{Capital}}$	$\frac{\text{Ingreso Neto}}{\text{Activo Total}}$
$\frac{\text{Pasivo Total}}{\text{Patrimonio}}$	$\frac{\text{Ingreso Neto}}{\text{Patrimonio}}$
$\frac{\text{Capital de Trabajo}}{\text{Activo Total}}$	$\frac{\text{Ingreso Neto}}{\text{Pasivo Total}}$
$\frac{\text{Activo Total}}{\text{Patrimonio}}$	$\frac{\text{Patrimonio}}{\text{Ventas}}$
$\frac{\text{Activo}}{\text{Pasivo Total}}$	$\frac{\text{Ventas}}{\text{Capital de Trabajo}}$
$\frac{\text{Efectivo}}{\text{Activo Total}}$	$\frac{\text{Ventas}}{\text{Activo Total}}$
$\frac{\text{Efectivo}}{\text{Pasivo Corriente}}$	$\frac{\text{Costo de Bienes Vendidos}}{\text{Inventario}}$
$\frac{\text{Activo Corriente}}{\text{Activo Total}}$	$\frac{\text{EBIT}}{\text{Activo Total}}$
$\frac{\text{Activo Corriente}}{\text{Pasivo Corriente}}$	$\frac{\text{EBIT}}{\text{Ventas}}$

$\frac{\text{Inventario}}{\text{Activo Corriente}}$	$\frac{\text{Cargos Fijos}}{\text{EBIT}}$
$\frac{\text{Inventario}}{\text{Capital de Trabajo}}$	$\frac{\text{Ventas} - \text{Costo de Bienes Vendidos}}{\text{Ventas}}$
$\frac{\text{Activos Rápidos}}{\text{Activo Total}}$	$\frac{\text{Ventas}}{\text{Planta Neta}}$
$\frac{\text{Activos Rápidos}}{\text{Pasivos Corrientes}}$	$\frac{\text{Flujo de Efectivo}}{\text{Capital Total}}$
$\frac{\text{Reservas}}{\text{Ventas}}$	$\frac{\text{Ingreso Total}}{\text{Capital Total}}$
$\frac{\text{Efectivo}}{\text{Ventas}}$	$\frac{\text{Ventas}}{\text{Capital Total}}$
$\frac{\text{Activos Corrientes}}{\text{Ventas}}$	$\frac{\text{Pasivo Corriente}}{\text{Activo Total}}$
$\frac{\text{Inventario}}{\text{Ventas}}$	$\frac{\text{Deuda}}{\text{Activo Total}}$
$\frac{\text{Activos Rápidos}}{\text{Ventas}}$	$\frac{\text{Pasivo Total}}{\text{Activo Total}}$

Adaptada de "The Hierarchical Classification of Financial Ratios", por G. Pnches et al., 1975.

<https://www.sciencedirect.com/science/article/abs/pii/0148296375900119>

Luego de calculadas las razones financieras se aplicó una transformación logarítmica para mejorar la normalidad, reducir los valores atípicos y mejorar la homocedasticidad de las distribuciones, por medio de procedimientos numéricos se determinaron 7 agrupaciones de razones basadas en similitudes empíricas entre las razones de interés. Cada agrupación se etiquetó representando el impulso predominante de las razones financieras dentro del grupo y las diferentes operaciones de una empresa: retorno en inversión, número de veces que rota el capital, número de veces que rotan los inventarios, apalancamiento financiero, rotación de cuentas por cobrar, liquidez a corto plazo y posición de caja.

Se emplearon técnicas de análisis factorial de primer orden para clasificar las 48 razones - factores y especificar las similitudes empíricas entre los grupos ayudando a determinar las interrelaciones de los ratios financieros pertenecientes a diferentes grupos debido a que las clasificaciones determinadas empíricamente deben tener un alto grado homogeneidad interna (dentro del grupo) y alta heterogeneidad externa (entre grupos).

No se pudo agrupar 8 de las 48 razones en ninguna de las 7 clasificaciones determinadas empíricamente por lo que se dejaron por fuera de los grupos, posteriormente mediante el análisis de factores de orden superior se agruparon los 7 grupos en 3 nuevas agrupaciones de orden superior: rendimiento de capital invertido, liquidez general y rotación de capital a corto plazo. Al seleccionar una razón de cada grupo de clasificación se puede identificar un conjunto de ratios financieros independientes entre sí, es decir baja interrelación a continuación en la Figura 15, se presenta los resultados de la clasificación jerárquica de los ratios.

Figura 15

*Clasificación Jerárquica de las Razones Financieras para Empresas Industriales Modelo de Pinches, Eubank, Mingo y Caruthers*

<b>Rendimiento de Capital Invertido</b>	<b>Retorno en Inversión</b>	$\frac{\text{Ingreso Total}}{\text{Ventas}}$
		$\frac{\text{Flujo de Efectivo}}{\text{Activo Total}}$
		$\frac{\text{Efectivo}}{\text{Activo Total}}$
		$\frac{\text{Ingreso Total}}{\text{Activo Total}}$
		$\frac{\text{Ingreso Neto}}{\text{Activo Total}}$
		$\frac{\text{Ingreso Neto}}{\text{Patrimonio}}$
		$\frac{\text{EBIT}}{\text{Activo Total}}$
		$\frac{\text{EBIT}}{\text{Ventas}}$
		$\frac{\text{Flujo de Efectivo}}{\text{Capital Total}}$
		$\frac{\text{Ingreso Total}}{\text{Capital Total}}$
<b>Rotación de Capital</b>	$\frac{\text{Deuda}}{\text{Planta Neta}}$	
	$\frac{\text{Deuda}}{\text{Capital Total}}$	
	$\frac{\text{Pasivo Total}}{\text{Patrimonio}}$	
	$\frac{\text{Activo Total}}{\text{Patrimonio}}$	
	$\frac{\text{Activo Total}}{\text{Patrimonio}}$	

		$\frac{\text{Deuda}}{\text{Activo Total}}$ $\frac{\text{Pasivo Total}}{\text{Patrimonio}}$
<b>Liquidez General</b>	<b>Rotación de Inventarios</b>	$\frac{\text{Flujo de Efectivo}}{\text{Ventas}}$ $\frac{\text{Activo Corriente}}{\text{Activo Total}}$ $\frac{\text{Activos Rápidos}}{\text{Activo Total}}$ $\frac{\text{Patrimonio}}{\text{Ventas}}$ $\frac{\text{Ventas}}{\text{Activo Total}}$ $\frac{\text{Ventas}}{\text{Planta Neta}}$
	<b>Apalancamiento Financiero</b>	$\frac{\text{Pasivo Corriente}}{\text{Patrimonio}}$ $\frac{\text{Activo Corriente}}{\text{Pasivo Corriente}}$ $\frac{\text{Inventario}}{\text{Capital de Trabajo}}$
		$\frac{\text{Activos Rápidos}}{\text{Pasivo Corriente}}$ $\frac{\text{Pasivo Corriente}}{\text{Activo Total}}$
	<b>Rotación de Cuentas por Cobrar</b>	$\frac{\text{Efectivo}}{\text{Activo Total}}$ $\frac{\text{Efectivo}}{\text{Pasivo Corriente}}$

<b>Rotación de Capital a Corto Plazo</b>		$\frac{\text{Efectivo}}{\text{Ventas}}$
		$\frac{\text{Efectivo}}{\text{Gastos del fondo}}$
	<b>Liquidez a Corto Plazo</b>	$\frac{\text{Activos Corrientes}}{\text{Ventas}}$ $\frac{\text{Inventario}}{\text{Ventas}}$ $\frac{\text{Ventas}}{\text{Capital de Trabajo}}$ $\frac{\text{Costo de Bienes Vendidos}}{\text{Inventario}}$
	<b>Posición de Caja</b>	$\frac{\text{Reserva}}{\text{Inventario}}$ $\frac{\text{Inventario}}{\text{Activo Corriente}}$ $\frac{\text{Reserva}}{\text{Ventas}}$ $\frac{\text{Activos Rápidos}}{\text{Ventas}}$

Adaptada de “The Hierarchical Classification of Financial Ratios”, por G. Pinches et al., 1975.

<https://www.sciencedirect.com/science/article/abs/pii/0148296375900119>

Adicionalmente Pinches, Eubank, Mingo y Caruthers también realizaron una revisión bibliográfica con la finalidad de encontrar las Razones Financieras Identificadas como predictores significativos y compararlas con la clasificación jerárquica de su estudio, a continuación, en la Figura 16 se presenta los ratios identificados en la revisión bibliográfica:

**Figura 16**

*Razones Financieras Identificadas como predictores significativos Modelo de Pinches, Eubank, Mingo y Caruthers*

<b>Clasificación Financiera de la Razones</b>	<b>Razones</b>
	$\frac{\text{EBIT}}{\text{Activo Total}}$
Retorno de la inversión	$\frac{\text{Utilida Neta}}{\text{Activo Total}}$
	$\frac{\text{EBIT}}{\text{Ventas}}$
Rotación de capital	$\frac{\text{Ventas}}{\text{Activo Total}}$
	$\frac{\text{Patrimonio}}{\text{Ventas}}$
Rotación de inventario	$\frac{\text{Ventas}}{\text{Capital de Trabajo}}$
	$\frac{\text{Pasivo Total}}{\text{Patrimonio}}$
Apalancamiento Financiero	$\frac{\text{Pasivo Total}}{\text{Activo Total}}$
	$\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$
Razones sin Clasificación	$\frac{\text{Capital de Trabajo}}{\text{Activo Total}}$
	$\frac{\text{Utilidad Neta} + \text{intereses}}{\text{Intereses}}$
	$\frac{\text{Utilidad retenida}}{\text{Activos Totales}}$
Ratios no incluidos en el Estudio	$\frac{\text{Valor de Mercado del Patrimonio}}{\text{Pasivo Total}}$

Adaptada de "The Hierarchical Classification of Financial Ratios", por G. Pnches et al., 1975.

<https://www.sciencedirect.com/science/article/abs/pii/0148296375900119>

El análisis de sus resultados indica que: 3 de las 8 razones financieras que no fueron agrupadas en ninguna categoría flujo de caja entre deuda total, capital de trabajo entre activos totales e Ingresos netos e intereses entre intereses y 3 de las 40 razones financieras si agrupadas en alguna de las 7 categorías ventas entre activos totales, deuda entre capital total, y deuda entre activos totales, además de ser las razones financieras más descriptivas en sus respectivas categorías coinciden con las razones predictivas más significativas de la anterior Figura.

Pinches, Eubank, Mingo y Caruthers concluyen por un lado que si la finalidad es describir o representar un sector industrial es conveniente para los investigadores y analistas usar la clasificación jerárquica de las razones producto de esta investigación ya que en las 7 categorías se representan las diferentes operaciones de un negocio por otro lado, si la finalidad es la predicción es conveniente seleccionar una revisión bibliográfica de las razones financieras más significativas a la fecha.

### ***Modelo de Ibarra (2001)***

El objetivo principal de Ibarra, 2001 en su trabajo de investigación titulado *Análisis de las dificultades financieras de las empresas en una Economía emergente: las bases de datos y las variables Independientes en el sector hotelero de la bolsa Mexicana de Valores* es estudiar en una economía emergente dentro de un sector muy específico proponiendo un modelo predictivo de quiebra y/o insolvencia financiera para el contexto mexicano basado en análisis factorial. Ibarra estudio el paradigma de la utilidad de la información contable, observando a las bases de datos de la información financiera como un problema central junto

con el efecto window dressing y su repercusión negativa en la utilidad del sistema contable preocupándose por:

- Estructurar una base de datos confiable.
- Definir mejor la variable independiente referente a las razones financieras generadas a partir la información financiera.
- Establecer la variable dependiente desde el enfoque del éxito o fracaso empresarial.

Una vez analizados los anteriores 3 elementos estructurales de los modelos predictivos, se revisaron otras variables como el tamaño de la empresa, la cantidad de activos, el volumen de ventas, el importe de los beneficios, los pasivos totales, el capital contable, la rotación de los activos y el tamaño de los activos fijos además de otros factores influyentes como el sector y la ubicación geográfica.

Para la muestra se seleccionaron 5 empresas - hoteles que pertenecen al sector servicios y forman parte de un total de 240 empresas que cotizan en la bolsa mexicana de valores, de las 5 empresas 1 esta categorizada como no sana con problemas de insolvencia o quiebra financiera y 4 sanas sin problemas de insolvencia o quiebra financiera.

Para compensar la limitada muestra consideraron la información financiera de 8 años, en su investigación utilizaron el modelo estadístico multivariable factorial a través de los componentes principales, la investigación toma como referencia el trabajo Libby, 1975 al ser el precursor o una de las primeras personas en investigar la aplicación del Análisis Factorial mediante el método de componentes principales antes de aplicar el análisis factorial común.

El modelo incluye 71 razones financieras recolectadas a partir de una revisión bibliográfica, las cuales se dividieron en 7 categorías las cuales son: razones de rentabilidad, razones de productividad, razones de eficiencia, razones de liquides, razones de flujo de efectivo, razones de solvencia y razones de endeudamiento, posteriormente la cantidad de razones se reduce de 71 a 21 debido a las limitaciones de la muestra, a continuación se presenta la Figura 17 con los ratios aplicados en el estudio.

**Figura 17**

*Selección de Razones Financieras Aplicadas al Modelo de Ibarra*

<b>1. RATIOS DE RENTABILIDAD</b>	
	<b>Rentabilidad de la Explotación o Económica. 1er criterio:</b>
R1	$\frac{\text{Beneficio neto antes de interes e impuestos}}{\text{Media del Activo Total (Recursos propios de explotacion y financieros)}}$
	<b>Rentabilidad de la Empresa 1er Criterio:</b>
R4	$\frac{\text{Beneficios Netos despues de intereses antes de impuestos}}{\text{Activo Total}}$
	<b>Rentabilidad Financiera con Base en el Capital Propio:</b>
R9	$\frac{\text{Beneficio neto despues de impuestos}}{\text{Capital contable}}$
	<b>Rentabilidad Financiera con Base en el Activo Bruto:</b>
R11	$\frac{\text{Beneficios netos antes de intereses e impuestos}}{\text{Activo Total Bruto}}$
	<b>Rentabilidad Acumulada:</b>
R12	$\frac{\text{Beneficios Netos no distribuidos}}{\text{Activo Total}}$
<b>2. RATIOS DE PRODUCTIVIDAD Y 3. EFICIENCIA</b>	
	<b>Ratio de Margen Neto de Beneficios o Tasa de Margen Neto:</b>
R13	$\frac{\text{Beneficio Neto despues de interes e impuestos}}{\text{Ventas Netas}}$
R14	<b>Ratio de Rotación de Activos:</b>

		$\frac{\text{Ventas Netas}}{\text{Activo Total}(\text{capital total invertido})}$
R15	<b>Tasa de Margen Bruto:</b>	$\frac{\text{Beneficio neto antes de interes e impuestos}}{\text{Ventas Netas}}$
<b>4. RATIOS DE LIQUIDEZ</b>		
R28	<b>Ratio o Coeficiente de Capital Circulante:</b>	$\frac{\text{Activo Circulante}}{\text{Pasivo Circulante}}$
R30	<b>Ratio o Coeficiente de Tesorería:</b>	$\frac{(\text{Tesorería} + \text{Equivalentes})}{\text{Pasivo Circulante}}$
R35		$\frac{\text{Activo Circulante}}{\text{Activo Total}}$
R36		$\frac{\text{Capital de Trabajo}}{\text{Activo Total}}$
R38		$\frac{\text{Activo Circulante}}{\text{Pasivo Circulante}}$
<b>5. RATIOS DE CASH FLOW</b>		
R50		$\frac{\text{Cash Flow}}{\text{Pasivo Total}}$
<b>6. RATIOS DE SOLVENCIA</b>		
R51	<b>Ratio de Solvencia Global:</b>	$\frac{\text{Pasivo Total}}{\text{Activo Total}}$
R52	<b>Ratio de Estructura Financiera:</b>	$\frac{\text{Valor de mercado de los fondos propios}}{\text{Valor Contable de los pasivos}}$
R56	<b>Ratio de Protección al Pasivo (2):</b>	$\frac{\text{Pasivo Fijo}}{\text{Capital Contable}}$
R57	<b>Ratio de Protección al Pasivo (3):</b>	

		$\frac{\text{Pasivo Total}}{\text{Capital contable}}$
	<b>Ratio de Protección al Activo (1):</b>	
R58		$\frac{\text{Activo Circulante}}{\text{Capital contable}}$
<b>7. RATIOS DE ENDEUDAMIENTO</b>		
	<b>Ratio de Independencia Financiera:</b>	
R64		$\frac{\text{Capital Contable}}{\text{Activo Total}}$
R70		$\frac{\text{Activo Total}}{\text{Capital Contable}}$

Adaptado de: “Análisis de las Dificultades Financieras de las Empresas en una Economía Emergente: las bases de datos y las Variables Independientes en el Sector Hotelero de la Bolsa Mexicana de Valores”, por L. Ibarra, 2001. <https://www.tesisenred.net/handle/10803/3941#page=1>.

En su investigación concluyen que la tendencia de las razones no siempre muestra si una empresa tiene problemas de insolvencia o va a la quiebra, pues puede existir una manipulación en los estados financieros, específicamente en aquellas partidas que son parte de los componentes de las razones más representativas a la hora de analizar la eficiencia financiera de una empresa, por esto es importante estructurar una base de datos confiable a fin de mitigar posibles inconsistencias de la información financiera de las empresas y mejorar los resultados de los modelos de predicción de insolvencia financiera y quiebra.

Los ratios que dieron mejores resultados predictivos en la investigación fueron: rentabilidad de la empresa (beneficios netos después de intereses antes de impuestos sobre activo Total), rentabilidad acumulada (beneficios netos no distribuidos sobre activo total), rotación de activos (ventas netas sobre activo Total), liquidez (activo corriente sobre pasivo total), cash flow (flujo de efectivo sobre pasivo total), solvencia global (pasivo total sobre activo total).

### 3.2.4. Modelos de Puntuación Z

Este documento fue realizado con el fin de identificar, analizar y evaluar todos aquellos modelos de puntuación Z además de presentar dos modelos inéditos identificando de manera crítica sus características sobresalientes y la utilidad de las técnicas. Así pues, se presentan los siguientes modelos.

#### *Modelo Lis (1972)*

El modelo se basa en ratios expresados por Altman, 1968 y se da mediante una función discriminante de cuatro variables, para lo cual se utiliza como muestra en quiebra 30 fracasos empresariales de distintos sectores dentro de los cuales están el sector de la fabricación, la construcción y comercialización, empresa de igual tamaño relacionada por sector con igual números de activos y con base en información del mismo periodo correspondiente. El modelo se expresó mediante la siguiente función:

**Figura 18**

*Razones utilizados por el Modelo de Lis (1972)*

<b><math>Z = 0.063 X1 + 0.092 X2 + 0.057 X3 + 0.0014 X4</math></b>	
X1	$\frac{\text{Capital Circulante}}{\text{Activos Totales}}$
X2	$\frac{\text{Beneficios antes de Intereses e Impuestos}}{\text{Activos Totales}}$
X3	$\frac{\text{Beneficios no Distribuidos}}{\text{Activos Totales}}$
X4	$\frac{\text{Patrimonio Neto}}{\text{Deuda Total}}$
Z	Indice General

*Nota:* Los beneficios no distribuidos son ajustados para las emisiones de acciones especulativas.

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), p.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

**Modelo Taffler (1974)**

Este modelo cuenta con una capacidad predictiva con antelación (ex ante). Para su desarrollo se tuvo en cuenta tres variables discriminantes que fueron: ratios convencionales, medidas de tendencia a cuatro años y variables de estado de los fondos, cabe destacar que el análisis del mismo se basó en 50 razones financieras; El modelo finalmente se transformó a 5 variables provenientes en gran medida del uso del análisis de componentes principales y estas son:

**Figura 19**

*Razones utilizadas por el Modelo de Taffler (1974)*

No	Razón Financiera
1	$\frac{\text{Beneficios antes de intereses e impuestos}}{\text{Activo Total al inicio del Ejercicio}}$
2	$\frac{\text{Deuda Total}}{\text{Capital}}$
3	$\frac{\text{Activos Defensivos}}{\text{Activo Total}}$
4	$\frac{\text{Activos Circulantes}}{\text{Fondos Propios}}$
5	Rotacion de Stocks

*Nota:* Los beneficios no distribuidos son ajustados para las emisiones de acciones especulativas.

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), p.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Es fundamental recalcar la importancia de las dos primeras variables en el desarrollo de este modelo; una vez realizada la prueba u de lachenbruch (método para seleccionar

variables productoras) se obtuvo un error de tipo I y ningún error de tipo II. Finalmente, el modelo no arrojó los resultados esperados ya que de las 52 empresas fallidas el 40,4% fueron clasificadas erróneamente, esto básicamente por la poca importancia que Taffler le dio a la capacidad del practicante de hacer la regla de los ajustes; su interés era examinar cuestiones estadísticas y conceptuales situación que lo condujo a centrarse mucho en cuestiones teóricas que sirvieran en el desarrollo de un modelo operativo Z-score.

### ***Modelo Tisshaw (1976)***

Este modelo fue inspirado en la necesidad de desarrollar una Z separado y distinto para aquellas empresas que cotizaban en mercados capitales y aquellas que no; el objetivo del mismo era analizar empresas manufactureras de propiedad privada; para esto se basó en 31 empresas que fallaron en un periodo de 18 meses y utilizó el enfoque convencional análisis discriminante lineal (LDA) y un conjunto de coeficientes financieros llegando finalmente a un modelo que consiste en 5 variables que son:

**Figura 20***Razones utilizados por Tisshaw (1977)*

No	Razón Financiera
1	$\frac{\text{Beneficio antes de interes e impuestos}}{\text{Exigible a Corto Plazo}}$
2	$\frac{\text{Activo circulante}}{\text{Exigible Total}}$
3	$\frac{\text{Exigible a Corto Plazo}}{\text{Activo Total}}$
4	Intervalo sin credito
5	Rotacion de Stocks

*Nota:* Los beneficios no distribuidos son ajustados para las emisiones de acciones especulativas.

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Este modelo obtuvo un error tipo I y un error de tipo II; lo que Tisshaw hizo fue modificar el límite para las probabilidades previas de 1:5 (posibles fallos: preocupaciones sólidas en la población) para de esta manera demostrar que el análisis de las cuentas de aquellas empresas que no cotizan en bolsa era factible.

### ***Modelo Taffler (1977)***

Este modelo fue realizado con el fin de demostrar la capacidad de predicción ex ante y los resultados de su uso en práctica, su desarrollo se basó en 46 empresas manufactureras cotizadas en bolsa y se utilizó un paquete discriminante lineal para llegar a las siguientes 4 variables:

**Figura 21***Razones utilizadas por Taffler (1977)*

No	Razón Financiera
1	$\frac{\text{Beneficio antes de impuestos}}{\text{Exigible a Corto Plazo}}$
2	$\frac{\text{Activo corriente}}{\text{Pasivo Corriente}}$
3	$\frac{\text{Pasivo Corriente}}{\text{Activo Total}}$
4	Intervalo sin credito

*Nota:* Los beneficios no distribuidos son ajustados para las emisiones de acciones especulativas.

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), p.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Una vez aplicada la prueba  $u$  de lachenbruch se obtuvo un error de tipo I y ningún error de tipo II; cabe destacar que los coeficientes median la rentabilidad, la situación del capital circulante, el riesgo financiero y a liquidez respectivamente. La gran mayoría de las empresas con las que se probó el modelo fueron correctamente clasificadas como empresas en situación de riesgos lo que, en efecto fue un resultado alentador y propicio para su validación.

Es importante destacar que una puntuación  $Z$  en riesgo hace referencia a la descripción de una empresa que en su conjunto tiene similitudes con un grupo de fallos anteriores; así mismo se demostró la verdadera capacidad de predicción *ex ante* pues los eventos que se pudieron predecir estaban relacionados con la existencia o no de dificultades financieras de una empresa; en esta misma línea Taffler describió un avance del enfoque  $Z$ -score que

recibiría el nombre de Z-score de Z-scores y que respondería a un pronóstico de la probabilidad real de que una empresa falle en un lapso de tiempo de 1 año.

Aquí se concluye que los resultados con mayor poder de discriminación fue el ratio de Beneficio antes de impuestos entre Exigible a corto plazo (Taffler, 1984).

Con respecto a las limitaciones del análisis discriminante a través de la utilización aislada de la *Z-Score*, opinaba que esta herramienta era más útil cuando se consigue un desarrollo complementario que proporciona una aproximación a los diferentes grados de riesgo (Taffler, 1984).

### ***Modelo del Banco de Inglaterra (1979)***

La base de justificación de este modelo fue que los modelos anteriores tenían una baja precisión y rendimiento. Para su desarrollo se usó la mejor relación (flujo de efectivo/deuda total) de Beaver, 1966 y las funciones de Deakin, 1977 y Taffler, 1977; la muestra que se utilizó provenía de empresas cotizadas del sector de la fabricación y la distribución las cuales habían fracasado en un lapso de tiempo de 3 años.

Las variables financieras usadas correspondían a 47 coeficientes convencionales y 12 calculados y así mismo se utilizó un enfoque de probabilidad lineal; luego entonces el modelo resultante comprendía las siguientes variables:

**Figura 22***Razones utilizadas por Taffler (1979)*

<b>No</b>	<b>Razón Financiera</b>
1	$\frac{\text{Activos Corrientes}}{\text{Activo Total Bruto}}$
2	$\frac{1000}{\text{Activos Totales Brutos}}$
3	$\frac{\text{Flujo de Efectivo}}{\text{Pasivo Corriente}}$
4	$\frac{\text{Flujo de Fondos}}{\text{Pasivo Total}}$

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

La importancia de los ratios en el desarrollo de este modelo fue establecida en un orden descendente así, X3 (rentabilidad), X2 (tamaño), X1 (liquidez) y X4 (flujo de fondos); a su vez, este modelo arrojó un error de tipo I y cinco errores de tipo II. Es así como se concluye que los malos resultados de este modelo provienen de distintos factores que se debieron tener en cuenta para su desarrollo como lo son la correcta elección de las empresas para la muestra, la evaluación de aspectos que difieren de unas empresas a otras tanto a sus características financieras y la selección correcta como a la medida de tamaño de las empresas que sirven de muestra.

**Modelo Mason y Harris (1978)**

Este modelo se centra específicamente en aquellas empresas de construcción que se encuentran en peligro de fracaso; para su desarrollo se usaron 28 variables y se usó el análisis discriminante lineal (LDA) convencional del cual se derivó la siguiente función:

**Figura 23**

*Razones utilizados por Taffler (1978)*

$Z = 25.4 - 51.2 X1 + 87.8 X2 - 4.8 X3 - 14.5 X4 - 9.1 X5 - 4.5 X6$	
X1	$\frac{\text{Beneficio antes de intereses e impuestos}}{\text{Activos Netos de Apertura}}$
X2	$\frac{\text{Beneficios antes de intereses e impuestos}}{\text{Capital Neto de Apertura Empleado}}$
X3	$\frac{\text{Deudores}}{\text{Acreedores}}$
X4	$\frac{\text{Pasivos Corrientes}}{\text{Activos Corrientes}}$
X5	Log10 (deudores por días)
X6	Medición de la tendencia de los acreedores
Z	Índice General

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), p.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Este modelo no arrojó ninguno error de tipo II ya que una vez aplicada la prueba de resustitución se encontró que ninguna de las 40 empresas se clasificó erróneamente, sin embargo, se obtuvo un error de tipo I en una muestra de validación de 11 empresas. Es así como este modelo demuestra la utilidad de Z-técnicas de puntuación incluso en la aplicación a lo que es una industria heterogénea, sin embargo, este a su vez es propenso al riesgo de sesgos

muéstrales, además de que la muestra de validación en su clasificación arroja una alta tasa de error.

### *Modelo de Datastream (1980)*

Este modelo fue diseñado con el fin de proporcionárselo al servicio de información financiera computarizada en línea, su uso esta enfatizado en el cribado y fue desarrollado por Marais. Para la muestra se establecieron un grupo de empresas que fracasaron en un lapso de 6 años más 60 empresas agregadas, las variables comprendían un conjunto de 40 coeficientes financieros junto con otros que provenían del estado de corrientes de fondos, una vez realizadas las depuraciones se obtuvieron las siguientes variables:

**Figura 24**

*Razones Utilizadas en el Datastream*

No. Razón	Componentes de la Razón
1	$\frac{\text{Beneficio antes de impuestos + amortizaciones}}{\text{Pasivos Corrientes}}$
2	$\text{Prueba de fuego: } \frac{\text{Activos Rapidos}}{\text{Pasivos Corrientes}}$
3	$\text{Valor contable: } \frac{\text{Deuda Total}}{\text{Pasivos Corrientes}}$
4	Volumen de negocios

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), p.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Una vez aplicada la prueba u de lachenbruch un error de tipo I y seis errores de tipo II; dentro de las características importantes del datastream se encuentra la clasificación de 15 tipos de tendencia de puntuación Z. En general, parece ser un modelo de utilidad si se tiene en cuenta el gran uso que las personas le dan, sin embargo, los problemas que presenta son

evidentes y tienen que ver en gran medida con aquellos problemas identificado con el modelo del Banco de Inglaterra, es decir factores como la correcta elección de las empresas para la muestra, entre otros.

***Modelo Betts y Belhoul (1982 y 1983)***

Este modelo fue construido como una alternativa para la inspección física de las empresas dentro de la economía, su desarrollo consta de las siguientes cinco variables:

**Figura 25**

*Razones Utilizadas en el Modelo de Betts y Belhoul*

<b>No. Razón</b>	<b>Componentes de la Razón</b>
1	$\frac{\text{Beneficio antes de impuestos + amortizaciones}}{\text{Activos Totales}}$
2	$\frac{\text{Activos rápidos}}{\text{Activos Corrientes}}$
3	$\frac{\text{Activos Corrientes}}{\text{Capital Neto Empleado}}$
4	$\frac{\text{Capital Circulante}}{\text{Patrimonio Neto}}$
5	Días acreedores

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

La muestra se basó en dos grupos de 26 y 131 grupos respectivamente cada uno, en donde se obtuvo un conjunto de 26 ratios financieros y se utilizó un enfoque discriminante

lineal, una vez aplicada la prueba u de lachenbruch se obtuvo un resultado de ningún error de tipo I y cinco de tipo II.

Una vez obtenido este resultado, al modelo se introdujeron medidas de estabilidad de ratio además de un índice de descomposición de balance, para proceder con un conjunto fallido de 50 empresas en donde se utilizaron 5 categorías de medidas que fueron ratios financieros convencionales, medidas de estabilidad, medidas de tendencia, índice de descomposición de balance y tamaño; luego entonces se realizaron una serie de procedimientos de los que resultaron las siguientes variables:

**Figura 26**

*Razones Finales usados en el Modelo Betts y Belhoul*

No. Razón	Componentes de la Razón
1	$\frac{\text{Beneficio Neto}}{\text{Activos totales}}$
2	$\frac{\text{Activos totales}}{\text{Deuda Total}}$
3	Prueba de fuego: $\frac{\text{Activos Rápidos}}{\text{Pasivos Corrientes}}$
4	$\frac{\text{Activos Rápidos}}{\text{Activos Corrientes}}$
5	$\frac{\text{Activos Circulante}}{\text{Capital Neto Empleado}}$
6	Acreeedores en días
7	Medida de descomposición del balance

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Este modelo proporcionó errores de tipo I en un 4,3% y errores de tipo II en igual medida.

***Modelo Hennaway y Morris (1983)***

El objetivo de este estudio es mejorar la capacidad de los modelos de predicción a través de la derivación de datos hecha años antes del fracaso, además de incluir indicadores económicos e industriales. La muestra utilizada provenía de 44 empresas de diferentes sectores además de datos extraídos de una base computarizada. Las variables estaban conformadas por 40 ratios financieros y cabe destacar que fue un modelo probado a lo largo de 5 años, así pues, la función derivada de 88 empresas en el primer año fue la siguiente:

**Figura 27**

*Razones utilizados por Hennaway y Morris (1983)*

<b><math>Z = -6.17 + 11.43 X1 + 14.07 X2 + 0.55 X3 - 1.57 X4 + 0.98 X5</math></b>	
X1	$\frac{\text{Beneficio de explotación antes de la amortización}}{\text{Activos Totales}}$
X2	$\frac{\text{Deuda a Largo Plazo}}{\text{Capital Neto Empleado}}$
X3	$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$
X4	Industria de Cantería y Construcción Ficticia
X5	Industria de Distribución Ficticia
Z	Índice General

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Luego, la función derivada de las 86 empresas en el quinto año fue la siguiente:

**Figura 28**

*Razones utilizados por Hennaway y Morris (1983)*

<b><math>Z = -4.86 + 13.50 X1 + 3.11 X2 + 4.80 X3 - 0.97 X4 + 0.68 X5</math></b>	
X1	$\frac{\text{Beneficio antes de intereses e impuestos}}{\text{Activos Totales}}$
X2	Prueba Acida
X3	$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$
X4	Industria de Cantería y Construcción
X5	Industria de Distribución
Z	Índice General

*Nota:* La Prueba Acida es igual al Activo Corriente menos los inventarios sobre los Pasivos Corrientes, Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9).

Una vez realizadas las diferentes pruebas (como la u de lachenbruch) y métodos (como el de validación cruzada) se obtuvo una tasa de éxito alta en todos los casos, llegando a la conclusión de que un modelo discriminante múltiple puede predecir la posible quiebra al menos un año antes de que ocurra y que además la diferencia entre las industrias es un factor que influye.

A continuación, el documento exhibiría dos nuevos modelos que tendrían como propósito el análisis de empresas de distribución y la revisión del modelo de Tisshaw de 1976, respectivamente. Para el desarrollo se llevó a cabo un examen univariado y multivariado de

570 empresas de fabricación y 200 empresa de distribución durante un lapso de tiempo de 6 años y se demostró que, a pesar de la medida de dimensiones similares había una cierta cantidad de las mismas que diferían de las muestras en mejor que índice de correlación  $IC=0.01$ , por tal razón se evidenció la necesidad de tratar por separado las distintas clases de empresas.

Luego entonces, para el desarrollo del siguiente modelo se tomaron 65 empresas y el modelo consistió en las siguientes variables, donde:

**Figura 29**

*Razones posteriores utilizados por Hennaway y Morris (1983)*

No	Razón Financiera
1	$\frac{\text{Flujo de Caja}}{\text{Pasivo Total}}$
2	$\frac{\text{Deuda}}{\text{Activos Rápidos}}$
3	$\frac{\text{Pasivos Corrientes}}{\text{Activo Total}}$
4	Intervalo sin credito

Adaptado de: "Empirical Models for the Monitoring of UK Corporations", R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Aquí los coeficientes fueron interpretados como: rentabilidad, posición de deuda, riesgo financiero y liquidez; así pues, una vez aplicada la prueba u de lachenbruch se obtuvo un error de tipo I y ningún error de tipo II. Cabe destacar que 3 años después de la

implementación de este modelo los resultados que arrojó demostraron un rotundo éxito y la utilidad operativa del mismo quedó demostrada.

### *Análisis de Jackknise*

Este método fue creado para reducir el sesgo en un estimador a través de la reutilización de muestras, su esencia está en separar el impacto de un subconjunto en una estimación proveniente de la muestra total, cabe destacar que la estadística Jackknise se deriva de una medida de los pseudovalores que se obtienen a través de una ponderación y una resta de aquellos valores provenientes del cálculo de la muestra total.

### *Modelo Jackknise modelo no cotizado (1983)*

Para el desarrollo de este modelo se usaron 56 empresas con ejercicios financieros en un lapso de 1 año, en donde además se ajustó el modelo inicial de Fisher y se obtuvieron las siguientes variables:

**Figura 30***Razones utilizados por Hennaway y Morris (1983)*

No	Razón Financiera
1	$\frac{\text{Ganacias antes de intereses e impuestos}}{\text{Ventas}}$
2	$\frac{\text{Deuda}}{\text{Patrimonio Neto}}$
3	$\frac{\text{Acreedores Medios}}{\text{Coste de Ventas}}$
4	$\frac{\text{Pasivos Totales}}{\text{Activos Totales}}$

Adaptado de: “Empirical Models for the Monitoring of UK Corporations”, R. Taffler, 1984, Journal of Banking and Finance, Vol. (8), P.2. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Una vez aplicada la prueba u de lachenbruch se obtuvieron dos errores de tipo I y dos de tipo II; luego entonces el modelo de Jackknise se ajustó al mismo conjunto de variables, pero utilizando 95 submuestras se obtuvo una clasificación idéntica además de coeficientes discriminantes de magnitud similares a la función de Fisher. Es así como la transformación de este modelo proporciona una medida holística del rendimiento, además de la obtención de información a través del desglosamiento de los distintos coeficientes.

### ***La transformación PAS-score***

Este es un modelo de fácil interpretación en tanto a la forma en cómo se deriva y calibra, además de que sus proporciones miden distintas facetas del desempeño de una empresa, también es importante recalcar que es posible observar los límites de escala fijos y la

calibración de la relación estricta con la relación lineal. La justificación del PAS-score en tanto a su transformación e interpretación se da sobre bases teóricas y empíricas.

### *El enfoque Z - score en la práctica*

El enfoque Z-score proporciona una herramienta analítica adicional que además funciona como un dispositivo de detección pues combina el juicio cualitativo y las puntuaciones Z obteniendo finalmente una visión general de una empresa y/o negocio; sus diferentes trayectorias y modificaciones han sido y son propicias para la medición del desempeño de una empresa y la historia de la dirección.

En conclusión, a través de este estudio quedó demostrada la necesidad de la construcción de modelos separados, además de la gran utilidad del enfoque Jackknise en el desarrollo de una función discriminante. Cabe destacar que la resistencia al enfoque Z-score viene dada en gran medida por una errónea idea de que este sustituye de algún modo la intuición y experiencia de aquella persona encargada de la toma de decisiones cuando en realidad su uso está mejor entendido como una herramienta de medición que ayuda en la toma de decisiones basada en habilidad, experiencia y juicios particulares.

### *Modelo Hernández (2014)*

El objetivo de la investigación de Hernández fue explorar, analizar y detallar los modelos de predicción de dificultades financieros de Altman conocidos como Z - Score o en español puntajes Z para el contexto latinoamericano, específicamente en empresas de capital

abierto del sector público en Costa Rica sin embargo, los modelos puede ser utilizados tanto por empresas de gran tamaño como Pymes sin importar que “sean de capital abierto o cerrado. Hernández recomienda su uso en actividades de medición del riesgo crédito y como indicador importante para evaluar los riesgos y la probabilidad de éxito o fracaso a futuro para las empresas del mercado latino” (Hernandez, 2014, p. 14).

El modelo de Altman usa la técnica estadística Análisis Discriminante Múltiple, luego de clasificar las empresas pertenecientes a la muestra de la población de estudio en 2 o más grupos y recolectar la información financiera necesaria para calcular las razones o indicadores se aplicó el MDA, su finalidad es intentar derivar una combinación lineal de las características que mejor discrimine entre los grupos, generalmente los modelos de predicción financiera de Altman manejan 3 grupos o zonas catalogadas como empresas seguras o sin problemas financieros, empresas en quiebra o con problemas financieros y empresas no bien definidas en alguna de las anteriores categorías ignoradas o zona gris.

El resultado del modelo es una función discriminante que tiene la forma  $Z = V_1X_1 + V_2X_2 + \dots + V_nX_n$ , la Función discriminante transforma los valores de las variables particulares en una puntuación discriminante única o valor Z, que luego es utilizado para clasificar el objeto, en donde:

- $V_i$  = son coeficientes discriminantes.
- $X_i$  = son las variables independientes.

El MDA calcula los coeficientes discriminantes  $V_i$ , mientras que las variables independientes  $X_i$  son los valores asociadas a las Razones o indicadores actuales ingresados al modelo, en la siguiente Figura se propuso la función discriminante final de Altman Z – Score

que es considerada como la más conveniente para las industrias manufactureras que se negocian públicamente (E. I. Altman, 2000).

**Figura 31**

*Modelo Score Z de Altman para empresas de Capital Abierto Manufactureras Modelo de Hernández*

$Z = 1.2 (X1) + 1.4 (X2) + 3.3 (X3) + 0.6 (X4) + 0.999 (X5)$	
X1	$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$
X2	$\frac{\text{Utilidades Retenidas}}{\text{Activos Totales}}$
X3	$\frac{\text{Utilidad antes de intereses e impuestos}}{\text{Activos Totales}}$
X4	$\frac{\text{Valor de Mercado del Patrimonio}}{\text{Valor en Libros de total de Deuda}}$
X5	$\frac{\text{Ventas}}{\text{Activos Totales}}$
Z	Índice General

Adaptado de “Modelo financiero para la detección de quiebras con el uso de análisis discriminante múltiple”, por M. Hernández, 2014.

<https://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095>

Posteriormente Altman adaptó su modelo Z – Score para el caso de industrias o empresas genéricas de capital cerrado, reemplazando el valor de mercado del capital - patrimonio por el valor contable del mismo, ya que esta variable si estaba disponible para cualquier empresa que utilice un sistema contable, generando un modelo capaz de predecir la insolvencia o quiebra para el caso de empresas manufactureras y otro modelo para empresas

no manufactureras, a continuación en las siguientes Figuras se muestran los resultados de su investigación.

**Figura 32**

*Modelo Score Z de Altman para Empresas de Capital Cerrado Manufactureras Modelo de Hernández*

$Z = 0.717 (X1) + 0.847(X2) + 3.107 (X3) + 0.420 (X4) + 0.998 (X5)$	
X1	$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$
X2	$\frac{\text{Utilidades Retenidas}}{\text{Activos Totales}}$
X3	$\frac{\text{Utilidades Retenidas Antes de Intereses e Impuestos}}{\text{Activos totales}}$
X4	$\frac{\text{Valor Contable del Patrimonio}}{\text{Valor en Libros del Total de Deuda}}$
X5	$\frac{\text{Ventas}}{\text{Activos Totales}}$
Z	Índice General

Adaptado de “Modelo financiero para la detección de quiebras con el uso de análisis discriminante múltiple”, por M. Hernández, 2014.

<https://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095>

**Figura 33**

*Modelo Score Z de Altman para Empresas de Capital Cerrado en General (Manufactureras y no Manufactureras) Modelo de Hernández.*

$Z = 6.56 (X1) + 3.26 (X2) + 6.72 (X3) + 1.05 (X4)$	
X1	$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$
X2	$\frac{\text{Utilidades Retenidas}}{\text{Activos Totales}}$
X3	$\frac{\text{Utilidades Antes de Intereses e Impuestos}}{\text{Activos Totales}}$
X4	$\frac{\text{Valor Contable del Patrimonio}}{\text{Valor en Libros del Total de Deuda}}$
Z	Índice General

Adaptado de “Modelo financiero para la detección de quiebras con el uso de análisis discriminante múltiple”, por M. Hernández, 2014.

<https://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095>

Adicionalmente Altman en sus modelos de 1968 y 2000 estableció los scores de corte o límites necesarios para poder realizar pronósticos prematuros de potenciales quiebras empresariales, en la siguiente Figura se presentan los rangos de los indicadores para cada score de Z:

**Figura 34**

*Puntos de corte o Límites de los Puntajes Z de Altman Modelo de Hernández*

Predicción	Empresa de Capital Abierto Manufacturera	Empresa de Capital Cerrado Manufacturera	Empresa de Capital Cerrado Genérica
Zona de Quiebra	< 1.81	< 1.23	< 1.1
Zona de Ignorancia (Área Gris)	de 1.81 a 2.99	de 1.23 a 2.90	de 1.1 a 2.6
Zona Segura	> 2.99	> 2.90	> 2.6

Adaptado de “Financial Ratios: Use, Predictive Power and the Z-Score”, por R. Rufus, 2003.

[http://www.rufusandrufus.com/financial\\_ratios.pdf](http://www.rufusandrufus.com/financial_ratios.pdf)

Hernández sugiere la enseñanza e introducción temprana a el modelo de Altman Z – Score tanto en el pregrado como en el postgrado de carreras Administrativas con énfasis en finanzas y contaduría, pues brinda a los estudiosos nuevas experiencias, potenciando el análisis fortaleciendo hipótesis con gran valor agregado, favorables en su misión profesional, mercadeables y pragmáticas Arnold, (2005).

En las conclusiones Hernández señala que el modelo de Altman Z – Score puede ser usado inicialmente para descubrir o anticipar si las empresas han están o estarán en dificultades financieras, útil para que los encargados participen de forma oportuna, también se pueden usar los scores del Z – Score para evaluar la eficiencia de una empresa y realizar estrategias oportunas que permitan tomar decisiones en cuanto a si la empresa se debe reorganizar, liquidar o mejorar su desempeño y su calificación en el mercado.

Finalmente Hernández recomienda dar gran importancia al estudio y análisis de los distintos modelos de salud financiera y pronósticos prematuros de dificultades financieras, insolvencia o quiebras, tomando como base información financiera particularmente de

empresas del contexto latinoamericano a partir de las bases de datos que manejen los países de dicho contexto para las pequeñas y medianas empresas PYMES a fin de corroborar la efectividad o no de estos modelos y determinar o anticipar las causas y efectos potenciales del fracaso empresarial.

### **3.2.5. Regresión Logística y Binaria Logit - Probit**

Logit y Probit son dos modelos introducidos por Joseph Berkson en 1944 y Chester Ittner en 1934 respectivamente. Se basan en la regresión logística y binaria la cual “es una técnica estadística multivariable destinada al estudio de una relación de dependencia entre una variable dependiente y un conjunto de variables independientes” (López-Roldán & Fachelli, 2016, p.5), su finalidad es poder realizar predicciones o pronósticos del comportamiento, se estiman probabilidades de un suceso identificado como la variable dependiente en función de unas variables de pronóstico conocidas como variables independientes, la regresión logística se desarrolló en la década de los años 60 con el trabajo de Cornfield, Gordon y Smith (1961), el estudio de regresión logística tiene dos modalidades:

La regresión logística binaria cuando se pretende explicar una característica o suceso dicotómico (estar desempleado o no, abstenerse en las elecciones o no), y la regresión logística multinomial en el caso más general de querer explicar una variable cualitativa politómica. Para ello se requiere convertir la variable en diversas variables dicotómicas ficticias, es decir, creando tantas variables dicotómicas (dummy) como categorías tenga la variable menos una, la que actuará de categoría de referencia. En este segundo caso se diferencia la

situación en que la variable categórica es politómica nominal (la elección de una marca de un producto o la filiación política) o politómica ordinal (el nivel salarial o el grado de acuerdo sobre una cuestión). (López-Roldán & Fachelli, 2016, p6)

La variable dummy o ficticia toma valores binarios es decir de 0 o 1 para indicar la ausencia o presencia de algún efecto y/o síntoma categórico que logre cambiar el resultado, puede usarse para variables cuantitativas y cualitativas. Los modelos basados en regresión logística y binaria se usan para modelos de predicción de insolvencia y/o quiebra por medio variables provenientes de ratios financieros, las cuales son usadas con el fin de predecir la fragilidad de las empresas, se les asigna una probabilidad que describe el posible resultado quiebra la cual es entendida como una situación de fracaso empresarial asociada a distintos factores provenientes de consecuencias de los ciclos económicos, deficiencias administrativas, insolvencia, cambios en el mercado, entre otros.

### ***Modelo Ohlson (1980)***

El trabajo de Ohlson titulado *Financial Ratios and the Probabilistic Prediction of Bankruptcy* presentó un importante cambio con respecto a otras investigaciones, este se basó en un método estadístico para evaluar la probabilidad de la quiebra. Por primera vez Ohlson utilizó la metodología econométrica del análisis de LOGIT condicional, asociado con el análisis multivariable o más conocido como el análisis discriminante (MDA) y esta vez centro más su atención en problemas metodológicos.

El objetivo era obtener datos de tres años antes de la fecha de quiebra, dentro de los datos seleccionado se destacó el balance general, el estado de resultado, el estado de fondos y el informe contable, si este último informe indicaba que la empresa estaba en quiebra se debía acudir a un cuarto informe denominado 10K (informe condensado de la *U.S. Securities and Exchange Commission*). Dentro de la información algunas empresas tuvieron que ser eliminadas de la muestra porque no se disponía de ningún informe otras porque eran estructuras corporativas y no tenía ventas.

Finalmente la muestra se compuso por 105 empresas que participaban en el mercado de valores y que fracasaron ente los años 1970 y 1976, a esta muestra se suma información de 2058 empresas calificadas como solventes participantes en mercado. Ohlson seleccionó nueve ratios e indico que, se debe prestar especial atención a cuatro elementos que podrían posibilitar el fracaso dentro de los que se evidencia el tamaño de la empresa, la capacidad de la estructura financiera, la rentabilidad, y la liquidez a corto plazo.

Figura 35

Razones Utilizadas en el modelo de Ohlson

No	Razón Financiera
1	$\frac{\text{Logaritmo de los Activos Totales}}{\text{Indice de Precios}}$
2	$\frac{\text{Pasivos Totales}}{\text{Activos Totales}}$
3	$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$
4	$\frac{\text{Activos Corriente}}{\text{Pasivo Corriente}}$
5	Variable Dummy de Solvencia: La razón es igual a 1 en el caso que el Pasivo Total es mayor al Activo Total, en caso contrario la razón es igual a 0.
6	$\frac{\text{Retorno sobre los activos o EBIT}}{\text{Activos Totales}}$
7	$\frac{\text{Resultado Operacional}}{\text{Obligaciones Totales}}$
8	Variable Dummy de Rentabilidad: La razón es igual a 1 cuando los ingresos en los últimos dos años han sido negativos, en caso contrario la razón es igual a 0.
9	$\frac{(\text{Ingreso Neto del Periodo Actual} - \text{Ingresos Netos del Periodo Anterior})}{\text{Ingreso Neto Actuañ} + \text{Ingreso neto del periodo Anterior}}$

*Nota:* El ratio número 9 muestra el cambio en el ingreso neto de un periodo a otro actuando, el valor absoluto del denominador es un indicador de nivel. Adaptado de: "Financial Ratios and the Probabilistic Prediction of Bankruptcy", J. Ohlson, 1980, Journal of Accounting Research, Vol. (18), p.1.  
<https://doi.org/10.2307/2490395>

Como resultados obtenidos del modelo realizado por Ohlson se encuentra que, clasificó correctamente:

El 96.12% de la muestra de empresas industriales con probabilidad de quiebra, los resultados indicaron que la combinación de la tasa de error tipo I y II arrojaron un punto límite de 0.38; lo que muestra que, el 17,4% y 12.4% de las empresas en quiebra No fueron correctamente clasificadas (Ohlson, 1980, p. 129).

Una característica del estudio de Ohlson es que los datos de las empresas quebradas fueron tomados de desempeño financiero requerido por la comisión de bolsa y valores, y con este proceso se podía conocer de acuerdo con la publicación si, la empresa había entrado en quiebra antes o después de su publicación, situación que evitaría sobre-predecir las empresas quebradas.

### **3.2.6. Predicción de Bancarrota: Una Comparación de Técnicas Estadísticas y de Aprendizaje Supervisado para Computadora. Modelo Z – Score y Procesos Gaussianos**

El modelo Z – Score introducido en 1968 por Edward I. Altman propone una generalización a través de la estimación de un estadístico multivariado mientras que los procesos Gaussianos son algoritmos recientemente desarrollados que se dan en términos de distribuciones de probabilidad a priori potencialmente de gran poder para predecir la bancarrota en las empresas, que es el punto central del estudio a realizar.

Un algoritmo es un conjunto de instrucciones, pasos, reglas u operaciones sistémicas ordenadas secuencialmente de forma lógica con el fin de resolver un problema, realizar una acción o llevar a cabo una actividad o tarea, un algoritmo puede ser más que una instrucción, dictan sus parámetros de funcionamiento porque “no solo determinan la elección de qué datos serán recolectados (y cuáles otros no) sino también qué tipo de correlaciones e instrucciones (y cuáles no) se tendrán con respecto a los mismos y cuales efectos y/o resultados se esperan”, (Gendler, 2019, p. 304).

Así pues se presenta una comparación entre las técnicas estadísticas clásicas y los más recientes algoritmos con el fin de predecir las bancarrotas de una forma probabilística, esto centrado particularmente en las entidades bancarias ya que estas instituciones tienen el deber de satisfacer requerimientos legales y contables al referirse a los tributos o impuestos propios de las autoridades financieras, de esta manera, el objetivo del estudio es determinar u obtener un sistema que prediga si una institución ira a quiebra en base a algún tipo de información financiera.

Para el desarrollo de dicho estudio se contó con información proporcionada por la Universidad de Essex, en donde se encuentra un igual número de bancos estadounidenses que fueron y no fueron a la quiebra, estos a su vez fueron separados en pares, es decir, un banco que fue a la quiebra junto con uno que no fue a la quiebra; una vez realizadas las respectivas depuraciones se obtuvo un total de cien pares y se establecieron once razones financieras a realizar, como se evidencia en la siguiente Figura.

Figura 36

*Razones Financieras Predicción de la Bancarrota Utilizadas por Peña, Martínez, Abudu*

No	Ratios	Composición
1	Margen de Interés Neto	No disponible.
2	$\frac{\text{Ingresos} - \text{Intereses}}{\text{Activos Productivos}}$	Basados en honorarios, comercio, resultados de actividades financieras y otros relacionados con el interés.
3	$\frac{\text{Gastos} - \text{Intereses}}{\text{Activos Productivos}}$	Comprende tres tipos de gastos: gastos personales, tendencia y otros gastos operativos.
4	$\frac{\text{Ingreso Operativo Neto}}{\text{Activos Totales}}$	Relacionado al ingreso bruto de una empresa asociado con sus propiedades menos los gastos operativos.
5	Rentabilidad Económica	Razón entre las ganancias totales de una empresa durante el año y sus activos totales.
6	Rentabilidad Financiera	Ganancias Netas.
7	Razón de Eficiencia	No disponible.
8	$\frac{\text{Activos Fijos}}{\text{Activos No corrientes} + \text{Otros Inmuebles}}$	No son convertibles fácilmente en efectivo
9	Razón de Efectivo + Obligaciones de gobierno y del tesoro de E.E.U.U.	No disponible.
10	$\frac{\text{Patrimonio Neto}}{\text{Activos Totales}}$	No disponible.
11	Razón de Capital primario	Recursos propios sobre Activos ponderados por el riesgo.

*Nota:* Otros inmuebles en posesión de los Activos. Adaptado de: "Predicción de Bancarrota: una comparación de Técnicas Estadísticas y de Aprendizaje Supervisado para Computadora Peña et al., 2009. <https://doi.org/10.36095/banxico/di.2009.18>.

Luego entonces, se probaron cinco algoritmos diferentes para los datos con los que se contaba, estos fueron: el análisis de discriminantes de Fisher (FDA), máquinas de soporte vectorial de mínimos cuadrados (LS-SVM, por sus siglas en inglés), procesos gaussianos

(GP's) para la clasificación (GPC), GP's deformados (WGP) y discriminante Bayesiano de Fisher (BFD), con el fin de obtener un promedio de clasificación en tanto a la media, la mediana, el máximo, el mínimo y la desviación estándar del área porcentual, como se evidencia en la siguiente Tabla.

**Tabla 2**

*Resultados promedio de clasificación de los datos del Seguro de Depósito de los Estados Unidos*

*Modelo de Peña, Martínez, Abudu*

	FDA	Logística	Probit	LS-SVM (rbf)	BFD(rbf)	GPC (rbf)
Media	0,866	0,839	0,825	0,823	0,817	0,815
Mediana	0,877	0,841	0,838	0,818	0,816	0,815
Máximo	0,962	0,949	0,940	0,956	0,950	0,919
Mínimo	0,672	0,679	0,678	0,687	0,681	0,676
STD	0,051	0,056	0,055	0,055	0,051	0,05

Adaptado de: "Predicción de Bancarrota: una comparación de Técnicas Estadísticas y de Aprendizaje Supervisado para Computadora Peña et al., 2009. <https://doi.org/10.36095/banxico/di.2009.18>.

De los resultados obtenidos se entiende que los GP's o procesos Gaussianos pueden ser una herramienta útil para las predicciones en la vida real, sin embargo, presentan dificultades en tanto que los efectos de varianza conllevan a una superposición de muestras lo que causa que el computo dé la desviación estándar.

Se procede a realizar un análisis de las características a través de algunas funciones de covarianza tipo determinación automática de relevancia (ARD) en tanto a la determinación de bancarrota en las entidades bancarias a través de las distintas razones financieras, las cuales fueron divididas en dos grupos, las más relevantes y las menos relevantes, dentro de las más

relevantes se encuentran la rentabilidad financiera (ROE), rentabilidad económica (ROA), razón de eficiencia (ER) e ingreso operativo neto (NOI) mientras que en las menos relevantes se encuentran el margen de interés neto (NIM), ingresos sin ingresos por interés (NII) y la razón de capital (CR).

Una vez realizada la revisión comprensiva de los modelos y/o métodos expuestos, se llega a la conclusión de que instancias diferentes de GP's generan resultados de clasificación potencialmente competitivos con respecto a modelos como Z – Score y regresión logística; cabe destacar que las GP's son una generalización de la densidad Gaussiana y se aplica para tareas de inferencia Bayesiana, dadas sus propiedades analíticas simples y su facilidad de uso.

### **3.2.7. Modelos y Aplicaciones para Colombia**

#### ***Modelo Rosillo (2002)***

Los criterios utilizados para clasificar las empresas en fuertes y débiles, fue clasificarlas y ordenarlas de acuerdo con el valor obtenido en el indicador o razón; los indicadores utilizados por Rosillo fueron 12 para el análisis discriminante, pero el sistema paso a paso los redujo a los siguientes tres ratios que finalmente establecieron la diferencia entre las empresas fuertes y débiles.

**Figura 37***Indicadores utilizados por Rosillo (2002)*

No	Indicadores Financieros
1	$\frac{\text{Endeudamiento Ganancias antes de intereses e impuestos}}{\text{Ventas}}$
2	Rentabilidad del Patrimonio (ROE)
3	Leverage

*Nota:* El ROE es igual a la Deuda sobre el Patrimonio Neto, el Leverage es igual a los Pasivos Totales sobre el Patrimonio. Adaptado de: “Moldeo de Predicción de Quiebras de las Empresas Colombianas” por J. Rosillo, 2002, Journal Revista de Ciencias Administrativas y Sociales, volumen (1), p.19. <https://revistas.unal.edu.co/index.php/innovar/article/view/23959>.

A partir del análisis discriminante catalogado como un instrumento estadístico, el principal uso que se le ha dado es, en la predicción de quiebra de empresas. El Análisis discriminante permite por el método paso a paso definir qué variables independientes del total de las tomadas inicialmente, son los mejores discriminadores que cumplen los niveles de significancia de la prueba.

Rosillo con la anterior aplicación probó ochenta (80) empresas y la clasificación que se obtuvo de esa prueba fue la siguiente: efectividad de un 94% para las empresas fuertes y un 87% para las débiles. Posteriormente tomó 26 empresas que no hicieron parte del cálculo en el proceso del Z discriminante, estas fueron clasificadas como 15 fuertes y 11 débiles. Los resultados fueron 100% de aciertos para las empresas fuertes y 82% para las débiles. Concluye que los tres indicadores clasificados como fundamentales para el diagnóstico de empresas fueron el nivel endeudamiento (pasivo total / patrimonio), rentabilidad del patrimonio

(utilidad o ingresos netos / patrimonio) y leverage (pasivo total / activo total), que dirigidos a las empresas fuertes, presentaron un valor promedio de endeudamiento de 26% que es relativamente bajo, puesto que 75% es el máximo exigido a los solicitantes de crédito por las entidades financieras, rendimiento de capital del 6,11%, y, compromisos cubiertos por el 50% del capital. Las empresas débiles mostraron una situación de apalancamiento cercano al 50%, un rendimiento negativo y un pasivo respaldado por el 100% del patrimonio con notable riesgo. Con estos resultados concluye que la función discriminante es un buen predictor o instrumentos de pronóstico.

### ***Modelo de Martínez (2003)***

La finalidad de Martínez en su investigación *titulada Determinantes de Fragilidad en las empresas colombianas* fue identificar los determinantes de la insolvencia en el año 2001, para dicho fin se utilizó los estados y/o información financiera que cada empresa declaró en el año 2000. Según (Martínez 2003) “la heterogeneidad en la estructura institucional, en las prácticas contables y en el comportamiento de las variables macroeconómicas a lo largo del tiempo, no permite una generalización de los resultados obtenidos en otros países” (p. 1), por esta razón fue conveniente desarrollar modelos financieros propios para cada país.

Para el contexto Colombiano por una parte Rosillo en el año 2002 fue el primer investigador en presentar un modelo de predicción de fragilidad empresarial y quiebra usando técnicas estadísticas como análisis discriminante en un limitado tamaño de muestra, por otro parte Martínez en su investigación utilizó la técnica probabilística y estadística de regresión

PROBIT para un considerable tamaño de muestra con la cual desarrollo un modelo en que se identificaron las variables relevantes para predecir el estrés o fragilidad financiera.

Su muestra se compuso de los estados financieros de 9000 empresas colombianas de diferentes sectores económicos reportados en el año 2000 a la Superintendencia de Sociedades. Se encontraron dentro del grupo de las 9000 empresas 171 empresas que presentaban estrés o fragilidad, inicialmente definió que la variable dependiente del modelo era el estado de estrés o fragilidad empresarial y como variable independiente el conjunto de razones o indicadores que cubrían los aspectos de liquidez, rentabilidad y endeudamiento generalmente aceptados en la literatura como decisivos de la fragilidad empresarial, a continuación en la Figura 38 se resumen las razones utilizadas en la investigación de Martínez, agrupadas por cada aspecto junto con una pequeña descripción:

**Figura 38**

*Razones utilizadas en el modelo de Martínez*

Categoría del Indicador o Razón	Indicador o Razón	Descripción
Liquidez	1. $\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$	Capturan la relación entre los activos fácilmente realizables y el endeudamiento de corto plazo, y el nivel de liquidez como proporción del activo de cada empresa. En la medida en que exista un amortiguador de liquidez que permita mantener la operación sin afectar el pago a deudores, la empresa estará más lejos de un posible estado de insolvencia.
	2. $\frac{\text{Efectivo}}{\text{Pasivo Corriente}}$	
	3. $\frac{\text{Capital de trabajo}}{\text{Activos Totales}}$	
	4. $\frac{\text{Efectivo}}{\text{Activos Totales}}$	
Rentabilidad	1. $\frac{\text{Utilidad Operacional}}{\text{Activos Totales}}$	Mide la cantidad de ingresos que cada unidad de activo es capaz de generar.

Categoría del Indicador o Razón	Indicador o Razón	Descripción
	2. $\frac{\text{Utilidad antes de Impuestos}}{\text{Activos Totales}}$	Miden la rentabilidad del negocio una vez se han pagado los servicios de deuda y operación.
	3. $\frac{\text{Utilidad Antes de Impuestos}}{\text{Utilidad Operacional}}$	
Endeudamiento	1. $\frac{\text{Pasivos Totales}}{\text{Activos Totales}}$	Miden el grado de apalancamiento de la empresa, que, en el caso de ser alto, compromete la capacidad de pago a deudores ante choques negativos no esperados.
	2. $\frac{\text{Obligaciones Financieras}}{\text{Activos Totales}}$	
	3. $\frac{\text{Egresos Financieros}}{\text{Ingreso Operacional} + \text{I. Financiero}}$	Captura el efecto de los flujos de efectivo necesarios para cumplir con el pago de intereses y que puede dar origen a presiones financieras

Adaptado de: “Determinantes de Fragilidad en las Empresas Colombianas”, por O. Martínez-A., 2003.

<http://repositorio.banrep.gov.co/handle/20.500.12134/2096>

A continuación a las razones financieras se les calculó mediante estadística descriptiva la media y su desviación estándar y se aplicó el logaritmo natural para cada una de las razones, posteriormente se realizaron las estimaciones Probit desarrollando de esta forma 4 modelos, para el modelo 1 y 3 se determinó que en ambos casos se rechazó la hipótesis nula de homocedasticidad, lo que llevó a la estimación Probit heteroscedástica en los modelos 2 y 4, para más información revisar la investigación de Martínez. Los modelos heteroscedásticos clasificaron correctamente el 82% de las empresas frágiles y no frágiles, los resultados confirmaron la importancia de 3 razones relacionadas con la rentabilidad, endeudamiento y liquidez en la solvencia presentada por las empresas, las cuales son las razones de utilidad o ingresos antes de impuestos sobre activos totales, obligaciones financieras sobre activos totales y efectivo sobre activos totales respectivamente.

*Modelo Pérez, González, Lopera (2011)*

El objetivo de su investigación fue estimar un modelo actualizado que permita medir la probabilidad de que una empresa entre en quiebra o insolvencia financiera usando los modelos probabilísticos de selección discreta PROBIT y LOGIT, para dicho fin se usó como fuente la base de datos de los estados financieros reportados por las empresas a la Superintendencia de Sociedades de Colombia.

Inicialmente la muestra se compuso de 28387 empresas vigiladas por la Superintendencia de Sociedades con información a 31 de diciembre de 2011, al depurar la muestra se eliminaron las empresas inconsistentes o que no contaban con la información requerida, eso redujo la cantidad de empresas a 25756. Las empresas se agruparon según su sector económico y posteriormente se clasificaron en dos categorías frágiles (con problemas de insolvencia o quiebra) y no frágiles, a través de los estados financieros reportados se construyeron razones financieras de liquidez, endeudamiento y rentabilidad, las cuales son variables explicativas en la estimación de los modelos estadísticos de regresión binaria

Estas razones o variables son conocidas dentro del modelo como variables independientes, las cuales emplean una función lineal con una variable dependiente que es definida como fragilidad empresarial y hace referencia a la posibilidad de que una empresa pueda llegar a la quiebra o tener insolvencia financiera. Las razones financieras fueron escogidas de la literatura y se incluyeron variables del modelo de Rosillo y Martínez para conservar un buen nivel predictivo ya que tienen una mejor interpretación económica, las razones se pueden categorizar en: razones de liquidez, razones de endeudamiento y razones de

rentabilidad, las razones de rotación no fueron tomadas en cuenta excepto la razón rotación de activos, en la siguiente Figura se puede observar las razones específicas de cada categoría.

**Figura 39**

Razones Financieras usadas en el Modelo de Pérez, González y Lopera

<b>Categoría del Indicador Financiero</b>	<b>Indicador Financiero</b>	<b>Definición Operativa del Indicador</b>
Indicadores de Liquidez	Razón Corriente	$\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$
	Prueba Acida	$\frac{\text{Activos Corrientes} - \text{Inventarios}}{\text{Pasivos Corrientes}}$
	Razón de Tesorería	$\frac{\text{Caja} + \text{Bancos}}{\text{Pasivos corrientes}}$
Indicadores de Rotación	Rotación de Activos	$\frac{\text{Ventas}}{\text{Activos Totales}}$
Indicadores de Rentabilidad	Margen Neto	$\frac{\text{Utilidad Neta}}{\text{Ventas}}$
	Rentabilidad del Activo	$\frac{\text{Utilidad}}{\text{Activos Totales}}$
Indicadores de Endeudamiento	Apalancamiento a Corto Plazo	$\frac{\text{Pasivo Corriente}}{\text{Patrimonio}}$
	Apalancamiento Total	$\frac{\text{Pasivos Totales}}{\text{Patrimonio}}$

Adaptado de “Modelos de predicción de la fragilidad empresarial: Aplicación al caso colombiano para el año 2011”, por J. Pérez et al., 2013. <https://www.redalyc.org/pdf/861/86131758010.pdf>

Las Razones fueron calculadas tanto para las empresas frágiles y no frágiles, a estas razones se les calculo las medidas estadísticas de tendencia central y dispersión media y desviación estándar respectivamente, se realizó una prueba de hipótesis nula con la finalidad de revisar cuales razones eran más confiables para discriminar entre empresas. Se estimaron los modelos de selección discreta LOGIT el cual clasifico correctamente el 83,20% de las observaciones y PROBIT el cual clasifico correctamente el 80,8% de las observaciones, uno de los principales problemas del modelo es que al ser datos gubernamentales existe la probabilidad de que la información sea afectada siendo manipulada o distorsionada por prácticas como la denominada contabilidad creativa o el efecto window dressing, siendo una problemática en la práctica contable en general, la Figura 40 a continuación resume los resultados obtenidos.

**Figura 40**

*Proporción de empresas clasificadas correctamente Modelo LOGIT y PROBIT modelo de Pérez, González y Lopera*

<b>EMPRESAS</b>	<b>LOGIT</b>	<b>PROBIT</b>
NO FRAGILES	88,88%	85,71%
FRAGILES	77,41%	75,80%
<b>PROMEDIO TOTAL DE CLASIFICACION</b>	<b>83,15%</b>	<b>80,76%</b>

Adaptado de “Modelos de predicción de la fragilidad empresarial: Aplicación al caso colombiano para el año 2011”, por J. Pérez et al., 2013. <https://www.redalyc.org/pdf/861/86131758010.pdf>

En las estimaciones de los modelos (LOGIT Y PROBIT) se destacó un problema de multicolinealidad el cual es entendido como un alto coeficiente de correlación entre las distintas variables, sin embargo, este no representa un impedimento para la realización de las

distintas estimaciones. En sus resultados se puede evidenciar que las razones financieras más relevantes para medir la probabilidad de riesgo de quiebra o insolvencia financiera son: la rentabilidad del activo, rotación del activo y el nivel de endeudamiento.

### **3.2.8. Redes Neuronales Artificiales (RNA)**

Las RNA (Redes Neuronales Artificiales) o ANN (Artificial Neural Networks) “se inspiran en las redes neuronales biológicas del cerebro humano. Se constituyen de elementos que se comportan de forma homologa a la neurona biológica en sus funciones más comunes” (Matich, 2001, p. 4) Aprender, Generalizar y Abstraer, son similares a los algoritmos o secuencias de instrucciones, existen numerosas formas de definir a las redes neuronales como un modelo matemático o “sistema de computación que está compuesto por un gran número de elementos simples, elementos de procesos muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas” (Matich, 2001, p. 8).

En sus diversas aplicaciones están el análisis de señales, reconocimiento de imágenes, control de procesos, filtrado de ruido, robótica, procesamiento de lenguaje, diagnósticos médicos, predicción, modelamiento de sistemas, modelos económicos y financieros, valoración del riesgo de los créditos, explotación de bases de datos, análisis de tendencias y patrones, previsión de la evolución de los precios. entre otros (Matich, 2001).

#### ***Modelo Pessoa de Oliveira (2016)***

Pessoa, (2016) estudio a través de los estados contables (información contable) la situación de solvencia de las empresas, empleando los modelos *estáticos* vs el modelo *dinámico*, el autor toma el diagnostico financiero de las empresas como uno de elementos

principales para el análisis de informes en un periodo determinado, lo cual permite comprender los resultados de la situación actual de las mismas, se encontró con ciertas limitaciones debido a que, no tuvo acceso a la información para realizar un diagnóstico completo, desde lo cuantitativo y cualitativo que pudieran ayudar a determinar qué otras causas pudieran provocar solvencia e insolvencia en las empresas.

El objetivo propuesto para esta investigación fue probar con objetividad cuál de los dos modelos (estático vs dinámico) utilizados en la evaluación de la solvencia de las empresas es más equilibrado. Para realizar esta tarea tomo dos grupos de empresas distintas. El primero lo conformaron empresas (S.A) que se encontraban en concurso de acreedores, clasificadas como insolventes. El segundo grupo lo conformo las empresas (S.A) que en el período de 2008 a 2012 cotizaban en el mercado de valores de la bolsa de Madrid, a excepción de las entidades financieras y aseguradoras, cuyo objetivo es aumentar su capital, crecer en el mercado y tener una alta participación operativa.

Dentro de las labores de trabajo realizadas por el autor de la investigación estuvo el sondeo y verificación de los ratios empleados por los investigadores relacionados con el tema de insolvencia así como los criterios de cuantificación y discriminación entre solventes e insolventes.

A partir de la identificación de los tipos de empresas a investigar y basándose en los informes anuales de dichas empresas, se determina algunos criterios para conformar cada grupo de empresas objeto de estudio. Para las empresas concursadas, el Sistema de Análisis de Balances Ibéricos (SABI) permitió una base de datos de 1795 empresas de las cuales tomaron las primeras 134 empresas con mayor volumen de actividad, y que hubiesen

reportado el estado de situación financiera y el estado de flujos de efectivo antes de la fecha en la que entro concursada; de estas 134 empresas finalmente se tomaron 102 empresas ya que las restantes 32 empresas no cumplieron con algunos parámetros necesarios para el análisis. Para las empresas cotizadas se tomó todas que pertenecían al mercado continuo español reportadas por la Bolsa de Madrid correspondientes al periodo 2008 – 2012 y que contaban con el balance de situación y el flujo de efectivo de periodos previos, como tampoco haber sido suspendidas durante el periodo objeto de estudio.

Con respecto a la selección de ratios el investigador realizo un profundo recorrido sobre la literatura que trata sobre el tema en cuestión, encontrando más de 120 publicaciones y un total de 368 ratios los cuales después de su análisis, selección 21 ratios que fueron usados con mayor frecuencia en estudios sobre insolvencia. Estos ratios se resumen en 5 grupos por su grado de interés y en función de la información (rotación, gastos financieros, rentabilidad, productividad, solvencia (Pessoa, 2016).

Posteriormente y con el propósito de clasificar las empresas solventes e insolventes a través del modelo *estático*, se seleccionan tres ratios significativos como se muestra en la Figura 41.

**Figura 41**

*Razones seleccionadas para el modelo estático Pessoa de Oliveira*

No Ratios	Componente del Ratio
R1	Razon de Solvencia a Corto Plazo $\frac{\text{Activos Corrientes}}{\text{Pasivo Corriente}}$
R2	Razon endeudamiento General $\frac{\text{Deudas Totales}}{\text{Pasivo Netos}}$
R3	Razon endeudamiento del Inmovilizado $\frac{\text{Deudas Largo Plazo}}{\text{Inmovilizado}}$

Adaptado de: “Análisis de solvencia de las empresas: modelo estático versus modelo dinámico O. Pessoa, 2016. <https://zaguan.unizar.es/record/48319/>

El primer ratio (R1) muestra la capacidad (solvencia) de pago que tiene la empresa en el corto plazo con criterio de cuantificación de solvencia  $R1 > 1$ .

El segundo ratio (R2) de endeudamiento general significativo en el análisis de la situación de la empresa con criterio de cuantificación de solvencia  $0 < R2 < 1$ .

El tercer ratio (R3) de endeudamiento del inmovilizado mide el grado de financiación con los recursos ajenos a L.P, con criterio de cuantificación de solvencia  $0 < R3 < 0,5$

Una vez aplicado el análisis de *insolvencia* a través del modelo *estático* en las empresas en concurso de acreedores se pudo determinar lo siguiente:

Pessoa, (2016) basado en cuantificaciones de sensibilidad en empresas objetivamente insolventes el ratio 2 catalogado como ratio de endeudamiento (R2), resulto ser el más preciso ya que de las 102 empresas analizadas pudo detectar el 95,1% como insolventes y el

4,9% como empresas solventes; demostrando de esta manera la capacidad de clasificar 97 empresas de las 102 objeto de estudio. Se puede detectar que este ratio muestra de alguna manera la proporción de las deudas con respecto al capital neto.

Con respecto al ratio 3(R3) definido como el ratio de endeudamiento del inmovilizado arrojó un 53,9% como insolventes equivalente a 55 empresas de las 102 analizadas y 46,1% como empresas solventes.

El ratio 1 (R1) denominado solvencia a corto plazo clasificó el 34,3% de empresas insolventes y un 63,7% como solventes, evidenciándose de esta manera que la asertividad de este ratio es menor en comparación con los ratios R1 y R3.

Se pudo detectar que de acuerdo,

*Al análisis de contingencia* existe una relación de independencia entre R1 y R2 en el momento de clasificar este tipo de empresas como solventes e insolventes, es decir el R1 no influye en R2 de manera significativa pues R1 trata del activo y pasivo a corto plazo y el R2 se relaciona las deudas a corto y largo plazo con respecto al capital neto (Pessoa, 2016, p. 205).

Respecto al *análisis de concordancia* R1 y R2 armonizaron en un 34,3%, es decir 35 de las 102 empresas analizadas fueron clasificadas tanto en R1 como en R2. El análisis de concordancia o armonización entre R1 y R3 corresponde a que 35 de las empresas clasificadas en R1 como insolventes y de las 55 en R3 solo 9 fueron coincidencia. El análisis entre R2 y R3, es mayor en comparación con las demás parejas de ratios coincidiendo en un 58,8%, es

decir de 60 de las 102 empresas, ya que, de las 55 empresas clasificadas como insolventes en R3 también fueron clasificadas en R2. (Pessoa, 2016).

En cuanto al análisis de correlación se evidencia cierto grado de dependencia entre R1 y R3, entre R2 y R3, sin embargo, no se ha encontrado una explicación financiera por lo que se deduce que hay independencia financiera entre los pares de ratios (R1 y R3/R2 y R3).

Pessoa (2016), basándose en el mismo grupo de empresas concursadas fundamento el modelo *dinámico* en el flujo de caja donde intervienen las actividades de operaciones, inversión y financiación. En este modelo se hace énfasis en tres momentos de gran connotación que merecen especial atención ya que se puede presentar posibles afectaciones por un lado están las referidas al comportamiento financiero, (obtención de recursos generados por las operaciones), la política financiera a corto plazo, y el proceso de inversión/financiación a largo plazo.

El primer indicador *dinámico* está representado por los recursos generados por operación (RGO), este se evidencia a través del estado de resultado una vez se haya calculado los impuestos y adicionando las amortizaciones. El segundo indicador tiene que ver con dos componentes: la tesorería generada por las operaciones (TGO), y de los recursos generados por operación (RGO). Es importante tener presente que la TGO es fundamental para la empresa porque, de ella dependerá la capacidad que tiene la empresa para hacer frente a sus obligaciones, el crecimiento y la distribución de dividendos (Pessoa, 2016).

La razón o indicador expresado como tesorería generada en operaciones (TGO) sobre los recursos generados por operación (RGO), mide de alguna manera la apropiada e inapropiada política financiera a corto plazo, medida por la capacidad de tesorería generada

por actividad. Según Pessoa es importante establecer algunos criterios de cuantificación para este indicador y, es que la empresa deberá tener por lo menos un de las siguientes condiciones:

- Es considerada solvente si la razón  $TGO/RGO$  es mayor que 0,95 siendo RGO mayor que 0, o
- Es considerada solvente si la razón  $TGO/RGO$  es menor que 1,05 siendo RGO menor que 0.

Otro aspecto relevante que tomo el autor para completar el análisis dinámico es el indicador de las actividades de financiación, inversión, dividendos y de la tesorería generada por las operaciones, expresado como:

$$\frac{F}{(I + Di - TGO)}$$

Donde  $F$  es la financiación,  $I$  la inversión,  $Di$  los dividendos y  $TGO$  es la tesorería generada por operación, este indicador responde al equilibrio en términos de la adecuada política financiera a largo plazo y el correcto uso de los recursos permanentes. Este indicador mide la eficacia de la política financiera y el correcto uso de recursos generados.

**Figura 42***Ratios Financieros modelo dinámico Pessoa de Oliveira*

<b>Ratio</b>	<b>Expresión</b>	<b>Criterio cuantificación de solvencia</b>
S1	RGO	$S1 > 0$
S2	$TGO/RGO$	$S2 > 0,95$ si $RGO > 0$ $S2 < 1,05$ si $RGO < 0$
S3	$\frac{F}{(I + Di - TGO)}$	$S3 \geq 1$

Adaptado de: "Análisis de solvencia de las empresas: modelo estático versus modelo dinámico O. Pessoa, 2016. <https://zaguan.unizar.es/record/48319/>

Para la aplicación de estas variables Pessoa acudió a medidas de valoración en períodos clave de un modelo de comportamiento financiero unificando la capacidad predictiva de los tres ratios (S1, S2, S3). Por consiguiente, el investigador identifica comportamientos distintos en tres momentos distintos del proceso de generación de solvencia, concluyendo que no se presenta reciprocidad estadística conjuntamente entre los ratios del modelo dinámico, ya que, cada uno de ellos procede de una manera diferente.

Como resultado del proceso aplicado con respecto al análisis descriptivo para las 102 empresas el S1 estuvo por debajo de 0 con un resultado de -1782,50 observando que un 87,3%, de empresas están correctamente clasificadas como insolventes.

Con respecto al S2 el resultado fue inferior a 0 con un valor de -0,3550, lo que indica que un 76,5% refleja un buen indicador de insolvencia.

El ratio S3 por el contrario arrojó un resultado de 1,0000, demostrando que, es el que menos diferencias muestra en las estadísticas descriptivas, con un 46,1% de clasificación de empresas como insolventes; con este indicador se detecta el buen manejo de la política financiera, lo que permite ser por el contrario una empresa solvente.

Como consecuencia del examen de reciprocidad de los ratios dinámicos en comparación con los ratios del modelo estático presentó consenso de empresas insolventes más sólidas, obteniendo relevancia en su resultado la combinación de los ratios S1 y S2 con un 71,6%, seguido de los ratios S1 y S3 con un 39,2% y S2 y S3 con un 35,3%.

Por otra parte, el investigador a partir de los resultados obtenidos de los modelos estáticos y dinámicos aplica técnicas de *inteligencia artificial* y para ello toma muestras de 130 empresas industriales españolas clasificadas en dos grupos de 65 empresas tanto solventes como insolventes a las cuales se le aplica el modelo Perceptrón Multicapa (MLP) catalogada como una herramienta no paramétrica la cual es usada para realizar trabajos de clasificación y predicción.<sup>8</sup>

Los resultados de categorización del MLP para un modelo de red neuronal de carácter global, fueron buenos, porque en la muestra sometida a prueba los resultados ascendieron al 82,6%, como predictores a nivel general. Los seis ratios de los modelos *estático* y *dinámico* alcanzaron a clasificar correctamente el 81,8% de las empresas en insolventes y el 83,3% de las empresas en solventes, detectando que los ratios S1 y S2 reflejaron un 100% de importancia significativa y un 69%, respectivamente para clasificar satisfactoriamente empresas solventes e insolventes (Pessoa, 2016).

### 3.2.9. Machine Learning - Aprendizaje Automático

El aprendizaje automático se puede definir como “el estudio de algoritmos que permiten a las máquinas mejorar su rendimiento en alguna tarea determinada conforme se le van suministrando nuevos datos”, (Vinet & Zhedanov, 2011, p. 217). Los desarrolladores no tendrán que invertir mucho tiempo en programar todos los posibles escenarios además de todas las excepciones posibles, solo deberán alimentar el algoritmo con un gran volumen de información (datos) para que el algoritmo memorice (aprenda) y reconozca qué hacer en cada uno de los posibles casos.

Alan Turing (1950) fue el hombre que abrió camino a la inteligencia artificial en su artículo denominado *Computing machinery and intelligence*, indico que la inteligencia artificial es una campo científico y tecnológico, cuyo objetivo es desarrollar procesos de la mente y su relación con el cuerpo a través de una representación computacional, sostiene que el cerebro posee un sistema de procesos de datos, similar al de una computadora digital (Matich, 2001, p. 6).

Las etapas del ciclo de vida del machine learning son la adquisición de colección de datos, preparación limpieza y calidad de datos, procesar y ejecutar herramientas de machine learning como algoritmos y reportar los resultados.

Según Liliana Sandoval (2018), en su artículo denominado *machine learning algorithms for data analysis and prediccion*, cuando estos dispositivos se entrenan, la habilidad de aprender y entender, se convierten en capacidades iguales a las del hombre ya que, tienen más desarrollo en velocidad y capacidad que cualquier humano puede tener. Los tipos de implementación de machine learning pueden clasificarse: en aprendizaje no

supervisado y supervisado, aunque existen más categorías de aprendizaje (Sandoval, 2018), son muy usados en la minería de datos.

### *Aprendizaje no supervisado*

En el aprendizaje no supervisado, a la máquina (software informático) se le suministra un conjunto de datos y no se le proporciona ninguna respuesta correcta. Dada la enorme cantidad de datos, la máquina puede identificar patrones y tendencias de similitud. Por medio de un algoritmo se identifican agrupaciones o grupos de elementos similares o semejanzas de un elemento nuevo con un grupo existente, a las agrupaciones identificadas se les conoce como clústeres, dichas agrupaciones se pueden categorizar o etiquetar (Rebala et al., 2019).

A continuación, un ejemplo del aprendizaje no supervisado:

Una máquina registradora en una tienda de comestibles recibe los datos sobre la compra de todos los clientes. En un período de tiempo, en la máquina se pueden identificar patrones de adquisición, como los clientes que compran mochilas escolares también compran lápices de colores o los que compran cerveza también compran papas fritas. La máquina aún no sabe qué son estos objetos en concreto. Sin embargo, podrá identificar la correlación entre ellos (Rebala et al., 2019, p.21).

De forma análoga se pueden identificar patrones o similitudes en los indicadores financieros de una muestra de empresas pertenecientes a un mismo sector económico de un país en un periodo determinado, es importante recordar que los indicadores financieros de una empresa no sirven de mucho si no se compara con los indicadores financieros de empresas de

un mismo sector o industria, por lo que es conveniente involucrar técnicas y/o algoritmos que permitan facilitar, agilizar y hacer más eficiente el proceso de análisis.

*Algoritmo K-Means – Aprendizaje No Supervisado*

El algoritmo k-means corresponde propiamente al modelo de aprendizaje no supervisado, consiste en agrupar un conjunto de objetos (datos) de entrada en  $K$  grupos o clústeres previamente definidos y cada objeto del conjunto no está etiquetado (categorizado). Sin embargo, el algoritmo agrupa los objetos de acuerdo con su cercanía o noción de distancia, identificando patrones y tendencias. Su desarrollo consiste en establecer  $K$  puntos, denominados centroides (centros) de los grupos (clústeres), para luego identificar cuáles son los objetos más cercanos y asignar ese objeto al clúster correspondiente a cada centroide.

El algoritmo trabaja encontrando las distancias geométricas<sup>1</sup> de los objetos a los centroides, por lo general, empleando la distancia euclídea. El objetivo del algoritmo es minimizar la distancia de cada objeto al centroide de su clúster. Es decir, establecido un centroide  $v_K$  y un objeto  $x_i$  en el clúster  $K$ , se busca minimizar la distancia dada por,

$$Distancia = || x_i - v_k ||^2.$$

Para minimizar dicha distancia, se deben tener en cuenta todos los objetos a considerar y, por lo tanto, se lleva a cabo mediante un proceso iterativo donde la posición del centroide se

---

<sup>1</sup> Desde una perspectiva lineal, es el espacio entre un punto y otro, la cual satisface los axiomas de Euclides en la geometría.

actualiza para cada clúster tomando como nuevo centroide la posición del promedio de los objetos pertenecientes a dicho clúster.

Para el desarrollo del algoritmo se hace necesario un conjunto de iteraciones de tal forma que ningún objeto cambie de clúster, en donde cada clúster  $K$  está representado por sus centroides. El pseudocódigo del algoritmo se presenta a continuación.

Dado un conjunto de  $i$  objetos,

Inicializar aleatoriamente  $[v_1, v_2, \dots, v_k]$   $K$  puntos centroides que representan cada cluster

Iterar hasta que ningún objeto cambie de cluster

{

Para cada objeto (de 1 hasta  $i$ )

Asignarlo al cluster cuyo centroide está más cerca al objeto

Fin para

Para cada cluster (1 a  $K$ )

Buscar la media de todos los objetos asociados con este cluster

Actualizar el centroide del cluster ( $v_k$ ) a esta nueva media encontrada.

Fin Para

}

El algoritmo realiza lo que se denomina una partición dura que cumple dos condiciones. La primera corresponde a que cada objeto tiene que pertenecer a alguno de los clústeres del conjunto de objetos y la segunda, que cada objeto pertenece exclusivamente a un cluster de la partición. En este sentido, la cantidad de clústeres debe ser especificada para inicializar el algoritmo. A continuación, se presenta la definición de una partición dura (Rojas et al., 2009).

**Definición 1.** Sea  $X$  un conjunto de  $m$  objetos y  $x_i$  un elemento que pertenece a  $X$ , se dice que una partición  $P = \{C_1, C_2, \dots, C_c\}$  donde  $c$  es un número entero mayor a 1 que indica la cantidad de clústeres, es una partición dura de  $X$  si y solo si:

1.  $\forall x_i \in X \exists C_k \in P$  tal que  $x_i \in C_k$
2.  $\forall x_i \in X \quad x_i \in C_k \rightarrow x_i \notin C_j$  donde  $k \neq j, C_k, C_j \in P$ .

Para lograr hallar el centroide de cada cluster se establece la función objetivo correspondiente y que permite determinar dicho centroide. La función habitualmente empleada, considerando la distancia euclídea corresponde con la suma de la distancia entre los objetos de cada cluster y su centroide de tal manera que la distancia de cada punto con respecto al centroide del clúster sea mínima. Por lo tanto, se tiene la función objetivo  $J$ , que determina el problema de optimización dado a continuación.

$$\min J(P, V) = \sum_{k=1}^n \sum_{i=1}^m \|x_i - v_k\|^2$$

Donde  $V$  es el vector de los centroides de cada cluster a ser identificados y  $v_k$  un centroide. Este criterio es útil porque dado un conjunto de centroides de los clústeres adecuado

o correcto, permitirá siempre encontrar un valor mínimo de la función  $J$  (Yen & Langari, 1999). Uno de los problemas que presenta este algoritmo está relacionado con el establecimiento de mínimos locales; con el fin de evitar esto, se hace necesario probar distintos puntos de partida y encontrar los clústeres. Este problema solo se presenta si el número de clústeres es pequeño. También puede darse la situación de que alguno de los grupos podría no tener algún punto asociado, es así como se puede dar lugar al clúster vacío.

#### *Algoritmo Fuzzy C-Means – Aprendizaje No Supervisado*

Fuzzy c-means (FCM) también es un algoritmo de agrupamiento de aprendizaje no supervisado en donde cada objeto tiene un grado de pertenencia difusa (fuzzy) a los clústeres. La finalidad del algoritmo FCM es similar al algoritmo k-means en cuanto a agrupar objetos de un conjunto en clústeres.

Se suele emplear cuando un objeto está muy cerca de dos clústeres de tal forma que es difícil agruparlo en uno o en otro, esto se debe a la relativa frecuencia con la cuál un objeto específico presenta características pertenecientes a distintos clústeres y como consecuencia no es fácilmente clasificado. El algoritmo FCM asigna a cada objeto un valor de pertenencia dentro de cada clúster y por consiguiente un objeto específico puede pertenecer o presentar características similares de forma parcial en más de un clúster (Rojas et al., 2009, p. 2).

Por lo tanto, k-means y FCM difieren en que el primero trabaja con una partición dura mientras que el segundo realiza una partición suave o borrosa al conjunto de objetos, en donde

los objetos pertenecen en algún grado a todos los clústeres (Rojas et al., 2009). A continuación, se presenta la formulación de una partición suave:

**Definición 2.** Sea  $X$  conjunto de objetos y  $x_i$  un elemento perteneciente a  $X$ , se dice que una partición  $P = \{C_1, C_2, \dots, C_c\}$  donde  $c$  es un número entero mayor a 1 que indica la cantidad de clústeres, es una partición suave de  $X$  si y solo si:

- $\forall x_i \in X \quad \forall C_k \in P \quad 0 \leq \mu_{C_k}(x_i) \leq 1$
- $\forall x_i \in X \quad \exists C_k \in P \quad \text{tal que} \quad \mu_{C_k}(x_i) > 0$

Donde  $\mu_{C_k}(x_i)$  indica el grado de pertenencia de  $x_i$  en los cluster  $C_k$  (Yen & Langari, 1999). Una partición suave es restringida si la sumatoria de los grados de pertenencia de un objeto específico en los  $c$  clústeres es igual a 1.

$$\sum_k \mu_{C_k}(x_i) = 1 \quad \forall x_i \in X$$

La función objetivo  $J$  de la partición dura en el algoritmo k-means se extiende en dos fases dando lugar a la partición suave restringida. En la primera fase, a la función objetivo  $J$  se le agrega los grados de pertenencia difusos de cada objeto en cada clúster. En la segunda fase, se agrega un parámetro adicional  $b$  como exponente que sirve de peso en la función de pertenencia. Por lo tanto, la función objetivo extendida  $J$ , determina el siguiente problema de optimización.

$$\min J(P, V) = \sum_{k=1}^n \sum_{i=1}^m (\mu_{C_k}(x_i))^b \|x_i - v_k\|^2$$

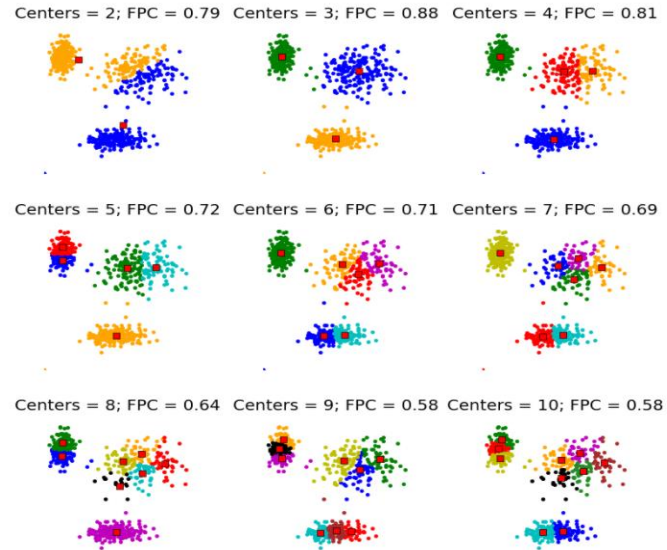
Donde  $P$  es una partición difusa del conjunto de objetos  $X$  formada por  $\{C_1, C_2, \dots, C_k\}$  el parámetro  $b$  es un peso que define el grado en el cual los objetos

parciales de un clúster afectan el resultado. Los algoritmos FCM y k-means intentan hallar la mejor partición mediante la búsqueda de los prototipos  $v_k$  que minimizan la función objetivo  $J$ , el algoritmo también debe hallar las funciones de pertenencia  $\mu_{C_k}$  que minimicen a  $J$  (Yen & Langari, 1999).

Existe tres categorías o medidas de la validez de una partición suave: las medidas basadas en el grado de pertenencia, las medidas basadas en la geometría y las medidas basadas en el desempeño. El objetivo de dichas medidas es establecer qué tan buena es la partición generada por el algoritmo (Rojas et al., 2009). Una de las medidas basadas en el grado de pertenencia es el coeficiente de partición (PC) desarrollado por Jim Bezdek en 1973 y el cual se expresa a continuación:

$$PC = \frac{1}{m} \sum_{k=1}^n \sum_{i=1}^m \mu_{C_k}(x_i)$$

El índice PC mide el grado de separación o solapamiento (grado de fuzziness) entre los  $c$  grupos formados, un PC igual a 1 indica que la partición es dura y, para el caso de la partición difusa el coeficiente de partición será igual a  $\frac{1}{c}$  por tanto, una partición difusa será mejor, si el valor de dicho coeficiente se aproxima a 1. Es decir, a partir del PC se puede establecer la cantidad de cluster  $c$ . A continuación, se presenta un ejemplo de una muestra de datos, en donde se observan los coeficientes de partición fuzzy PC para diferentes cantidades de cluster (de 2 a 10).

**Figura 43***Cluster Fuzzy C-Means*

*Nota.* Ejemplo del FPC para diferentes cluster fuzzy c-means Adaptado de Cluster Fuzzy C-Means por Python, s.f. [https://pythonhosted.org/scikit-fuzzy/auto\\_examples/plot\\_cmeans.html](https://pythonhosted.org/scikit-fuzzy/auto_examples/plot_cmeans.html)

Se puede observar los centroides como puntos de color rojo y los datos pertenecientes a cada clúster alrededor de cada centroide representados por diferentes colores para cada partición. De acuerdo al ejemplo el PC que más se acerca a uno y es el mejor es 0.88, el cual corresponde a una partición con tres clústeres. Existen otros índices que de la misma manera son muy usados para establecer que tan buenas son las particiones como:

El coeficiente de entropía (CE), presentado también por Jim Bezdek, es una medida escalar de la cantidad de borrosidad en los  $k$  centroides, entre más bajo sea el valor, la partición obtenida será mejor, la formulación de CE está dado por:

$$CE = -\frac{1}{m} \sum_{k=1}^n \sum_{i=1}^m \mu_{C_k}(x_i) \log_a \mu_{C_k}(x_i)$$

El coeficiente de partición modificado (MPC) propuesto por R.N. Dave en 1996, busca reducir la tendencia evolutiva monótona que presentan el PC y CE con respecto al número de clústeres  $c$ . El índice puede variar de 0 a 1 y al igual que el PC, cuanto más se aproxime el valor obtenido a 1 la partición obtenida será mejor. La formulación de MPC está dada por la siguiente transformación lineal de PC:

$$MCP = 1 - \frac{n}{n-1} (1 - PC)$$

El coeficiente de Xie–Beni (XB), planteado por X. Xie y G. Beni en 1991, se centra en la medición de dos propiedades, compacidad y separación. El numerador de la siguiente formulación mide la compacidad de la partición borrosa y el denominador la fuerza de separación de los clústeres,

$$XB = \frac{\sum_{k=1}^n \sum_{i=1}^m (\mu_{C_k}(x_i))^b \|x_i - v_k\|^2}{n \min_{k,i} \|v_k - v_i\|^2}$$

El coeficiente Kwon, propuesto por H. Kwon en 1991, busca eliminar la tendencia a disminuir monótonamente cuando el número de agrupaciones se acerca al número de puntos de objetos del índice de Xie y Beni. Para lograr esto, se introdujo una función de penalización en el numerador del índice XB por lo tanto, el índice Kwon se define como

$$Kwon = \frac{\sum_{k=1}^n \sum_{i=1}^m (\mu_{C_k}(x_i))^b \|x_i - v_k\|^2 + \frac{1}{c} \sum_{k=1}^n \|v_i - v_k\|^2}{n \min_{k \neq i} \|v_k - v_i\|^2}$$

### ***El aprendizaje supervisado***

Se habla de aprendizaje supervisado cuando se entrena un algoritmo de machine learning dándole las preguntas y las respuestas. En este tipo de aprendizaje hay dos algoritmos: el de clasificación que como su nombre lo indica es cuando el algoritmo encuentra patrones en los datos que le asignan y los clasifica en grupos; y el de *regresión*, que a diferencia del de clasificación no lo ubica en un grupo, sino que devuelve un número o valor en específico. Bajo dicho aprendizaje, la máquina recibe un conjunto de datos que conforman el algoritmo de aprendizaje, es importante destacar que en este caso los datos están etiquetados o categorizados, es decir, que cuentan con su respectiva respuesta que además es correcta; una vez establecidos los datos, la función del algoritmo es identificar aquellas características del conjunto de datos para así determinar la respuesta correcta. Así pues, si llegase a haber un nuevo conjunto de datos el algoritmo debería poder arrojar un resultado provisional a través de una predicción basada en las características o etiquetas previamente identificadas (Rebala et al., 2019).

### ***Algoritmo K-Nearest Neighbor (KNN) – Aprendizaje Supervisado***

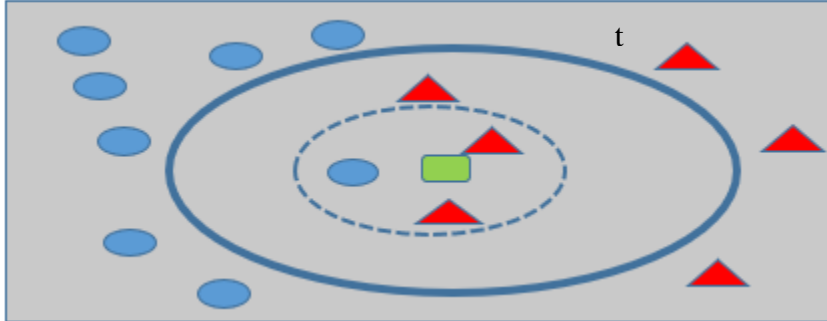
K-Nearest Neighbor (KNN) hace referencia a un algoritmo diseñado bajo el modelo de aprendizaje supervisado y es usado tanto para problemas de clasificación como problemas de regresión. En ambos casos, la esencia del algoritmo consiste en, dado un conjunto de datos previamente etiquetados o clasificados en varias clases (dos o más), se busca clasificar un nuevo dato considerando los  $k$  elementos que se encuentran *más cercanos* al dato a clasificar.

(Martin & Idoate, 2015). Por lo tanto, la noción de distancia es relevante para el desarrollo del algoritmo. En este sentido, se suele usar la distancia euclídea, sin que sea la única alternativa.

Por lo tanto, para localizar los puntos similares más cercanos, como entrada del algoritmo, se debe especificar la métrica que permita medir la proximidad o distancia entre los puntos utilizando medidas de distancias o nociones de disimilitud. Como se mencionó anteriormente, la opción más popular corresponde con la distancia *euclídea*, aunque hay otro tipo de nociones de distancia como la *manhattan*, la distancia *minkowski*, *Cosine Similarity* entre otras.

A diferencia de otros algoritmos que utilizan el conjunto de entrenamiento para aprender sobre algunos parámetros o variables y con base en estos, se clasifica el nuevo dato, el algoritmo KNN clasifica un nuevo dato empleando el conjunto de entrenamiento sin necesidad de aprender ningún parámetro. (Martin & Idoate, 2015)

En esencia, KNN calcula la distancia entre el nuevo dato a clasificar y todos los elementos del conjunto y, selecciona los  $k$  elementos con menor distancia. Una vez se han escogidos tales elementos, determina a que clase pertenecen la mayoría de ellos para establecer la clase o categoría del nuevo dato. Por ejemplo, la Figura 44, presenta la situación en la cual dadas dos clases, círculos y triángulos, se busca clasificar un nuevo objeto (cuadrado) estableciendo los vecinos más cercanos. Considerando una elección de  $k=3$ , la cual se representa con la línea punteada, el nuevo objeto por mayoría, es clasificado en la clase de los triángulos. En una segunda situación, con una elección de  $k= 3$ , nuevamente el objeto queda clasificado en la clase de los triángulos.

**Figura 44***KNN - Vecinos más cercanos**Elaboración propia.*

Claramente, el valor de  $k$  es el que establece cuantos vecinos selecciona el algoritmo para predecir la clase a la cual pertenecerá el nuevo dato. En general, el algoritmo funciona bastante bien con un conjunto amplio de variables o atributos para cada dato, sin embargo, se hace necesario escalar las variables para que cada variable tenga el mismo peso o importancia dentro de la determinación de la distancia. A continuación, se presenta el pseudocódigo para el clasificador KNN básico.

Inicio

*Entrada:*  $D = \{(X_1, C_1), \dots, (X_N, C_N)\}$

$X = (x_1, \dots, x_n)$  nuevo caso a clasificar

Para todo objeto ya clasificado  $(x_i, c_i)$

Calcular  $d_i = d(X_i, X)$

Ordenar  $d_i (i = 1, \dots, N)$  en orden ascendente

Seleccionar con los  $K$  casos  $D_X^K$ , ya clasificados más cercanos a  $X$ .

Asignar a  $X$  la clase más frecuente en  $D_X^K$ .

Fin

Según Mendialdua (2015), el anterior pseudocódigo para el clasificador K-NN,  $D_X^K$  representa los  $k$  datos ya clasificados más cercanos a  $X$ .  $C_N$  Representa la clase a la cual pertenece el objeto  $X_N$  con  $n$  atributos o variables. El algoritmo recibe el conjunto de entrenamiento  $D$  y por cada nuevo objeto  $X$  se calcula la distancia entre ese objeto y los elementos del conjunto de entrenamiento

Una de los aspectos que se debe tener en cuenta es, si el valor de  $k$  es muy alto, dentro de los  $k$  más cercanos se puede incluir datos que son de otra clase. Lo anterior se puede observar en la Figura 45. Por lo tanto, lo más conveniente es establecer un umbral bajo el cual dicha clasificación no se altera o cambia. De igual forma, en el caso bajo el cual la cantidad de vecinos más cercanos sea la misma para ambas clases, se debe establecer algún tipo de regla para el desempate, por ejemplo, seleccionar la del primer vecino más cercano. (Martin & Idoate, 2015).

**Figura 45**

*Distintos vecindarios para diferentes valores de k*



*Nota:* Aprendizaje de distancias basadas en disimilitudes para el algoritmo de clasificación KNN, Adaptado de: “Aprendizaje de distancias basadas en disimilitudes para el algoritmo de clasificación kNN”. M. Martín & M. Idoate, 2015. <http://academicae.unavarra.es/handle/2454/18430?show=full>

En los algoritmos de aprendizaje supervisado, existen dos fases fundamentales en el proceso, la fase de entrenamiento y la fase de prueba.

#### *Fase de entrenamiento*

En esta fase se tiene una cantidad enorme de datos, de la cual se separa una parte para entrenar el aprendizaje del algoritmo al darle toda esa información para que encuentre los patrones necesarios y después pueda hacer predicciones, puede ser el 70% o 80% del total de datos (Sandoval, 2018, p. 39).

#### *Fase de prueba*

En la fase de prueba el resto de los datos que quedan, son usados para hacer las pruebas. Así se puede revisar el algoritmo, evaluar si los resultados están bien o mal, y saber si está aprendiendo o no. Si se observa que no coinciden los datos resultantes, se tendrá que agregar más datos o cambiar el método utilizado. En

caso contrario, si se observa que hay entre un 80% a 90% resultados correctos, se podrá decir que hay un buen grado de aprendizaje y el algoritmo es útil para su uso (Sandoval, 2018, p. 39).

#### **4 Metodología**

En esta sección, se busca presentar el método propuesto. Para este fin, se estable la muestra empleada para el estudio y aplicación, el tipo de investigación empleada y, se describen cada una de las etapas del diseño metodológico.

La información base sobre la cual se seleccionó la muestra para realizar el presente trabajo, se descargó de la página web de la Superintendencia de Sociedades (Insolvencia Empresarial en Colombia 2016, 2017 y 2018). Módulo de insolvencia financiera, sociedades en reorganización y validación judicial a 31 de diciembre de 2018. A través de esta página se reporta el listado de empresas que a 31 de diciembre de 2018 se acogieron a la ley 1116 de insolvencia o de reorganización financiera en Colombia, bajo esta ley se clasifican las empresas en reorganización, reestructuración, o liquidación obligatoria o judicial.

Los sectores empresariales tienen una clasificación económica asignada por la Cámara de Comercio de Colombia identificado como el código CIU (Clasificación Industrial Internacional Uniforme). Cada categoría de empresas se clasifica de acuerdo con el grupo al cual pertenecen según las Normas Internacionales de Información Financiera (NIIF); es así como se toman empresas pertenecientes al grupo 1 NIIF-Plenas individuales, específicamente del sector de la construcción.

La finalidad de este trabajo es proponer un método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada aplicado en el sector Económico de la Construcción en Colombia tomando información de los estados financieros de los años 2017 y 2018. Como referencia, basado en algoritmos de machine learning de aprendizaje no supervisado y supervisado respectivamente, para este método la variable dependiente corresponde con la situación de insolvencia empresarial, la información de los estados financieros, y como parte fundamenta la revisión bibliográfica referente a los modelos de predicción de fallas o fracaso empresarial que permitió la selección de los ratios más usados como variable independiente en aras de asegurar la efectividad en la aplicación del método.

#### **4.1 Tipo de Investigación**

La presente investigación se desarrolla bajo un enfoque cuantitativo al usar variables numéricas (datos y muestras) en este caso en particular información contable de los estados financieros de las empresas del sector de la construcción en colombiana reportados a la Superintendencia de Sociedades para el año 2017 además de la base de datos referente a las Empresas en proceso de Insolvencia Financiera para el 2018, en donde se busca realizar inferencias sobre un conjunto de variables numéricas. Así mismo el alcance de la investigación es propositivo por cuanto se busca diseñar un modelo con su correspondiente validación.

## 4.2 Diseño de la Metodología

El diseño de la metodología consta de cinco etapas. En la primera etapa se escogió la población y muestra de análisis, es decir el sector económico empresarial sujeto de estudio. En dicha etapa también se seleccionó el uso de los algoritmos de machine learning ya que por el volumen de información (datos) presenta ciertas ventajas que facilitan y agilizan el procesamiento de información eliminando el sesgo humano lo que hizo más eficiente el proceso de análisis.

Para la segunda etapa se seleccionaron los ratios usados dentro del método mediante la revisión bibliográfica del tema de estudio, posteriormente se calcularon para cada empresa perteneciente a la muestra dichos ratios, siendo estos la información de entrada y base de datos para el método.

En la tercera etapa se determinó el número óptimo de cluster para la base de datos por medio de la aplicación del algoritmo fuzzy c-means sobre el conjunto de ratios, se emplearon diferentes valores de  $c$  y a partir de ciertos índices como el coeficiente de partición fuzzy se definió el mejor  $c$  para el conjunto de datos particular. Con base en esa clasificación, se etiquetaron las empresas, es decir, el algoritmo agrupa las empresas en los clústeres según la similitud de los ratios financieros entre las empresas, posteriormente se etiqueta y/o clasifica la información de acuerdo al clúster al que pertenece. Por lo tanto, cada empresa queda agrupada en una de las dos categorías de, solvente o insolvente respectivamente para un número de clústeres óptimo igual a dos. En caso de haber sido el número de clústeres mayor a dos se hubiera definido categorías intermedias entre solvente e insolvente.

Para la cuarta etapa empleando las etiquetas o clústeres definidos a través del fuzzy c-means, se procede a aplicar el KNN, el algoritmo divide la base de datos categorizada en dos subgrupos uno para entrenamiento y el otro para la prueba basándose en las etiquetas del fuzzy c-means de la sección tres. Finalmente se introducen al método los ratios de nuevas empresas del mismo sector para su categorización como solvente o insolvente.

En la quinta etapa se validan los resultados de las posibles clasificaciones solvente o insolvente del modelo con la situación real de la empresa, vale la pena aclarar que se trabajó con información del año 2018, por medio del procedimiento metodológico se infirió el estado solvente o insolvente de las empresas para luego compararlos con la situación real, es decir, si para ese año se presentó o no problemas de insolvencia registrados por la superintendencia de sociedades. En la Figura 46 se resumen las anteriores etapas del método y en la Figura 47 se presenta un diagrama de flujo que expresa el método y en general, el desarrollo de las etapas.

Figura 46

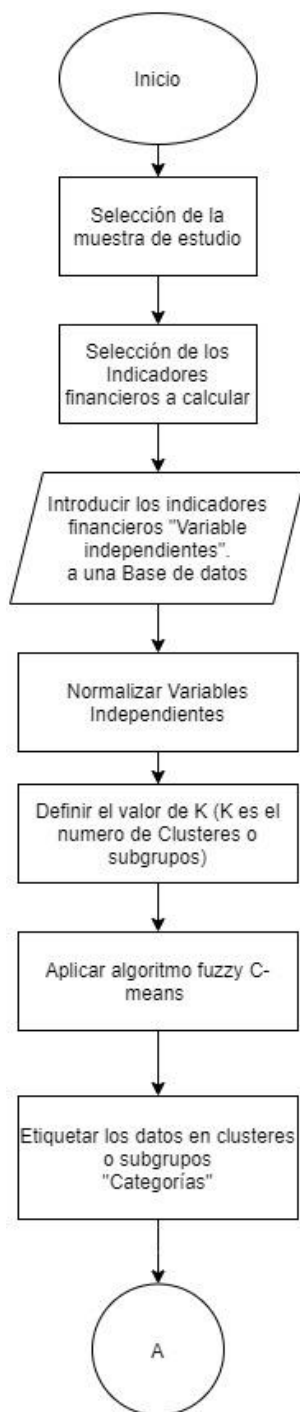
*Diseño Metodológico del Modelo de Clasificación de Insolvencia*

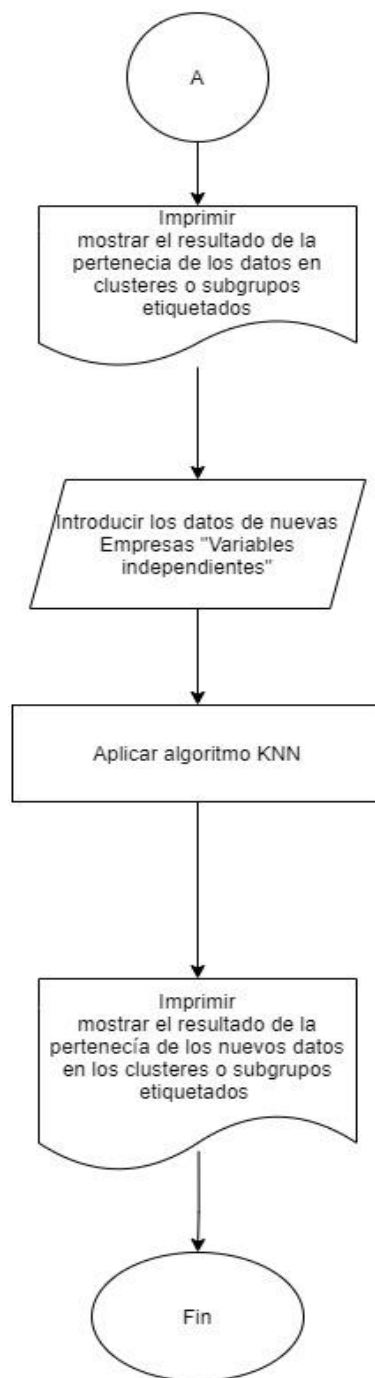
<b>ETAPA 1: Diseño de la Muestra</b>			
Selección de un Sector Económico específico y muestra de estudio.			
<b>ETAPA 2: Selección y toma de datos</b>	<b>ETAPA 3: Machine Learning</b>	<b>ETAPA 4: Parte 1 Machine Learning</b>	<b>ETAPA 4: Parte 2 Machine Learning</b>
Ratios Financieros Relevantes para utilizar en el modelo.	Aprendizaje no Supervisado uso del Algoritmo Fuzzy C- Means.	Aprendizaje Supervisado uso del Algoritmo K- Nearest Neighbor (KNN) o K-vecinos más cercanos.	Aprendizaje Supervisado uso del Algoritmo K-Nearest Neighbor (KNN) o K- vecinos más cercanos.
Se seleccionan y calculan las variables independientes para cada empresa.	Se hace particiones y calcula el número óptimo de clústeres para la clasificación y etiquetado usando el algoritmo Fuzzy C-Means	Se divide el conjunto de datos en dos clústeres, uno para entrenamiento del aprendizaje del algoritmo de clasificación Knn y el otro para poner a prueba la clasificación del algoritmo.	Se prueba la capacidad clasificatoria del algoritmo Knn con información de nuevas empresas del mismo sector.
<b>ETAPA 5: Análisis de Resultados</b>			
Método para la Determinación de Insolvencia Financiera			

Elaboración Propia

**Figura 47**

*Diagrama de flujo de la propuesta procedimiento clasificación de la insolvencia financiera*





*Nota.* Diagrama de flujo de la propuesta procedimiento clasificación de la insolvencia financiera, elaboración propia.

### Etapa 1: Diseño de la Muestra

Para la investigación se consideró el sector de la Construcción como población de análisis, ya que su crecimiento en términos del PIB para el año 2018 fue uno de los más bajos 0,30% según el DANE - Departamento Administrativo Nacional de Estadística, quedando por encima del sector de explotación de minas y canteras con -0,80% y por debajo de las Actividades Artísticas, de entretenimiento y recreación con 1,40%. Se puede considerar que el desempeño financiero en términos del PIB para el sector de la construcción fue estable o no tuvo mayor incidencia en la economía para el periodo comprendido en el año 2018.

**Figura 48**

*Crecimiento de la economía total en Colombia por sectores económicos para el año 2018*



*Nota.* Crecimiento de la economía total en Colombia por sectores económicos para el año 2018, adaptado de “La economía en 2018” por Departamento Administrativo Nacional de Estadística [DANE], 2018. <https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-trimestrales/historicos-producto-interno-bruto-pib>

El sector de la construcción se subdivide en los subsectores de la construcción de edificios y otros, acabado de edificios e instalaciones eléctricas y obras de ingeniería Civil. Según el Registro Único Empresarial y Social - RUES para el año 2018 en el sector de la construcción hubo 13525 empresas inscritas y para el año 2019 hubo 13751 con una variación o crecimiento del 1.7%. En la Figura 49 se resume la información agrupada por sectores económicos.

**Figura 49**

*Cantidad de Empresas por Sector Económico con Registro Mercantil*

Ene-Dic 2019/18				
Sector	Ene-Dic 2018	Ene-Dic 2019	Variación %	Contribución
Servicios	133443	137943	3,4%	1,5
Comercio	113751	115415	1,5%	0,5
Industria	28827	29414	2,0%	0,2
Construcción	13525	13751	1,7%	0,1
Extracción	1168	1091	-6,6%	0,0
Agricultura	4865	4790	-1,5%	0,0
Resto	7448	7059	-5,2%	-0,1
<b>Total General</b>	<b>303027</b>	<b>309463</b>	<b>2,1%</b>	<b>2,1</b>

Adaptado de: “Empresas por Sector Económico con Registro Mercantil”, por: Registro Único Empresarial y Social [RUES], 2018. <http://www.rues.org.co/>

Una vez seleccionado el sector económico de la construcción como población de análisis, se seleccionó la muestra de la investigación, para este caso se tuvo en cuenta solo los estados financieros de 197 grandes empresas que reportaron información bajo NIIF plenas individuales requeridas y publicadas por la Superintendencia de Sociedades para el año 2018

de acuerdo con la base de datos, *Revisar Anexo 1*. No se consideró la información de las grandes empresas que presentaron sus estados financieros bajo NIIF plenas en operaciones conjuntas ya que, se consolida la información perteneciente a diferentes grupos empresariales lo que puede afectar la realidad de económica y financiera de las mismas.

La información financiera fue tomada de la base de datos del sistema de información y reporte empresarial de la superintendencia de sociedades por sus siglas SIREM, el cual se encuentra en la página web de la Superintendencia de Sociedades. La base de datos del SIREM contiene la información financiera que reportan todas las personas naturales y jurídicas vigiladas o inspeccionadas por la Superintendencia de Sociedades con corte a 31 de diciembre de cada año (Superintendencia de Sociedades, 2019, párr. 1).

Se obtuvo la información de tres estados financieros: Estado de situación Financiera (E.S.F), Resultados Integrales - ORI y Estado de Cambios en el Flujos de Efectivo (E.C.F.E) reportados por las empresas para el año 2018, la muestra se compuso inicialmente de un total de 197 empresas, pero solo 163 empresas reflejaron la información completa para el periodo de análisis, la Figura 50 resume la anterior información:

**Figura 50***Muestra Preliminar*

Estado de la Información	Cantidad de Empresas
Completo	163
Incompleto	25
No Disponible	9
Total	197

Elaboración Propia

Al depurar la información de las 163 empresas, la muestra posteriormente se compuso de 111 empresas debido a que no todas presentaron la información financiera completa en sus estados financieros, es decir, no existía información sobre alguna cuenta. Para la muestra también se consideró la base de datos de las empresas en procesos de insolvencia empresarial publicada también por la Superintendencia de Sociedades en su página web. Se consideró específicamente la información relacionada al sector de la Construcción, se encontraron 69 registros de empresas que se sometieron a régimen de Insolvencia desde el año 2014 hasta el año 2019, pero solo 32 empresas cumplieron con los tres estados financieros requeridos para el periodo de análisis entre los años 2017 al 2018, la Figura 51 resume la anterior información:

**Figura 51***Muestra de Empresas en proceso de Insolvencia*

Estado de la Información	Cantidad de Empresas
Completa	32
Incompleta	20
No Disponible	17
Total	69

Elaboración Propia

Al depurar la información de las 32 empresas con problemas de insolvencia, solo 12 empresas entraron a régimen de insolvencia en el año 2018, *Revisar Anexo 2*. Es así como la muestra final se compone de la muestra compuesta por las 111 empresas sin problemas de insolvencia más las 12 que entraron a procesos de insolvencia en el año 2018 para un total de 123 empresas pertenecientes al Sector de la Construcción como muestra para el método. En la Figura 52 se resume la información relacionada.

**Figura 52**

*Muestra Final para el Modelo de Clasificación de Insolvencia*

Estado de la Empresa en términos de Solvencia	Cantidad de Empresas
Sin problemas	111
Con problemas	12
Total Muestra	123

Elaboración Propia

### **Etapa 2: Selección y toma de datos**

Para la selección de las variables independientes se tuvieron en cuenta 2 criterios: el primer criterio se relaciona con la revisión bibliográfica realizada en el marco teórico de la presente investigación, específicamente los ratios más populares usados en el modelamiento de la predicción y/o clasificación del fracaso empresarial, el segundo criterio tiene en cuenta los ratios recomendados y con los mejores resultados pertenecientes a la revisión bibliográfica, inicialmente se encontraron 70 ratios *revisar Anexo 3*, pero se seleccionaron y calcularon solo 14, la Figura 53 resume los 14 ratios que se incluyeron dentro del análisis propuesto.

Figura 53

*Lista de Ratios Seleccionados para el Modelo de Clasificación de Insolvencia*

<b>LIQUIDEZ (Efectivo/Liquidez)</b>		
<b>Estado de Cambios en la Situación Financiera (E.C.S.F)</b>	<b>R1</b>	$\frac{\text{Flujo de Efectivo}}{\text{Pasivos Totales}}$
	<b>R2</b>	$\frac{\text{Flujo de Efectivo}}{\text{Ventas}}$
	<b>R3</b>	$\frac{\text{Flujo de Efectivo}}{\text{Patrimonio Neto}}$
<b>RENTABILIDAD</b>		
<b>Estado de Resultados- ORI</b>	<b>R4</b>	$\frac{\text{Utilidad Neta}}{\text{Ventas}}$
	<b>R5</b>	$\frac{\text{Utilidad Neta}}{\text{Activos Totales}}$
	<b>R6</b>	$\frac{\text{Utilidad antes de Impuestos e Interes (EBITDA)}}{\text{Activos Totales}}$
	<b>R7</b>	$\frac{\text{Utilidad antes de Impuestos e Interes (EBITDA)}}{\text{Pasivos a Corto Plazo}}$
<b>SOLVENCIA</b>		
<b>Estado de Cambios en la Situación Financiera (E.C.S.F)</b>	<b>R8</b>	$\frac{\text{Pasivos Totales}}{\text{Activos Totales}}$
<b>LIQUIDEZ (Activos Líquidos / Activos)</b>		
<b>Estado de Cambios en la Situación Financiera (E.C.S.F)</b>	<b>R9</b>	$\frac{\text{Efectivo}}{\text{Activos Totales}}$
	<b>R10</b>	$\frac{\text{Activos Corrientes}}{\text{Activos Totales}}$

	R11	$\frac{\text{Capital de Trabajo}}{\text{Activos Totales}}$
<b>LIQUIDEZ (Activos Líquidos / Deuda Corriente)</b>		
Estado de Cambios en la Situación Financiera (E.C.S.F)	R12	$\frac{\text{Activos Corrientes}}{\text{Pasivos Corrientes}}$
<b>ACTIVIDAD (Rotación)</b>		
Estado de Resultados- ORI	R13	$\frac{\text{Capital de Trabajo Neto}}{\text{Ventas}}$
	R14	$\frac{\text{Ventas}}{\text{Activos Totales}}$

Elaboración Propia

Los 14 ratios financieros se calcularon para las 123 empresas del Sector de la construcción pertenecientes a la muestra del método *revisar Anexo 4*, posteriormente se reemplazaron los resultados de los ratios R3, R4, R5, R6, R7, R11 y R13 por variables dummy ya que se estableció como criterio de decisión en el estudio de la solvencia que dichos ratios determinan un estado de insolvencia si son negativos y un estado de solvencia si son positivos entonces, el valor de la variable dummy fue 0 cuando el ratio es negativo y 1 cuando el ratio es positivo.

También se utilizó un proceso de estandarización para los resultados de los ratios R1, R2, R8, R12, R14, el proceso se conoce como reescalado, este proceso consiste en ubicar los valores de variable entre un valor máximo y mínimo, por lo regular entre 0 y 1, el objetivo es que las variables sean comparables y los valores muy extremos no afecten la clasificación. La fórmula utilizada para el proceso de reescalado es la siguiente:

$$x.escalada = \left( \frac{x - \min(x)}{\max(x) - \min(x)} \right)$$

Donde  $x$  hace referencia a los valores del ratio,  $\min(x)$  es el valor mínimo del ratio y  $\max(x)$  es el valor máximo del ratio para las 123 empresas de la muestra respectivamente. A los ratios R9 y R10, no se les realizó ningún ajuste o normalización por lo que se dejaron igual que los calculados inicialmente. Una vez reemplazados los ratios correspondientes por variables dummy y aplicado el proceso de reescalado en los ratios requeridos, se definió que esas variables independientes serían la base de datos para el presente modelo de clasificación *revisar Anexo 5*. Para la aplicación y evaluación del método sólo se tomó la información de los años 2017 y 2018.

### **Etapa 3: Aprendizaje No Supervisado**

Para el análisis y modelamiento de la información se usó el software informático R<sup>2</sup>, la decisión de usar esta herramienta obedece a que, es un software estadístico que permite no solo el análisis sino la clasificación de variables de datos contribuyendo de esta manera a cumplir con el objetivo de catalogar las empresas a partir de ratios financieros se agrupan como solventes e insolventes financieramente.

Para iniciar el proceso en esta etapa se instaló y compiló un conjunto de paquetes<sup>3</sup> específicos de R `advclust` (Bagus & Pramana, 2016), el cual es necesario para implementar y

---

<sup>2</sup> Rstudio es un entorno de desarrollo integrado para el lenguaje de programación R.

<sup>3</sup> Librería (`install.packages("advclust")` y `library(advclust)`)

ejecutar el algoritmo fuzzy c-means de Bezdek (1984), es decir, a través del agrupamiento difuso. Se realizaron diferentes pruebas para encontrar la partición adecuada o número de clústeres adecuado entre 2 y 8 clústeres. Se establece un máximo de 8 clústeres debido a su baja interpretabilidad, es decir, clasificar a una empresa en uno de ocho estados de insolvencia no permite comprender una tendencia en empresas con ratios muy variables. Lo ideal, de acuerdo con la literatura encontrada son máximo 5 clústeres esto debido al criterio de convergencia, ya que entre más clústeres halla del conjunto inicial de objetos igual número de centroides se generaran, uno por cada cluster lo que reducirá la distancia entre los centroides afectando la distribución y los grados de pertenecía de cada objeto en cada cluster, lo anterior se verá reflejado en la variación de los índices como por ejemplo el (PC) y (CE), los cuales presentan resultados favorables en cuanto a la dispersión hasta 5 clústeres en su aplicación a la minería de datos como se puede evidenciar más adelante en la Figura 54.

Para realizar el análisis de los clústeres es importante tener presente que, el objetivo del algoritmo fuzzy c-means consiste en agrupar datos basándose en las características que poseen. Por lo tanto, para el análisis e interpretación, se busca encontrar clústeres de tal manera que cada objeto en su interior, sea los más *parecido* o similar a los que hay en él y, lo más disimilar a otro elemento de otro clúster. En este sentido, se empleó la distancia euclídea.

Para establecer el número final de conglomerados a formar, no hay un procedimiento determinado, lo que indica que una posible solución a este problema, corresponde con la interpretación de los índices de evaluación propio de particiones borrosas, presentados en la sección anterior. En el caso particular de la muestra seleccionada, se calcularon dichos índices para un número entre 2 y 8 clústeres y los resultados se presentan en la Figura 54.

**Figura 54***Numero de clústeres vs índices de partición Rstudio*

Índices No Clústeres	PC	MPC	CE	S	XB	Kwon	Tang
2	0,8456	0,6912	0,2546	0,1166	0,1166	15,345	15,345
3	0,7875	0,6813	0,398	0,1425	0,1425	21,14	19,34
4	0,6656	0,5423	0,6329	0,935	0,935	169,22	133,14
5	0,7706	0,7713	0,5126	0,068	0,068	16,48	10,459
6	0,609	0,531	0,798	1,19	1,19	319,43	180,98
7	0,622	0,559	0,7978	0,954	0,954	348,91	155,98
8	0,548	0,483	0,964	1,79	1,79	786,97	302,18

Elaboración Propia

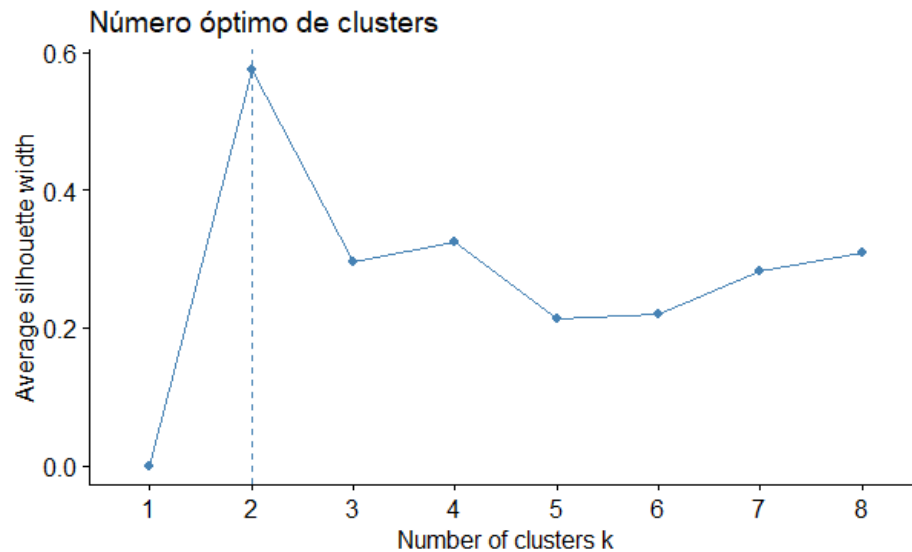
De acuerdo con lo anterior, la cantidad de clústeres apropiado para el conjunto de datos se seleccionó con base en los siete índices más representativos en la literatura de acuerdo a su amplio uso y efectividad. Como se puede observar en la Figura 53 para una partición con dos clústeres, los índices PC, CE y Kwon coinciden. Por lo tanto, de acuerdo a dicha coincidencia se ha decidido establecer una clasificación con dos clústeres, como la mejor.

Con el fin de validar la anterior selección se ha empleado adicionalmente el método de validación de clusteres average silhouette (Batool & Hennig, 2021), este coeficiente mide o cuantifica cual es la mejor asignación que se ha hecho de una observación al comparar su similitud con el resto de observaciones del clúster perteneciente frente a los demás clústeres. Su valor puede variar entre -1 y 1, siendo valores altos un indicativo de que la observación se

ha clasificado en el clúster correcto. El método muestra que el número óptimo clúster es igual a dos como se puede observar en la Figura 55.

**Figura 55**

*Gráfica de número óptimo de clústeres para clasificación de insolvencia en muestra de empresas de construcción.*



*Nota:* Gráfica del número óptimo de cluster's para determinar el Método de Insolvencia Financiera, Elaboración propia.

En cuanto al algoritmo, una clasificación borrosa o difusa con dos clases produce una matriz de pertenencia difusa ( $U$ ) y un centroide de agrupación difuso ( $V$ ). Existen dos maneras para determinar la etiqueta de las particiones, la primera es tomar el mayor valor de la membresía en el punto de datos y determinar la etiqueta del grupo  $V_1$  y  $V_2$ . La segunda manera hace referencia al ordenamiento de la base de datos donde se sabe de antemano cuales son las empresas solventes e insolventes.

Al analizar la membresía de una empresa sin problemas de solvencia en la agrupación, se determinó que en la partición  $V_1$  están las empresas sin problemas de insolvencia y en la

partición V2 las empresas con problemas de insolvencia. Es importante especificar que en la base de datos se agrega manualmente a una nueva columna llamada categoría con los resultados de la agrupación V1 y V2, es decir con la etiqueta solvente o insolvente para cada empresa, se guardó como base de datos categorizada y se importó de nuevo al software informático. *Revisar Anexo 6.*

#### **Etapas 4: Aprendizaje Supervisado**

Esta etapa se compone de dos partes. En la primera parte se instalaron y compilaron en el software informático los paquetes<sup>4</sup> correspondientes para la aplicación de algoritmos supervisados, los cuales contienen funciones que sirven para agilizar procesos de entrenamiento de algoritmos con problemas complejos de regresión y clasificación.

Una vez conformada la base de datos con las etiquetas establecidas en la fase anterior, se aplicó el algoritmo *KNN*, para lo cual, la información se repartió de la siguiente manera: la información de las 123 empresas se distribuyó así: el 70% de los datos (aproximadamente 87 empresas) se emplearon para entrenar el algoritmo y el 30% restantes (aproximadamente 36 empresas) se utilizaron para poner a prueba la capacidad de clasificación o predicción del mismo. Los porcentajes de la división del conjunto de datos se introdujeron manualmente al algoritmo, al igual que el valor de  $k$  correspondiente a los vecinos más cercanos variando  $k$

---

<sup>4</sup> Librería (install.packages("class"), library(class) )  
Librería (install.packages("caret"), library(caret) y library(lattice))

desde  $k = 1$  hasta  $k = 10$  vecinos, en la Figura 56 se presentan las correspondientes matrices de confusión<sup>5</sup> que permiten determinar la precisión del algoritmo en el caso estudiado.

**Figura 56**

*Matrices de confusión para  $k$  vecinos más cercanos*

$k$	Clasificación	Predicción Insolvente	Predicción Solvente
1	Actual Insolvente	7	0
	Actual Solvente	0	29
2	Actual Insolvente	7	0
	Actual Solvente	0	29
3	Actual Insolvente	7	0
	Actual Solvente	0	29
4	Actual Insolvente	7	0
	Actual Solvente	0	29
5	Actual Insolvente	7	0
	Actual Solvente	0	29
6	Actual Insolvente	7	0
	Actual Solvente	0	29
7	Actual Insolvente	7	0
	Actual Solvente	0	29
8	Actual Insolvente	7	0
	Actual Solvente	0	29
9	Actual Insolvente	7	0
	Actual Solvente	0	29
10	Actual Insolvente	7	0
	Actual Solvente	0	29

Elaboración Propia

<sup>5</sup> Herramienta empleada en la inteligencia artificial para medir la precisión en los algoritmos de clasificación bajo aprendizaje supervisado.

En la anterior Figura se puede evidenciar que la clasificación del conjunto de prueba es consistente para los diferentes valores de  $k$ . Es decir, sin importar el valor de  $k$ , la clasificación obtenida para las empresas de prueba siempre es la misma. De igual forma, se evidencia que no existen fallos, en el caso particular estudiado, en cuanto a empresas mal clasificadas. Para la segunda parte de la etapa se revisa el aprendizaje del algoritmo, es decir la capacidad de clasificación y/o predicción del mismo. Para dicho fin, se emplean como variables de entrada del algoritmo, los ratios financieros de dos nuevas empresas A y B creadas sin etiquetas o categoría, para la empresas A se han tomado los promedios de los ratios de las empresas solventes y, para la empresa B el promedio de la empresas insolventes. La Figura 57 presenta los valores asignados para los ratios de entrada del algoritmo.

**Figura 57**

*Valores ajustados de los ratios de las nuevas empresas sin clasificar*

<b>Ratios</b>	<b>Empresa A</b>	<b>Empresa B</b>
R1escalada	0,02437012	0,00386958
R3escalada	0,00927584	0,00010256
R4 Dummy	1	0
R5 Dummy	1	0
R6 Dummy	1	0
R7 Dummy	1	0
R8 Dummy	1	0
R9escalada	0,03390446	0,11647311
R12	0,07335481	0,03044446
R13	0,71072716	0,59115521
R14 Dummy	1	0
R15escalada	0,01957428	0,00597847
R17 Dummy	1	0
R18escalada	0,09828219	0,15509397

Elaboración Propia

Al utilizar de nuevo el algoritmo *KNN* se asignó la posible clasificación entre solvente o insolvente a las nuevas empresas sin etiquetar, el algoritmo toma como referencia al conjunto de datos de entrenamiento para dicha clasificación. En la Figura 58 se presenta la clasificación obtenida.

**Figura 58**

*Clasificación de las nuevas empresas*

Empresas	Clasificación
A	Solvente
B	Insolvente

Elaboración Propia

La clasificación para el promedio de las empresas solventes o empresa A fue solvente y la clasificación para el promedio de empresas insolventes o empresa B fue insolvente, ambos resultados son coherentes con respecto a las clasificaciones esperadas para la empresa A y B usadas para ejemplificar y validar la consistencia del algoritmo empleado en el método propuesto.

### **Etapas 5: Análisis y Resultados**

En esta sección se presenta de forma más detallada los resultados del procedimiento de clasificación. Para la validación de los resultados obtenidos, se ha empleado la información de la Superintendencia de Sociedades sobre las empresas que se han acogido a la ley de insolvencia 1116 en el año 2018. Por lo tanto, bajo el método propuesto se ha pronosticado un estado de solvencia o insolvencia de 111 empresas para el año 2018 empleando como

referencia los datos de los ratios financieros del año 2017. En primer lugar, el algoritmo *fuzzy c-means* clasifico correctamente 92 de las 111 empresas requeridas por la superintendencia de sociedades para el año 2017 como posibles candidatas al estado de *solventes* para el año 2018 y, ha clasificado incorrectamente 19 de las 111 empresas como posibles candidatas *insolventes* para el mismo año.

También el algoritmo clasifico 7 de las 12 empresas requeridas por la superintendencia de sociedades para el año 2017 y que se acogieron al régimen de insolvencia financiera de Colombia en el 2018 como posibles candidatas a insolventes para el mismo año 2018 y las restantes 5 de 12 como posibles candidatas a solventes. Es importante resaltar en esta parte la capacidad de clasificación ya que el algoritmo fue capaz de clasificar correctamente para el año 2018, 7 de las 12 empresas que entraron a régimen de insolvencia como insolventes, en la Figura 59 se resumen los resultados.

**Figura 59**

*Resultados del Método para la determinación de Insolvencia Financiera Machine Learning*

<b>Diagnostico Modelo de Clasificación de Insolvencia Financiera</b>	<b>Cantidad</b>	<b>Posibles Solventes %</b>	<b>Posibles Insolventes %</b>
Empresas que requeridas por Superintendencia de Sociedades para el 2017	111	82,9	17,1
Empresas que se acogieron al Régimen de Insolvencia en 2018	12	41,6	58,3
Nivel de Confianza %		80,48	
Error %		19,51	

Elaboración Propia

De acuerdo con lo anterior, el método propuesto presenta un nivel de precisión del 80,48 % con un margen de error promedio del 19,51%, esto quiere decir que en promedio clasifiqué correctamente 8 de cada 10 empresas para un total de 99 de las 123 empresas pertenecientes a la muestra de estudio, Revisar Anexo 7.

En la Figura 60 se muestran las 24 empresas que no fueron clasificadas correctamente, 19 empresas con error tipo I es decir, que no tuvieron problemas de solvencia para el año 2018 y fueron clasificadas como insolventes y 5 empresas con error tipo II refiriéndose a las que fueron clasificadas como solventes para el año 2018 y presentaron problemas de insolvencia financiera para el mismo año, adicionalmente se muestra para cada objeto el grado de pertenencia difusa (fuzzy) en los clústeres.

**Figura 60**

*Grado de pertenencia de las empresas clasificadas incorrectamente*

<b>Error tipo I (Empresas solventes clasificadas como insolventes)</b>		
<b>Solvente</b>	<b>Insolvente</b>	<b>Clasificación</b>
0,134953759	0,865046241	Insolvente
0,176689245	0,823310755	Insolvente
0,201359496	0,798640504	Insolvente
0,165611605	0,834388395	Insolvente
0,121473123	0,878526877	Insolvente
0,172913173	0,827086827	Insolvente
0,125838469	0,874161531	Insolvente
0,367404384	0,632595616	Insolvente
0,117768282	0,882231718	Insolvente
0,118766119	0,881233881	Insolvente
0,178031849	0,821968151	Insolvente
0,171812339	0,828187661	Insolvente
0,178992531	0,821007469	Insolvente

0,130459927	0,869540073	Insolvente
0,110826444	0,889173556	Insolvente
0,178794714	0,821205286	Insolvente
0,266910922	0,733089078	Insolvente
0,124691926	0,875308074	Insolvente
0,121161495	0,878838505	Insolvente
Error tipo II (Empresas insolventes clasificadas como solventes)		
0,975236115	0,024763885	Solvente
0,533884618	0,466115382	Solvente
0,532167022	0,467832978	Solvente
0,559316169	0,440683831	Solvente
0,995823937	0,004176063	Solvente

Elaboración Propia

Como se puede evidenciar en la Figura 60 el algoritmo FCM asignó a cada objeto un valor de pertenencia dentro de cada cluster y por consiguiente el objeto específico puede pertenecer parcialmente a más de un clúster. Uno de los potenciales del método corresponde con precisamente poder analizar el conjunto de empresas bajo el error tipo I, debido a que presentan características de insolvencia (de acuerdo con los ratios estudiados) aunque aún no han ingresado a proceso de insolvencia bajo la normativa colombiana. De forma clara, se establece entonces una alerta o un indicador que permite detectar de forma temprana posibles estados de insolvencia.

En cuanto a las empresas bajo el error tipo II, una mirada más detalla a los grados de pertenencia a cada estado, *solvente e insolvente*, permite identificar justamente que dichas empresas para el año 2017 presentaban en su mayoría altos grados de pertenencia al estado de insolvente. Por ejemplo, una empresa que para 2017 presentaba un grado de insolvencia de 0,4661, requería en aquel momento especial atención sobre su desarrollo futuro.

## 5 Conclusiones y trabajos futuros

La contabilidad no solo sirve para registrar los hechos económicos pasados de una empresa o para el análisis y seguimiento de la eficiencia financiera en un determinado punto, entre otras aplicaciones también sirve como herramienta para evaluar a futuro la situación de una empresa en el mercado según la posible tendencia de sus datos referente a la solvencia o insolvencia financiera, con impacto en el proceso de toma de decisiones a priori o antes de posibles eventos de fracaso empresarial.

El método propuesto para la determinación de insolvencia financiera está ajustado para el contexto Colombiano en el sector de la construcción específicamente las empresas que reportan sus estados financieros bajo NIIF Plenas individuales en el año 2017. Dicho método diagnóstica la clasificación de las empresas en términos de posible *solvencia* e *insolvencia* para el año 2018. Para dicha clasificación se tiene en cuenta que no es posible considerar a las empresas como organizaciones estáticas sino como organizaciones dinámicas ya que se toman decisiones y, las redes, procesos, recurso humano y el conocimiento se transforma en el tiempo.

Por la anterior razón, no se realizó un estudio longitudinal para las empresas, en contraste se analizó un año en particular. En este sentido, el uso de machine learning y algoritmos de aprendizaje supervisado y no supervisado para el modelamiento de la información facilita el procesamiento e interpretación de los ratios de un conjunto de empresas en términos de solvencia e insolvencia respectivamente sin exigir condiciones o comportamientos sobre las variables de entrada del algoritmo.

Los ratios seleccionados fueron apropiados y se resalta que es importante evaluar los ratios de liquidez, endeudamiento y rentabilidad como los puntos fuertes en donde las empresas del sector de la construcción deben concentrar su mayor atención para el estudio de la insolvencia financiera.

El algoritmo refleja un 80,48% de efectividad logrando clasificar de forma correcta 99 de las 123 empresas consideradas en la investigación y por lo tanto, determinando para el periodo de análisis el posible estado de solvencia e insolvencia empresarial. Se puede concluir que los objetivos planteados al principio del trabajo se han alcanzado y la metodología seguida para ello ha demostrado ser adecuada.

Una de las ventajas más importantes del uso del machine learning para la determinación de la insolvencia financiera a partir de los ratios es que se puede trabajar un mayor volumen de información sin que ello conlleve a un mayor gasto en tiempo y recursos para el análisis de los resultados, reduciendo el sesgo humano y por ende el error.

No se requiere emparejar las muestras de análisis entre empresas frágiles y no frágiles lo que permite trabajar una mayor cantidad de información al no estar limitada la muestra por el número de empresas frágiles del periodo de estudio, así mismo no se hace necesaria la información empresarial de los estados financieros por más de uno o dos periodos para su clasificación al no ser un estudio longitudinal y considerar la dinámica de las empresas y del sector económico.

Otra ventaja significativa de los algoritmos de clasificación de machine learning con respecto a las técnicas estadísticas es que no se requieren considerar supuestos iniciales para el

análisis como tampoco se hace necesario estudiar la distribución que siguen los datos, pudiéndose emplear criterios de insolvencia y técnicas para que los datos sean comparables.

Además, el uso del FCM permite aplicar una partición suave sobre el conjunto de datos lo que posibilita un mejor análisis ya que parcialmente los datos pueden presentar características en mayor o menor grado en más de un cluster o estados de solvencia y no pertenecer exclusivamente a uno “partición dura”, permitiendo evidenciar la transición de un estado a otro e inclusive identificar síntomas o tendencias negativas que conlleve a fallos futuros relacionados con insolvencia.

El algoritmo de clasificación de la insolvencia puede ser de gran utilidad como herramienta para la toma de decisiones. Por ejemplo, las entidades financieras pueden agrupar las empresas en conglomerados o clústeres en términos de su solvencia facilitando la opción de otorgar o no un crédito a una empresa en cuestión.

Otra aplicación, puede ser en el ámbito empresarial, bajo el cual determinada firma quiere compararse con otras firmas similares del mismo sector, pero no sabe cuáles elegir, el algoritmo le permite agrupar las firmas en términos de sus indicadores financieros, para luego comparar con las más semejantes dentro mismo clúster y así enfocar el proceso de toma de decisiones en pro de mejorar la competitividad de dicha firma.

El algoritmo puede ser una ventaja para agrupar empresas en términos de sus indicadores de acuerdo con su naturaleza de lo que se quiera medir como la liquidez, endeudamiento, rotación y rentabilidad. Es cuestión de modificar la base de datos la cual es el principal insumo del algoritmo, determinar el número de clúster óptimo y las etiquetas de cada clúster. En este sentido, es importante resaltar que el algoritmo y en general el método

propuesto, se aplica de forma particular a cada sector económico específico, junto con ratios específicos según el sector. Lo anterior implica que el método propuesto funciona de forma específica para un sector económico y debe ser adaptado, en cuanto a las variables de entrada, para un nuevo caso.

De igual forma, se logró determinar que uno de los principales inconvenientes que puede tener el método propuesto, y en general, cualquier análisis con indicadores financieros corresponde con la denominada contabilidad creativa ya que es una práctica común que usan las empresas para no reflejar en sus estados financieros la realidad económica de una empresa, ocultando falencias o malos manejos.

Tanto los métodos basados en k vecinos más cercanos como los basados en K-means corresponden con algoritmos que se pueden aplicar a cualquier grupo de empresas y en cualquier sector de la económica colombiana, por lo tanto, el método propuesto no presenta ninguna restricción en términos de sector o en termino de comportamiento de las variables.

Finalmente, el método propuesto para la predicción o clasificación de la insolvencia financiera se puede complementar con otros indicadores que por ejemplo midan el retorno de la inversión como el valor económico agregado [EVA], la tasa interna de retorno [TIR] y el valor presente neto [VPN] entre otros, con el objetivo de realizar un análisis más detallado y valorar de forma más óptima la situación financiera de las empresas en términos del pronóstico prematuro de su solvencia.

## 6 Bibliografía

- Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. <https://doi.org/10.1111/jofi.12742>
- Altman, E. I. (2000). Predicting financial distress of companies: En revisiting the z-score and zeta® models. *Journal of Finance*. 1-54. <http://pages.stern.nyu.edu/~ealtman/Zscores.pdf>
- Altman, E. I., & Hotchkiss, E. (2011). Corporate Credit Scoring-insolvency Risk Models. En *Corporate Financial Distress and Bankruptcy. Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt* (3 ed. 231-264) Wiley. <https://doi.org/10.1002/9781118267806.ch11>
- Arnold, P. J. (2005). Disciplinar a la reglamentación nacional: la Organización Mundial del Comercio y el mercado de servicios profesionales. *Accounting, Organizations and Society* (30), 299–330. Elsevier.
- Bagus, F., & Pramana, S. (2016). Package “advclust” Object Oriented Advanced Clustering. *CRAN*, 1-24. <https://cran.r-project.org/web/packages/advclust/advclust.pdf>
- Batool, F., & Hennig, C. (2021). Clustering with the Average Silhouette Width. *Computational Statistics & Data Analysis*, (158), 1–41. <https://doi.org/10.1016/j.csda.2021.107190>
- Banco Bilbao Vizcaya Argentaria [BBVA]. (2019). “Machine learning”: ¿qué es y cómo funciona? <https://www.bbva.com/es/machine-learning-que-es-y-como-funciona/>
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting*

*Research*, (4), 71-111. <https://doi.org/10.2307/2490171>

Blum, M. P. (1974). The Failing Company Doctrine. *Boston College Law Review*, 16(1), 11–12. <http://lawdigitalcommons.bc.edu/bclr/vol16/iss1/2>

Charpentier, J. (2014). Modelos de Beaver, Ohlson y Altman. *Tec Empresarial*, 8(3), 29–40. [https://revistas.tec.ac.cr/index.php/tec\\_empresarial/article/view/2078](https://revistas.tec.ac.cr/index.php/tec_empresarial/article/view/2078)

Chen, H., Yang, B., Wang, G., Liu, J., Xu, X. & Wang, S. (2001). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowledge-Based Systems* (24), 1348-1359.

Cvalora (2020, 15 abril). Sistema de Recuperación Empresarial. <https://www.cvalora.co/sistema-de-recuperacion-empresarial/>

Decreto 560/20, abril 15, 2020. Ministerio de Comercio, Industria y Turismo. (Colombia). Obtenido el 20 de febrero de 2021. <https://cijuf.org.co/normatividad/decreto-legislativo/2020/decreto-legislativo-560.html>

Departamento Administrativo Nacional de Estadística [DANE] (2018). Históricos Producto Interno Bruto PIB. Obtenida el 11 de noviembre de 2020 <https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-trimestrales/historicos-producto-interno-bruto-pib>

Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167. <https://doi.org/10.2307/2490225>

Deakin, E. B. (1976). Distributions of Financial Accounting Ratios: Some Empirical

Evidence. *American Accounting Association*, 51(1), 90–96.

<http://www.jstor.com/stable/245375>

Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business

Failure Prediction. *The Journal of Financial and Quantitative Analysis*, 7(2), 1477-1493.

<https://doi.org/10.2307/2329929>

García, C. & Gómez, I. (2006). Algoritmos de aprendizaje: knn & kmeans. Universidad

Carlos III de Madrid. <http://www.it.uc3m.es/jvillena/irc/practicas/08-09/06.pdf>

Gendler, M. (2019). Personalización algorítmica y apropiación social de tecnologías.

*CLACSO*, 1–408.

<http://biblioteca.clacso.edu.ar/clacso/se/20191128031455/Tecnologias-digitales.pdf>

Gordon, M. J. (1971). Towards a Theory of Financial Distress. *The Journal of Finance*, 26(2),

347–356. <https://doi.org/10.1111/j.1540-6261.1971.tb00902.x>

Hernández, M. (2014). Modelo financiero para la detección de quiebras con el uso de análisis

discriminante múltiple. *InterSedes*, 15(32), 4–19.

<https://doi.org/10.15517/isucr.v15i32.17792>

Horrigan, J. (1968). A Short History of Financial Ratio Analysis. *The Accounting Review*,

43(2), 284–294.

Leaño, H. M. (2004). Bankruptcy Prediction Model Using Discriminant Analysis on

Financial. *Corporate Balance Sheets*.

Ley 1116/06, diciembre 27, 2006. Diario Oficial [D.O.]: 46.494. (Colombia). Obtenida el 15

agosto de 2020

[https://www.sic.gov.co/recursos\\_user/documentos/normatividad/Ley\\_1116\\_2006.pdf](https://www.sic.gov.co/recursos_user/documentos/normatividad/Ley_1116_2006.pdf)

Libby, R. (1975). Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence. *Journal of Accounting Research*, 13(1), 150–161. <https://doi.org/10.2307/2490653>

López-Roldán, P., & Fachelli, S. (2016). Análisis de Regresión Logística. Metodología de La Investigación Social Cuantitativa. *Universitat Autònoma de Barcelona*, (1), 1–18. [https://ddd.uab.cat/pub/caplli/2016/163570/metinvsocua\\_a2016\\_cap3-10.pdf](https://ddd.uab.cat/pub/caplli/2016/163570/metinvsocua_a2016_cap3-10.pdf)

Ibarra, A. (2001). Análisis de las dificultades financieras de las empresas en una economía emergente: Las bases de datos y las variables independientes en el sector hotelero de la bolsa mexicana de valores. *Universitat Autònoma de Barcelona*, 1–307. <https://www.tesisenred.net/handle/10803/3941#page=1>

Martin, M. & Idoate, M. (2015). *A. prendizaje de distancias basadas en disimilitudes para el algoritmo de clasificación kNN* [Tesis de grado, Universidad Pública de Navarra]. Repositorio Institucional. <http://academicae.unavarra.es/handle/2454/18430?show=full>

Mendialdua, I. (2015). *Contributions on Distance-Based algorithms , Multi-Classifer Construction and Pairwise Classification*. [Tesis doctoral inédita]. Universidad del País del Vasco. Repositorio Institucional. <https://addi.ehu.es/handle/10810/15943>

Martínez, O. (2003). Determinantes de fragilidad en las empresas colombianas. *Temas de Estabilidad Financiera*, (1). <http://repositorio.banrep.gov.co/handle/20.500.12134/2096>

- Matich, D. J. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. *Universidad Tecnológica Nacional Facultad Regional Rosario*, 1–55.  
[https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5\\_ano/orientadora1/monograis/matich-redesneuronales.pdf](https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_ano/orientadora1/monograis/matich-redesneuronales.pdf)
- Min, J. & Jeong, C. (2009). A binary classification method for bankruptcy prediction. *Expert Systems with Applications* (36), 5256-5263.  
<http://www.ncbi.nlm.nih.gov/pubmed/27157931>
- Montoya, O. (2007). Aplicación del análisis factorial a la investigación de mercados, Caso de estudio. *Scientia Et Technica*, 1(1), 281–286.  
<http://dialnet.unirioja.es/servlet/articulo?codigo=4804281&info=resumen&idioma=ENG>
- Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109-131. <https://doi.org/10.2307/2490395>
- Peña, T., Martínez, S., & Abudu, B. (2009). Predicción de bancarrota: Una comparación de técnicas estadísticas y de aprendizaje supervisado para computadora. Predicción de Bancarrota: Una Comparación de Técnicas Estadísticas y de Aprendizaje Supervisado Para Computadora. *Banco de México*, 1–28. <https://doi.org/10.36095/banxico/di.2009.18>
- Pérez, J., González, K. & Lopera, M. (2013). Modelos de predicción de la fragilidad empresarial: aplicación al caso colombiano para el año 2011. *Perfil de Coyuntura Económica*, 0(22), 205–228. <https://www.redalyc.org/pdf/861/86131758010.pdf>
- Pessoa, A. (2016). Análisis de solvencia de las empresas: modelo estático versus modelo

dinámico.[Tesis doctoral inédita].Universidad de Zaragoza. Repositorio institucional

<https://zaguan.unizar.es/record/48319/>

Pnches, G., Caruthers, J., Mingo, K. & Eubank, A. (1975). The Hierarchical Classification of Financial Ratios. *Journal of Business Research*, 3(4), 295–310.

<https://www.sciencedirect.com/science/article/abs/pii/0148296375900119>

Python (2020, 19 de noviembre). Fuzzy c-means clustering. [https://pythonhosted.org/scikit-fuzzy/auto\\_examples/plot\\_cmeans.html](https://pythonhosted.org/scikit-fuzzy/auto_examples/plot_cmeans.html)

RAE. (2020). Insolvencia. Definición. *Diccionario de la lengua española*. RAE - ASALE.

<https://dle.rae.es/insolvencia>

Rebala, G., Ravi, A. & Churiwala, S. (2019). Classification. En Springer. *An introduction to machine learning*. (57-66). Springer. [https://link.springer.com/chapter/10.1007%2F978-3-030-15729-6\\_5](https://link.springer.com/chapter/10.1007%2F978-3-030-15729-6_5)

Rojas Diaz, J., Chavarro Porras, J. C., & Moreno Laverde, R. (2009). Tecnicas de logica difusa aplicadas a la mineria de datos. *Scientia et Technica*, 3(40), 1–6.

<https://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095>

Rosillo, J. (2002). Modelo de predicción de quiebras de las empresas colombianas. *Revista Innovar Journal Revista de Ciencias Administrativas y Sociales*, 1(19), 109–124.

<https://revistas.unal.edu.co/index.php/innovar/article/view/23959>

Registro Único Empresarial y Social [RUES], (2018). <http://www.rues.org.co/>

Rufus, R. (2003). Financial Ratios: Predictive Power and Z-Score. *The Valuation Examiner*,

14–17. [http://www.rufusandrufus.com/financiam\\_ratios.pdf](http://www.rufusandrufus.com/financiam_ratios.pdf)

Sandoval, L. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica N° 11*, 36–40.

[http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6\\_RT2018.pdf](http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf)

Superintendencia de Sociedades. (2012). *Causas de la insolvencia empresarial. Revista Supersociedades*, (4), 4-34.

[https://www.supersociedades.gov.co/delegatura\\_aec/Paginas/Base-completa-EF-2019.aspx](https://www.supersociedades.gov.co/delegatura_aec/Paginas/Base-completa-EF-2019.aspx)

Superintendencia de Sociedades. (2019).

[https://www.supersociedades.gov.co/prensa/Documentos\\_publicaciones/4-Causas-de-la-Insolvencia-en-Colombia.pdf](https://www.supersociedades.gov.co/prensa/Documentos_publicaciones/4-Causas-de-la-Insolvencia-en-Colombia.pdf)

Taffler, R. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking and Finance*, 8(2), 199–227. [https://doi.org/10.1016/0378-4266\(84\)90004-9](https://doi.org/10.1016/0378-4266(84)90004-9)

Torrado, M., & Berlanga, V. (2013). Análisis discriminante mediante SPSS. REIRE. *Revista d’Innovació i Recerca En Educació*, 6(2), 150–166.

<https://doi.org/10.1344/reire2013.6.26210>

Vinet, L., & Zhedanov, A. (2011). A “missing” family of classical orthogonal polynomials. *Journal of Physics A: Mathematical and Theoretical*, 44(8), 1689–1699.

<https://doi.org/10.1088/1751-8113/44/8/085201>

Walteros, M. (2018). Modelo de predicción de insolvencia en el sector comercio de Bogotá.

*Universidad Piloto de Colombia*, (14), 168.

<http://polux.unipiloto.edu.co:8080/00004619.pdf>.

Yen, J., & Langari, R. (1998). *Fuzzy Logic: Intelligence, Control, and Information*. *New Jersey: Prentice Hall*, 1999, p. 351-362.

**Anexos**

Anexo 1. Base de datos Información Financiera Empresas requeridas por la Superintendencia de Sociedades con corte a 31 de diciembre de 2018.

Anexo 2. Base de datos Soporte de Insolvencia 31 de diciembre de 2019 Superintendencia de Sociedades.

Anexo 3. Selección de Ratios.

Anexo 4. Resumen Base de Datos Ratios Calculados.

Anexo 5. Base de datos Depurada.

Anexo 6. Base de datos Etiquetada.

Anexo 7. Resultados Predicción 2017.