



MODELO DE SEGUIMIENTO DE RIESGO DE CRÉDITO PARA EL CLIENTE  
INDEPENDIENTE, PARA UNA ENTIDAD FINANCIERA DEL VALLE DEL CAUCA

Por

Vanessa N. Granja García<sup>1</sup>

Nathalia García Montealegre<sup>2</sup>

Tesis presentada como requisito parcial para obtener el título de  
Magíster en Administración Financiera

Asesor: Mg. Jaime Alberto Ospina Mejía

Universidad EAFIT  
Escuela de Administración  
Maestría en Administración Financiera  
Cali  
Agosto de 2020

---

<sup>1</sup> vngranjag@eafit.edu.co

<sup>2</sup> ngarciaml@eafit.edu.co

© 2020 por Vanessa Granja y Nathalia García

Todos los derechos reservados

## **Dedicatoria**

Dedicamos con mucho amor el presente trabajo a nuestras familias y seres queridos por su apoyo moral y su paciencia; y de manera especial consagramos este trabajo a Dios, que nos guía y sostiene en tiempos adversos.

## **Agradecimientos**

Expresamos nuestra gratitud a todas las personas que de una u otra forma nos alentaron y apoyaron para concluir con éxito el presente trabajo, de manera especial a nuestras familias, por su amor y por inspirarnos a intentar ser mejores cada día.

A la Universidad EAFIT, por procurar una excelente calidad académica. A sus docentes, que se esforzaron cada día para contribuir a nuestra formación profesional. Y a nuestro asesor, el profesor Jaime Alberto Ospina Mejía, por aceptar el reto de orientarnos en tan poco tiempo y por sus conocimientos y paciencia, que fueron pilares fundamentales para la consolidación y culminación de este proyecto.

Finalmente, a la entidad financiera estudiada y a todas las personas que estuvieron involucrados en la obtención de información y que de manera desinteresada y sin inconvenientes brindaron lo necesario para construir el proyecto.

## Resumen

El propósito de esta investigación es proponer un modelo de seguimiento de riesgo de crédito para la cartera de clientes de perfil independiente de una entidad financiera de Valle del Cauca, que aporte información relacionada con variables relevantes ligadas al análisis y la gestión del riesgo de crédito de este segmento. El desarrollo del modelo de seguimiento permite agruparlos según sus características y clasificarlos por su nivel de riesgo, predecir con más certeza un eventual impago y tomar acciones preventivas frente a este hecho, cumpliendo a su vez con las políticas de riesgo definidas en el Sistema de administración de riesgo de crédito (SARC), regulado en Colombia por la Superintendencia Financiera y fundamentado en el marco de referencia de los Acuerdos de Basilea. La metodología usada es la de técnicas de minería de datos como el *clustering*; la evaluación se lleva a cabo mediante árboles de decisión; al final se presentan las recomendaciones.

Palabras claves: riesgo de crédito, modelos de seguimiento, trabajador independiente, árboles de decisión, clúster.

## Abstract

The purpose of this research is to propose a credit risk monitoring model for the independent profile client portfolio of a financial entity in Colombia's department of Valle del Cauca, which provides information related to relevant variables associated to the analysis and management of the credit risk of this segment. The development of the monitoring model allows the grouping of the clients according to their characteristics and classifying them by their level of risk, predicting a possible default with more certainty and taking preventive actions against this fact, allowing in turn to comply with the risk policies defined in the Credit Risk Management System (Sistema de administración de riesgo de crédito, SARC), regulated in Colombia by the Superintendencia Financiera and based on the reference framework of the Basel Accords. The methodology used is that of data mining techniques such as clustering; the evaluation is carried out using decision trees; at the end the recommendations are presented.

Keywords: Credit risk, monitoring models, independent worker, decision trees, cluster.

## Tabla de contenidos

1. Introducción .....	1
2. Objetivos .....	3
2.1 Objetivo general.....	3
2.2 Objetivos específicos .....	3
3. Marco teórico .....	4
3.1 El riesgo de crédito (RC) .....	4
3.1.1 Definición y elementos.....	4
3.1.2 La Superintendencia Financiera de Colombia.....	4
3.1.3 El Sistema de administración de riesgo de crédito (SARC) y su Capítulo II.....	5
3.2 Modelos de medición, seguimiento y control .....	7
3.2.1 Modelos de medición del riesgo de crédito (RC).....	7
3.2.2 Modelos de seguimiento y control del riesgo de crédito (RC).....	11
3.3 Técnicas de minería de datos ( <i>data mining</i> ) .....	13
3.4 Contexto del trabajador independiente en Colombia.....	15
4. Metodología .....	19
4.1 Metodología y técnica de recolección de datos .....	19
4.1.1 Tipo de investigación.....	19
4.1.2 Definición de los parámetros y las fuentes .....	19
5. Resultados .....	33
6. Conclusiones y recomendaciones.....	35
6.1. Conclusiones .....	35
6.2. Recomendaciones .....	36
Referencias .....	37

## Lista de tablas

Tabla 1. Elementos del riesgo de crédito (RC) .....	4
Tabla 2. Modelos de medición del riesgo de crédito (RC).....	8
Tabla 3. Variables de mayor uso en los modelos de medición de riesgo.....	12
Tabla 4. La Entidad. Cartera de consumo e hipotecaria por actividad económica (diciembre de 2016-diciembre de 2019) .....	17
Tabla 5. La Entidad. Comportamiento interno a corto plazo – rango días (últimos seis meses) ...	21
Tabla 6. La Entidad. Comportamiento interno a largo plazo (12 meses).....	22
Tabla 7. La Entidad. Calificación interna a corto plazo (últimos seis meses) .....	22
Tabla 8. La Entidad. Calificación interna a largo plazo (últimos 12 meses) .....	23
Tabla 9. La Entidad. Rango del porcentaje amortizado .....	23
Tabla 10. La Entidad. Tipo de cobranza .....	23
Tabla 11. La Entidad. Porcentaje de participación de la deuda con el sector .....	23
Tabla 12. Número de alertas en el sector .....	24
Tabla 13. Calificación interna de la Entidad frente a la calificación externa del sector .....	24
Tabla 14. Porcentaje de participación del saldo de calidad con el sector .....	24
Tabla 15. La Entidad. Indicador de cartera vencida interno .....	25
Tabla 16. La Entidad. Nivel académico .....	25
Tabla 17. La Entidad. Menores a cargo .....	25
Tabla 18. La Entidad. Estrato.....	25
Tabla 19. La Entidad. Tipo de vivienda .....	26
Tabla 20. La Entidad. Patrimonio .....	26
Tabla 21. La Entidad. Rango de ingresos.....	26
Tabla 22. La Entidad. Rango de mora actual .....	26
Tabla 23. La Entidad. Clústeres – Segmentación de los clientes independientes .....	29
Tabla 24. Importancia del predictor de las variables de entrada.....	30

## Lista de figuras

Figura 1. Gestión del riesgo de crédito (RC) .....	6
Figura 2. Campo de aplicación de los modelos de medición del riesgo de crédito (RC) .....	7
Figura 3. División del modelo de análisis de conglomerados.....	11
Figura 4. Clasificación de las técnicas de minería de datos ( <i>data mining</i> ).....	14
Figura 5. Colombia. Tasa de ocupación asalariada y no asalariada – 23 ciudades (noviembre 2009-noviembre 2019) .....	16
Figura 6. Colombia. Tasa de ocupación formal e informal – 23 ciudades (octubre 2009-octubre 2019).....	16
Figura 7. La Entidad. Evolución de la participación del perfil del cliente independiente (diciembre de 2016-diciembre de 2019).....	17
Figura 8. La Entidad. Evolución del saldo de cartera del cliente independiente (diciembre de 2016-diciembre de 2019).....	18
Figura 9. La Entidad. Evolución de la cartera vencida del cliente independiente (diciembre de 2016-diciembre de 2019).....	18
Figura 10. La Entidad. Variables internas y externas – base de datos .....	21
Figura 11. La Entidad. Ventana del SPSS Modeler. Indicador de clientes buenos y malos.....	27
Figura 12. La Entidad. Ventana del SPSS Modeler. Entradas de K-Medias para definir los grupos de riesgo .....	28
Figura 13. La Entidad. Tamaños de los clústeres.....	28
Figura 14. La Entidad. Diagrama de árbol para el modelo de nivel de riesgo de la cartera independiente .....	31
Figura 15. La Entidad. Conglomerados de riesgo .....	33
Figura 16. La Entidad. Ventana del SPSS Modeler. Estrategia de seguimiento y asignación de riesgo .....	33
Figura 17. La Entidad. Clasificación del modelo.....	34
Figura 18. La Entidad. Métricas de evaluación.....	34
Figura 19. La Entidad. Matriz de coincidencias.....	34

## 1. Introducción

La Superintendencia Financiera de Colombia –en adelante la Superfinanciera– regula el sector financiero del país con el propósito de preservar la confianza pública y la estabilidad del sistema que, en su actividad de intermediación a través de la administración de recursos financieros y debido a sus características genera la mayor parte de sus beneficios y rentabilidad íntimamente ligados a los riesgos que asume. Uno de los principales riesgos financieros asumidos por las entidades que componen este sector es el riesgo de crédito o riesgo crediticio –en adelante RC–, propio de la administración de cartera de cuentas por cobrar (los préstamos). Al considerarse el riesgo como algo negativo, la gestión de riesgos financieros cobra vital importancia y refuerza la necesidad de una administración adecuada, puesto que toda entidad necesita identificar de forma anticipada las medidas que reduzcan su exposición al riesgo y no afecten de forma negativa su rentabilidad.

Los responsables de la gestión de riesgos en las entidades financieras deben garantizar la creación de procedimientos y desarrollar los sistemas que permitan una evaluación del crédito, facilitando la toma de decisiones tanto preventivas como correctivas, y mantener su posición dentro de los parámetros establecidos en la norma y por su entidad financiera en las diferentes etapas del Sistema de administración de riesgo de crédito –en adelante SARC– (Rankia, s. f.). Aunque algunas de estas entidades confían parte de la selección de sus clientes a la experiencia de sus analistas de crédito –que realizan el estudio de cada solicitud por separado–, es importante establecer modelos acertados no solo para el otorgamiento del crédito, sino también para el seguimiento y control que fortalezca el proceso realizado por ellos.

En el sector financiero existe un segmento de clientes que es considerado como riesgoso debido a su vulnerabilidad, al estar sujeto a todo tipo de cambios económicos y sociales. La mayor parte de los que conforman este segmento no cuenta con documentos que puedan soportar su capacidad de pago (ingresos), lo que ocasiona que las entidades requieran conocer de forma más específica sus condiciones y actividad económica para analizarlo. Esta situación causa que muchos clientes prefieran acudir a otros métodos de financiamiento –como los préstamos informales–, al percibir desconfianza por parte de este sector.

En este contexto, con el fin de mejorar y robustecer la gestión del RC en la etapa de seguimiento, se busca desarrollar un modelo de seguimiento para el segmento de clientes independientes de una entidad financiera de Valle del Cauca –en adelante la Entidad– a partir del apoyo en factores humanos para la primera etapa –selección de variables relevantes a incluir de acuerdo al criterio de la Entidad– y de técnicas de minería de datos como clústeres y árboles de decisión para la segunda – ejecución y evaluación del modelo en el software SPSS Modeler– (International Business Machines Corporation, IBM, s. f.) que se ajusten a sus necesidades, cumpliendo con la normatividad vigente en materia de RC y ajustándose a las particularidades de dichos clientes. Se espera que el modelo propuesto para esta investigación facilite el seguimiento y permita proponer acciones oportunas orientadas a reducir el no cumplimiento en la cartera. A fin de preservar su seguridad y la privacidad en la información, la entidad financiera sobre la cual se propone el modelo del presente trabajo es confidencial.

El presente trabajo se desarrolla en siete secciones:

- Introducción.
- Objetivos.
- Marco teórico, que contextualiza la definición de riesgo, la función de la Superfinanciera como ente regulador, las teorías relacionadas con los modelos de medición, seguimiento y control del RC, y finaliza con una contextualización del trabajador independiente en Colombia, todo ello con el propósito de presentar de manera general los modelos que usualmente se usan tanto para la medición del RC (el otorgamiento) como para el seguimiento y control, y que sirven de base para el modelo de seguimiento propuesto.
- Metodología, que describe las técnicas utilizadas y desarrolla el modelo de seguimiento a la cartera de clientes independientes de la Entidad, proponiendo los indicadores para dicho modelo a partir de su base de datos y de variables seleccionadas basadas en la experiencia y las necesidades de la Entidad, así como su posterior ejecución en el software Modeler SPSS del modelo K-Medias (International Business Machines Corporation, IBM, s. f.) y su evaluación a través de los árboles de decisión CHAID.
- Resultados, en la que se lleva a cabo la observación e interpretación de los hallazgos que llevan a la segmentación por nivel de riesgo de los clientes, y que será la base para definir y realizar un seguimiento adecuado por parte de la Entidad.
- Conclusiones y recomendaciones, a partir de los resultados del modelo de seguimiento del RC desarrollado para esta cartera.
- Referencias.

## 2. Objetivos

### 2.1 Objetivo general

Diseñar un modelo de seguimiento del riesgo de crédito (RC) para el cliente independiente en una entidad financiera de Valle del Cauca.

### 2.2 Objetivos específicos

Proponer los indicadores para la segmentación de cartera de los clientes independientes de la entidad financiera estudiada (la Entidad).

Seleccionar las variables cualitativas y cuantitativas relevantes para la entidad financiera estudiada (la Entidad) que serán empleadas en el modelo que se va a desarrollar.

Extraer la base de datos de los clientes independiente con información a corte del 31 de diciembre de 2019 de la entidad financiera estudiada (la Entidad).

### 3. Marco teórico

Esta sección despliega la información teórica más relevante relacionada con el propósito del trabajo. Inicia con la presentación general del RC y del SARC en Colombia, algunos de los principales modelos de medición y seguimiento del RC y las diferentes técnicas estadísticas y de minería de datos presentados en la literatura que se utilizan para medirlo y gestionarlo. A partir de esta información se resuelve usar el modelo de clúster K-Medias contenido en el software Modeler SPSS (International Business Machines Corporation, IBM, s. f.) para la segmentación de los clientes, en consideración a que está diseñado para la clasificación de individuos, no de variables; posteriormente se trabaja con el modelo de árboles de decisión CHAID (*Chi-square automatic interaction detection*, detector automático de Chi-cuadrado de interacción), también contenido en el Modeler SPSS (s. f.) para la evaluación del modelo y la reclasificación de los segmentos por niveles de riesgo; y finalmente se plantea el modelo de seguimiento. Todo lo anterior, en el marco de las condiciones y el contexto del trabajador independiente en Colombia.

#### 3.1 El riesgo de crédito (RC)

##### 3.1.1 Definición y elementos

El *riesgo de crédito* o *riesgo crediticio* (RC) se entiende como la probabilidad de que un cliente no cumpla con una parte o la totalidad del pago de su deuda u obligación financiera por causas relacionadas con quiebra o de otra índole (Chorafas, 2000). El RC se puede analizar a partir de los siguientes elementos [Tabla 1]:

Tabla 1. Elementos del riesgo de crédito (RC)

Elementos del riesgo de crédito
Riesgo de no cumplimiento: en el cual se rompe el acuerdo de pago de un contrato de crédito; sin embargo, se otorgan períodos de gracia antes de determinarlo como incumplimientos en el pago.
Exposición: relacionada con la falta de certeza sobre los montos en riesgo futuros, por ejemplo, aquellos otorgados mediante tarjetas de crédito o créditos rotativos, entre otros, en los cuales no se conoce exactamente el tiempo en el cual el cliente pagará su deuda.
Recuperación de cartera dado un incumplimiento en el pago: dificulta su predicción, pues depende de la garantía de pago dada por el cliente a la entidad financiera, la cual puede minimizar el riesgo de no pago

Fuente: Saavedra García, M. L. y Saavedra García, M. J. (2010).

##### 3.1.2 La Superintendencia Financiera de Colombia

Esta entidad fue creada en 1923 por la Ley 45 (Colombia, Congreso de la República, 1923), como un organismo adscrito al Ministerio de Hacienda y Crédito Público, con el fin de supervisar toda actividad en el sistema financiero, bajo el nombre de Sección Bancaria. Posteriormente, en 1979, por la Ley 32 (Colombia, Congreso de la República, 1979), se creó la Superintendencia de Valores, como encargada de estimular, organizar y regular el mercado público de valores.

Debido al crecimiento del sistema financiero, y en razón de que contaba con dos supervisores separados con tareas distintas, el Gobierno evaluó las gestiones llevadas por ambas

entidades y consideró su unión. El resultado fue el surgimiento de la Superintendencia Financiera de Colombia –la Superfinanciera– en 2005 por la Ley 964 (Colombia, Congreso de la República, 2005) y el Decreto 4327 (Colombia, Presidencia de la República, 2005), trasladándole así las funciones, los objetivos, las entidades vigiladas y las facultades asignadas. El objetivo de la Superfinanciera es preservar la estabilidad, la seguridad y la confianza en el sistema financiero colombiano al ejercer actividades de vigilancia, inspección y control sobre las personas naturales y jurídicas que realicen actividades de seguros, bursátiles o financieras, permitiéndole así organizar, desarrollar y promover el mercado de valores del país y la protección de los asegurados, los inversionistas y los ahorradores.

### 3.1.3 El Sistema de administración de riesgo de crédito (SARC) y su Capítulo II

El Comité de Basilea (Economipedia, s. f.) recomienda a las entidades financieras realizar análisis financieros de sus clientes para conocer su capacidad para cumplir o no con sus obligaciones de crédito, usando variables tradicionales como deuda, rendimientos o iliquidez, entre otras; no obstante, cabe resaltar que mientras más variables –relacionadas– se puedan involucrar, más fácil será clasificar a un conjunto de población frente a las decisiones de crédito.

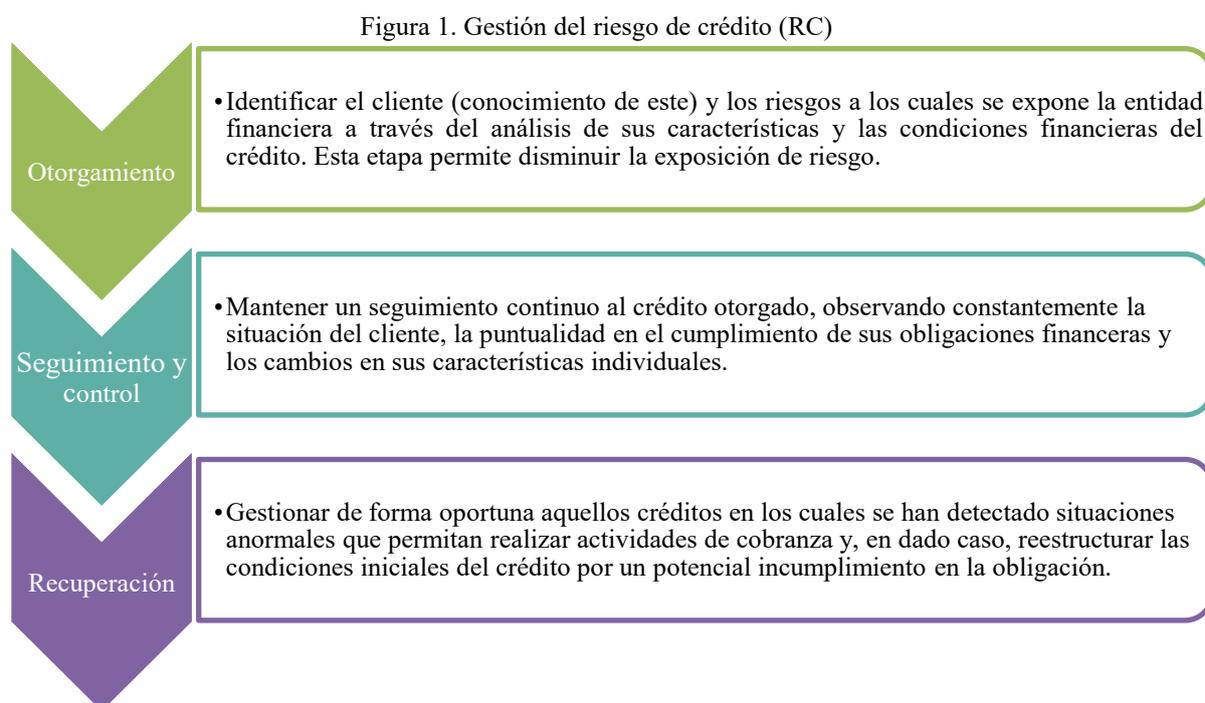
En el caso de Colombia, la Superfinanciera es el ente gubernamental encargado de llevar a cabo las funciones de control y vigilancia y de supervisar el sistema financiero y bursátil del país. Con este propósito, la Superfinanciera expidió la *Circular Básica Contable y Financiera* –Circular externa 100– (Colombia, Superintendencia Financiera, 1995), que reúne los requerimientos de información que las entidades le deben reportar y las instrucciones en materia contable y financiera vigentes actualmente que no están expresamente regladas en el Plan Único de Cuentas (Plan general contable, s. f.). En particular para este trabajo se destaca el Capítulo II, “Reglas relativas a la gestión de riesgo crediticio” (Colombia, Superintendencia Financiera, 1995), en el que se señalan los criterios generales, los principios y los parámetros mínimos que las entidades vigiladas deben observar para el diseño, el desarrollo y la aplicación del SARC; allí se define el RC como “la posibilidad de que una entidad incurra en pérdidas y se disminuya el valor de sus activos como consecuencia de que un deudor o contraparte incumpla sus obligaciones” (1995: 2). Por consiguiente, la administración de riesgos se constituye en una necesidad primordial para las entidades financieras, que día a día se ven expuestas a incurrir en pérdidas por el no pago de las obligaciones por parte de sus clientes, en razón de que les permite evitar arriesgar su estabilidad, solvencia y liquidez.

El SARC está compuesto por los siguientes elementos: a) Políticas de administración del RC; b) Procesos de administración del RC; c) Modelos internos o de referencia para la estimación o la cuantificación de las pérdidas esperadas; d) Sistema de provisiones para cubrir el RC; y e) Procesos de control interno (Colombia, Superintendencia Financiera, 1995: 2).

Como parte de las estrategias que las entidades financieras han implementado para la prevención del RC, se hace necesario establecer, junto a su perfil de riesgo, el perfil de aquellos que serán sujetos de crédito. En este caso hay un segmento que no tiene una estabilidad económica alta y que automáticamente supone un mayor riesgo de pérdida esperada. Así, dentro de la etapa de seguimiento, es importante contar con esquemas adecuados de caracterización y/o perfilamiento de los clientes, como principio de identificación de los riesgos, con el propósito de prevenir las pérdidas potenciales en las que las entidades podrían incurrir. Por ello es importante en este tipo de análisis poder incluir o considerar los diversos impactos de las

variables internas, la composición del portafolio, los criterios de calificación, las variables macroeconómicas y sectoriales, y las características históricas de la cartera que se va a analizar. En relación con las variables macroeconómicas, existen canales de transmisión entre cómo evoluciona este contexto y su relación con el no pago de obligaciones de los deudores, que influyen naturalmente sobre las decisiones de incumplimiento por su parte. Lo anterior se deriva de la insolvencia que afrontaría el deudor a raíz de un *shock* a nivel macro o microeconómico (Mermelstein, 2006).

El SARC debe contar con sistemas de seguimiento y control y definir las responsabilidades y funciones de cada una de las personas y organismos involucrados en el proceso de administración del riesgo de manera clara y expresa. *Grosso modo*, la gestión del crédito se muestra en las siguientes etapas [Figura 1]:



Fuente: elaboración de las autoras a partir de Colombia, Superintendencia Financiera (1995).

El cumplimiento de estas etapas se convierte en un ciclo en el cual, una vez finalizado el proceso de cualquier cliente relacionado con el seguimiento y control y el de recuperación, en caso de que sea necesario, se generan nuevas recomendaciones para la etapa de otorgamiento de créditos en procesos futuros. Paralelamente, las entidades deben priorizar la gestión del RC dando cumplimiento a las exigencias propuestas en el Capítulo II de la *Circular externa básica contable* (Colombia, Superintendencia Financiera, 1995), a través de la implementación del SARC (Rankia, s. f.).

En este orden de ideas, los recursos provenientes de las modalidades de captación deben ser administrados con suma prudencia, buscando al mismo tiempo mantener el apetito de riesgo definido por las entidades. Asimismo, las políticas y criterios establecidos para las etapas del riesgo (identificación, medición, control y evaluación) y la administración de riesgos deben estar en conformidad con las normas vigentes expedidas por la Superfinanciera (Colombia, Superintendencia Financiera, 1995).

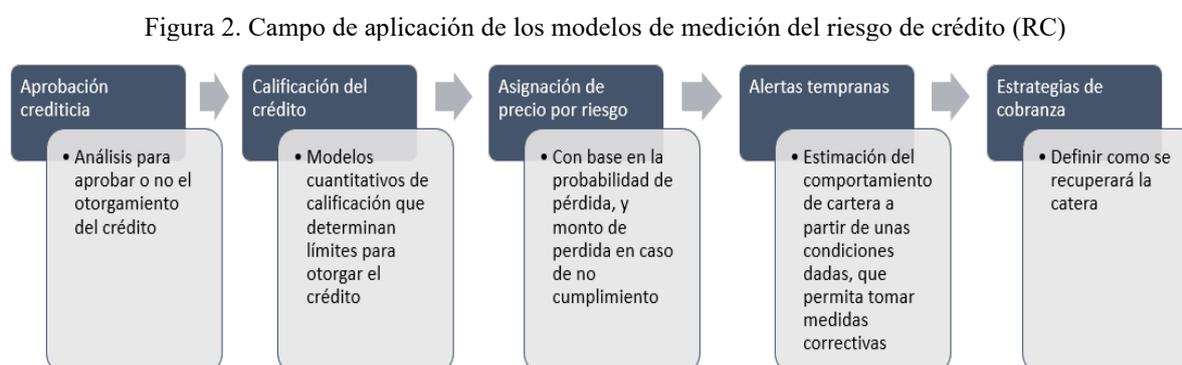
De manera general, el Comité de Supervisión de Basilea (Economipedia, s. f.) plantea que un modelo de RC cubre las políticas, los procedimientos y las prácticas usadas por las entidades financieras para estimar la función de distribución de una cartera de activos financieros expuestos al RC. En ese sentido, la finalidad de un sistema de medición del RC es proporcionar la o las medidas que le permitan a una entidad financiera cuantificar el RC de los activos que tiene en su estructura económica o de cualquiera de sus derivados.

### 3.2. Modelos de medición, seguimiento y control

En este punto se hace necesario anotar la diferencia entre los modelos orientados a la *medición del RC*, donde lo que se busca es, a través de análisis cuantitativos de riesgo, determinar las posibles pérdidas que se esperan en relación con el otorgamiento de un crédito, cuantificando los montos que los clientes posiblemente no pagarán y teniendo como base el nivel de riesgo asociado; y, por otra parte, aquellos que se orientan al *seguimiento y control*, que son principalmente de tipo cualitativo y previsorio, donde se trata de evitar que un cliente incumpla sus obligaciones y se concrete la pérdida calculada mediante el modelo de medición. Para efectos del presente trabajo, el modelo que se va a desarrollar es del tipo de seguimiento y control.

#### 3.2.1 Modelos de medición del riesgo de crédito (RC)

La aplicación de los modelos para medición del RC se realiza de modo general, como se muestra en la Figura 2.



Fuente: elaboración de las autoras a partir de Elizondo y López Romero (1999).

La Tabla 2 muestra los diferentes tipos de modelos y una descripción general de algunos de ellos. Cabe aclarar que, dado que los modelos de seguimiento y control se desarrollan internamente en las entidades financieras y no son de obligatoria divulgación, su planteamiento está sujeto a la discrecionalidad de su experiencia y necesidades, lo cual implica el uso de modelos menos estandarizados.

Tabla 2. Modelos de medición del riesgo de crédito (RC)  
Tipos de modelos de medición del RC

Tipos de modelos de medición del RC	
Modelos basados en el factor humano	a) Aquellos que se basan únicamente en el capital humano: el modelo de las cinco C. b) Aquellos que se basan en el factor humano en combinación con las normas.
Modelos estadísticos	a) Univariantes: análisis de tendencias; análisis de ratios financieros; test de clasificación dicotómica; análisis de ratios de probabilidad; medidas de descomposición del balance de situación. b) Multivariantes: análisis discriminante; modelos de regresión (probabilidad lineal, logístico y probabilístico); supervivencia.
Modelos de aprendizaje de máquinas	a) Modelos de inteligencia artificial: de adquisición de conocimiento; de aprendizaje, b) Modelos de redes neuronales artificiales, c) Modelos de árboles de decisión: binarios; no binarios.

Fuente: elaboración de las autoras a partir de Trigo Martínez (2009).

### *Modelo de las cinco C*

El modelo de las cinco C<sup>3</sup> es un modelo tradicional que basa su decisión en los analistas de crédito, que, a partir de la revisión de cinco elementos y su amplia experiencia, proceden a evaluarlos y a determinar si otorgan o no el crédito al cliente. Los elementos que se tienen en cuenta para este análisis son los siguientes: a) carácter: relacionado con la reputación del cliente y su voluntad de pago; b) capital: mide el capital total del cliente y el de endeudamiento; c) capacidad: verifica la capacidad de pago del cliente en relación con sus ingresos y ganancias; d) colateral: en caso de que se genere incumplimiento de pago (*default*), la entidad financiera tiene derecho sobre la garantía del crédito; e) ciclo económico: relacionado con la exposición crediticia en sectores económicos que se puedan ver afectados por los ciclos económicos (Saunders y Allen, 1999).

Este modelo presenta dos inconvenientes importantes, puesto que se presta para subjetividad por parte de los analistas frente a los diferentes tipos de clientes: la primera, no es consistente y no es objetivo, factores que lo han ido desplazando por nuevas metodologías que reducen los costos –menor número de analistas–; y la segunda, brindan mayor homogeneidad en el análisis de crédito ante una mayor cantidad de clientes.

### *Modelo KMV Moody's, Monitor de crédito*

El KMV Moody's es un modelo de diversificación basado en correlaciones del mercado de activos, buscando establecer la probabilidad de no cumplimiento entre pasivos y activos. Esta probabilidad se da en función de la estructura de capital, la volatilidad de los retornos y el valor de activos, los cuales, al ser indicadores sensibles, hacen que sea una buena opción en los países emergentes. El modelo busca relacionar la distancia al *default* con la probabilidad de que se produzca, y basa su metodología en información del mercado de valores, los cuales deben representar el valor real de las entidades (Saavedra García, M. L. y Saavedra García, M. J., 2010).

### *Modelo Credit Metrics de RiskMetrics Group y J. P. Morgan & Co.*

Este modelo se desarrolla en cinco etapas: 1) establece un esquema de calificaciones y una matriz a partir de información obtenida por las entidades calificadoras; 2) determina un período de tiempo, que generalmente es de un año; 3) propone un modelo de valoración; 4) analiza

<sup>3</sup> Character, capacity, capital, collateral, conditions.

variaciones en la cartera de crédito; 5) determina el no cumplimiento a partir del momento en que el valor de los activos es menor que el valor nominal del crédito.

Con este modelo se incorpora la diversificación –usada en la teoría de portafolios– a la medición del RC; no obstante, es necesario tener en cuenta que los mercados accionarios y de crédito tienen comportamientos diferentes (RiskMetrics Group y J. P. Morgan & Co., 1997/2007).

#### *Modelo Z-Score*

El Z-Score, que hace parte de los modelos paramétricos, tiene como base el análisis discriminante que permite obtener una combinación lineal de variables que posibiliten la clasificación por grupos, donde se maximice la varianza entre ellos y se minimice dentro de ellos. Este modelo, desarrollado por Edward Altman en 1968, se generó para predecir la quiebra o la no quiebra de las empresas a partir del análisis de información financiera por año, tamaño e industria a la que pertenecen (Hernández Ramírez, 2014).

En general, los modelos de *scoring* han tenido gran aceptación como parte de las actividades relacionadas al otorgamiento de créditos; además son útiles en las siguientes actividades: a) seguimiento y control, donde permiten identificar a clientes que puedan ser fidelizados por las entidades en vista de sus características y comportamiento, y advertir frente aquellos deudores que empiezan a mostrar señales de alerta y que requieran atención para asegurar que no lleguen a incumplimiento; y b) en actividades de recuperación, puesto que permiten un uso más eficiente de los recursos orientados a la cobranza, en términos de ejercer mayor presión entre quienes tengan el potencial de cumplir con sus obligaciones o en aquellos que finalmente se perfilan para ejecutar castigo de cartera o puedan ser objeto de venta de esta (Henaó Jassan, 2013).

#### *Sistemas expertos*

Estos sistemas, que hacen parte de las primeras técnicas que aplican inteligencia artificial para el desarrollo de sistemas de análisis del RC, integran la experiencia del conocimiento basado en hechos, así como los de tipo procedimental o de actuación, lo que les permite razonar de forma similar a expertos humanos. Los sistemas expertos usados habitualmente se basan en reglas que asumen premisas del tipo “si, entonces” (*if, then*). También se pueden encontrar sistemas basados en casos resueltos mediante similitudes con casos anteriores, buscando el que mayor semejanza tenga con la situación actual; al resolver los nuevos casos, estos pasan a ser parte de la base de casos resueltos (De Andrés Suárez, 2000).

#### *Árboles de decisión*

En general, los árboles de decisión permiten la segmentación, estratificación, predicción, reducción de datos y filtración de variables, la identificación de interacciones y la conversión de variables continuas en discretas (Berlanga Silvente, Rubio Hurtado y Vilà Baños, 2013).

Esta técnica no paramétrica facilita la identificación de clientes como buenos o malos a partir de reglas de clasificación que permiten llegar a una categorización final, dividiendo un conjunto de observaciones o datos en conjuntos que no tienen ningún elemento en común con los demás conjuntos (disjuntos) y aumentando de esta manera su homogeneidad. Estos últimos se dividen a su vez en nuevos subconjuntos, repitiendo el proceso tantas veces como sea

necesario para encontrar resultados aceptables. Los árboles de decisión son un método útil para seleccionar variables significativas, pues su resultado muestra aquellas variables independientes que presentan mayor relación con la variable dependiente, ya sea de forma inversa o directa; adicionalmente, cada rama generada desde el nodo raíz hasta el terminal puede entenderse como una regla de decisión (Breiman, Friedman, Stone y Olshen, 1984). Su ventaja más representativa es que aporta un mayor conocimiento de la relación causa/efecto que tienen las variables involucradas, apoyando así la selección de parámetros más relevantes para la determinación de clientes buenos y malos.

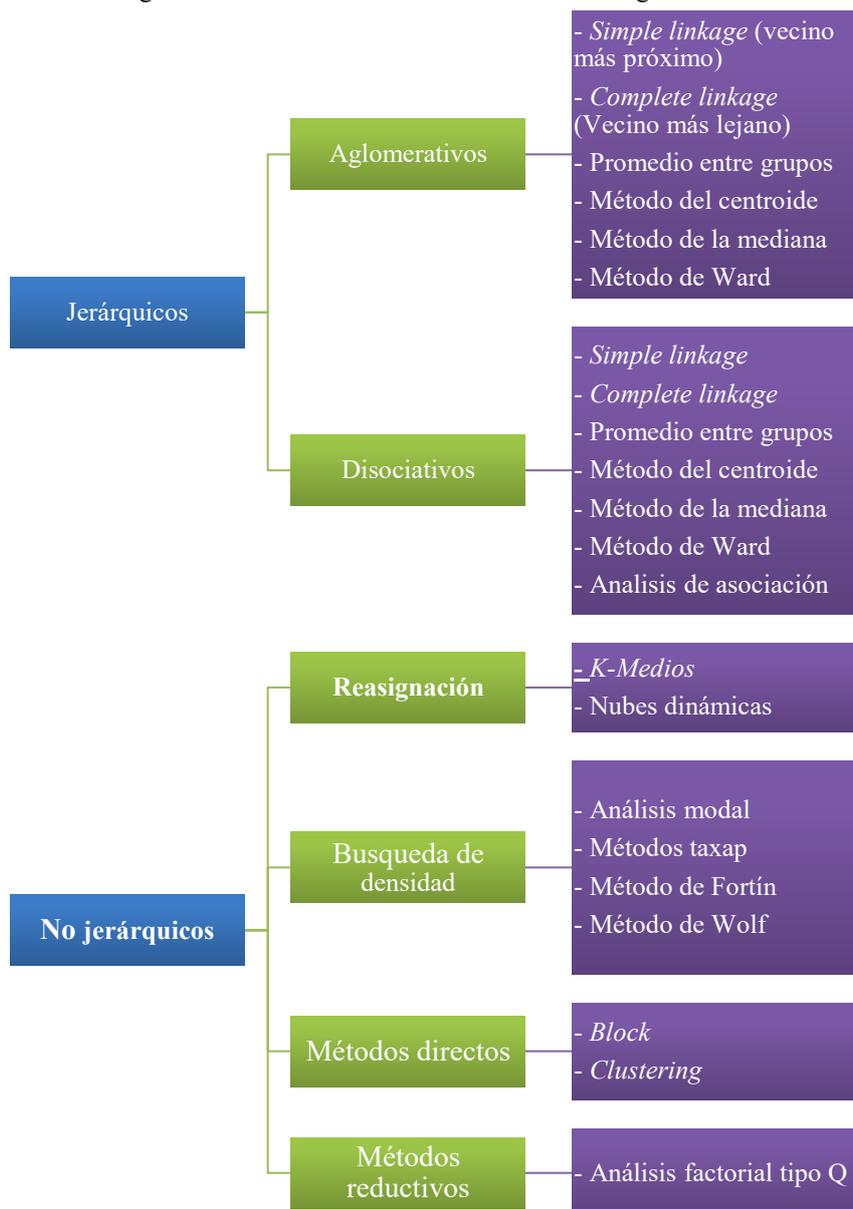
Los métodos de división de los árboles de decisión son los siguientes: CHAID (*Chi-square automatic interaction detector*; detector automático de Chi-cuadrado de interacción), CHAID exhaustivo, CRT (*classification and regression trees*; árboles de clasificación y regresión) y QUEST (*quick, unbiased, efficient, statistical tree*; árbol rápido, imparcial, eficiente y estadístico). El presente trabajo se evaluará con el método CHAID, propuesto por Kass (1980), que es particularmente útil cuando se trata de estudios exploratorios y la variable dependiente es de respuesta binaria; adicionalmente, este método, al ser de segmentación, permite que los resultados sean interpretados de manera más sencilla; no obstante, dichos resultados se soportan en análisis estadísticos, lo cual hace que sean técnicas robustas.

### *Modelos clúster*

También conocidos como “análisis de conglomerados”, hacen parte de las técnicas multivariantes en las cuales se busca agrupar elementos o variables para lograr la máxima homogeneidad dentro de un grupo y la máxima diferencia entre grupos diferentes. Por la esencia de los datos que los identifican, estos modelos sitúan individuos en grupos homogéneos (clústeres), de manera que los que tengan similitudes sean asignados a un mismo clúster y los disímiles sean ubicados en clústeres diferentes. Los modelos clúster difieren del análisis discriminante, en el que, dado un objetivo, se hace necesario determinar previamente los grupos (De la Fuente Fernández, 2011). Al aplicar la técnica de clústeres en el desarrollo del modelo presentado en este trabajo, se logra clasificar al cliente en un grupo específico según determinadas variables; cabe anotar que esto no implica establecer qué variables tienen mayor o menor peso frente a otras.

La Figura 3 muestra la división de los análisis de conglomerados.

Figura 3. División del modelo de análisis de conglomerados



Fuente: elaboración de las autoras a partir de De la Fuente Hernández (2011).

### 3.2.2 Modelos de seguimiento y control del riesgo de crédito (RC)

En un sentido básico, un modelo de seguimiento y control permite prever el riesgo de no pago en el que incurriría un cliente frente al otorgamiento de un crédito. Este modelo, esencialmente de tipo cualitativo, busca prevenir que un cliente llegue al incumplimiento en el pago y se materialice la pérdida para la entidad prestamista; así facilita su toma de decisiones sobre los criterios de alerta y los métodos de cobranza (Caicedo Cerezo, Claramunt Bielsa y Casanovas Ramón, 2011).

El seguimiento al RC se puede analizar tanto desde el punto de vista del riesgo del cliente como desde la cartera de crédito. El primero, que se da de forma regular, está basado en la revisión de sus condiciones macroeconómicas, el cumplimiento del contrato y su situación financiera; esto les permite a las entidades advertir posibles cambios en su capacidad de pago y puede evaluarse a partir de indicadores como disciplina de pago, rentabilidad, rotación y

liquidez, y, dependiendo de la solvencia del cliente, se pueden realizar seguimientos mensuales, trimestrales o incluso anuales. En caso de identificar que la capacidad de pago del cliente presenta variaciones, las entidades pueden iniciar acciones que les permitan evitar el deterioro del este mediante estrategias como una negociación de condiciones adicionales, el uso de instrumentos de cobertura o la reestructuración de la deuda, entre otras. Por otra parte, el seguimiento desde la cartera de crédito está orientado a los tipos de créditos ofrecidos por las entidades (hipotecas, créditos de consumo, etc.) que les aportan diversificación de riesgo; de esta forma, genera alertas sobre los montos acumulados de crédito, las tendencias en las carteras de crédito, los límites de minimización del riesgo y la maximización de las ganancias (Spuchl'áková, Valášková y Adamko, 2015).

Cabe destacar que cada entidad financiera se encarga de definir su propio modelo de seguimiento y control, claro está, con base en los lineamientos establecidos por la entidad reguladora –la Superfinanciera en el caso colombiano–; sin embargo, este hecho impide comparar los modelos entre ellas, puesto que no están obligadas a hacer públicos sus desarrollos internos; adicionalmente, al ser de discreción de cada una, les permite el uso de modelos menos estandarizados basados en su experiencia y necesidades. En este contexto, los modelos de seguimiento al RC dependen de las variables que sean más relevantes o consideradas de mayor riesgo por dichas entidades, según el segmento que deseen analizar (Benjumea Velásquez, 2013). Esta herramienta de seguimiento es especialmente útil para el caso de las personas naturales, dado que reduce la subjetividad en los análisis crediticios y permite usar la información histórica de los solicitantes apoyándose en la experiencia y criterio de los analistas y los gestores de crédito (Saavedra García M. L. y Saavedra García, M. J., 2010). Las principales variables sociodemográficas cualitativas y cuantitativas que pueden ser usadas para este análisis se muestran en la Tabla 3.

Tabla 3. Variables de mayor uso en los modelos de medición de riesgo

Tipo de variable	Variable
Cualitativa	Estado civil
	Tipo de vivienda
	Profesión
	Educación
	Sexo
	Estado civil
Cuantitativa	Tiempo vinculación en el empleo actual
	Tiempo vinculación en el empleo anterior
	Ingresos
	Capacidad de pago declarada
	Capacidad de pago calculada
	Capacidad de pago declarada/ingresos
	Edad
	Número de dependientes económicamente
	Tiempo de habitación en la vivienda actual
	Tiempo de habitación en la vivienda anterior
	Valor de los activos
	Número de veces vencido > 30 días en mora
	Número máximo de días vencido
	Monto refinanciado por mes
	Monto refinanciado al año
Cantidad de veces refinanciado	

Fuente: elaboración de las autoras.

Adicional a estas variables que identifican al cliente, se validan aspectos relacionados con el crédito en específico. La valoración de cada cliente también debe acompañarse con información de terceros que permitan analizar otro tipo de comportamientos por parte del solicitante: antecedentes bancarios, reportes en las centrales de riesgo o registro de propiedades; estos datos, además de apoyar el perfilamiento de la solicitud, brindarán información para futuras labores comerciales asociadas a otros servicios financieros ofrecidos.

Según las políticas establecidas en la *Circular Básica Contable y Financiera*, las principales actividades en el seguimiento y control del RC que debe realizar una entidad financiera agrupan los siguientes pasos:

- Diseñar un modelo de seguimiento y control en conformidad con sus políticas y sus actividades de crédito; dado que las condiciones de mercado son cambiantes, dicho modelo debe ser evaluado por lo menos dos veces al año, a fin de hacer los ajustes pertinentes.
- Evaluar de forma periódica a sus clientes al menos dos veces al año, de forma tal que se revise su situación de forma independiente previo a la revisión del modelo.
- Solicitar al cliente la actualización de su situación financiera al menos una vez al año.
- Realizar informes consolidados mensuales y trimestrales del comportamiento de la cartera, que deben ser entregados a la Superfinanciera.
- Determinar aquellos clientes que presenten anomalías en su comportamiento y/o cambios importantes en su situación financiera.
- Realizar comités de seguimiento trimestrales basados en los resultados del modelo de seguimiento y control para evaluar la evolución de cada cliente.
- Realizar las respectivas validaciones para el caso de clientes que sean objeto de reporte a las centrales de riesgo y evaluar aquellos casos que deban ser reestructurados (Colombia, Superintendencia Financiera, 1995).

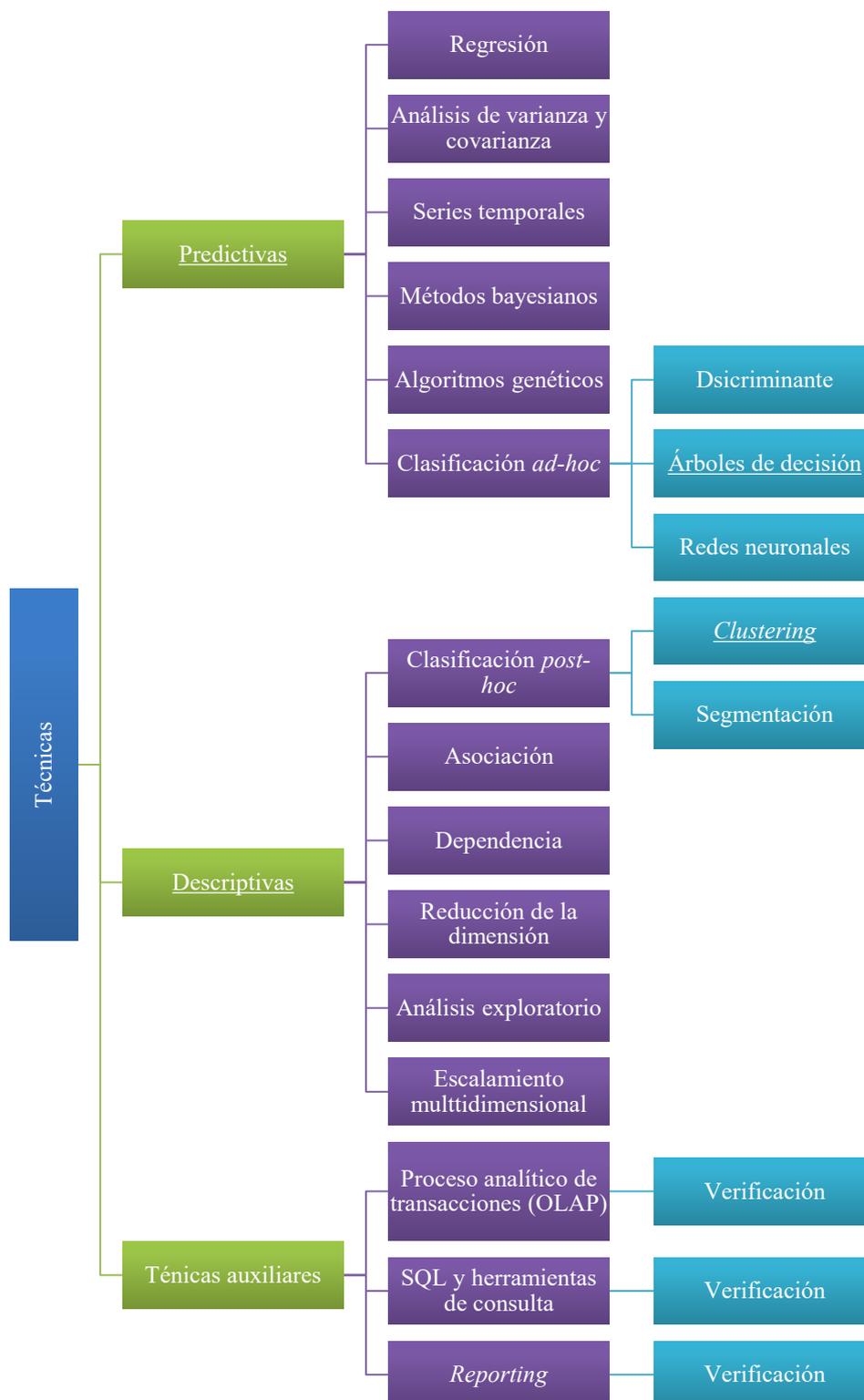
Todo lo anterior debe ser revisado en conjunto con las políticas de crédito de las entidades financieras.

### 3.3 Técnicas de minería de datos (*data mining*)

Las técnicas de minería de datos (*data mining*) provienen de la estadística y la inteligencia artificial, donde, a través de algoritmos aplicados a un conjunto de datos, se obtienen resultados. Las más reconocidas son las siguientes: regresión lineal, redes neuronales, modelos estadísticos, árboles de decisión, agrupamiento o *clustering* y reglas de asociación (Berlanga Silvente, Rubio Hurtado y Vilà Baños, 2013). Las técnicas predictivas y descriptivas aportan al descubrimiento, mientras que las auxiliares sirven para los procesos de verificación (Pérez López y Santín González, 2007). Ambas asisten como apoyo en el desarrollo del presente trabajo mediante el uso del *clustering* y de los árboles de decisión.

La Figura 4 muestra la clasificación de las técnicas de minería de datos.

Figura 4. Clasificación de las técnicas de minería de datos (*data mining*)



Fuente: elaboración de las autoras a partir de Pérez López y Santín González (2008).

### 3.4 Contexto del trabajador independiente en Colombia

Varias son las formas como se ha intentado medir la informalidad laboral, pero su definición aún no es clara. Lo que sí es evidente es que esta es un fenómeno que ha ganado importancia en la economía tanto en la generación de empleo como en la producción. La trascendencia del sector informal en la economía hace necesario poder entenderlo y saber sus determinantes.

Existen dos enfoques importantes que han tratado de definir los determinantes asociados a la informalidad: a) el enfoque estructuralista, en el que se asocia la informalidad con la posición ocupacional y el tamaño de la planta a partir de planteamientos empíricos que tienen en cuenta la productividad, la marginalidad laboral y el capital de trabajo; sin embargo, no desconoce el hecho de que la informalidad sea el resultado de las variables estructurales de la economía; b) el enfoque institucionalista, que la asocia con la evasión de normas, es decir, el no pago de los impuestos ni de la seguridad social, y además la relaciona con las dificultades que existen para la formación y el funcionamiento de empresas bajo los términos de la legalidad y el papel ineficiente del Estado (García Cruz, 2011).

En su publicación del 11 y 12 de julio de 2018, basada en datos de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), el Sistema Económico Latinoamericano y del Caribe, SELA, informó que Colombia, con el 61,3 % del total de la fuerza laboral, era el país del mundo con mayor cantidad de empleados autónomos, cifra aún mayor que la de 2016, cuando el porcentaje era del 51,3 %, siendo también en este momento el más alto del planeta, todo ello bajo el concepto de *trabajo autónomo*, en el cual no se generan las debidas garantías laborales (Sistema Económico Latinoamericano y del Caribe, SELA, 2018).

Según la Resolución 5858 de 2016 (Colombia, Ministerio de Salud y Protección Social, 2016), en Colombia existen tres principales tipos de trabajadores independientes, que se distinguen según el tipo de afiliación y el aporte que hacen al sistema de riesgos laborales o a las cajas de compensación. El primero es el *trabajador independiente*, que debe aportar obligatoriamente a los sistemas generales de salud y pensión como persona natural que realiza una actividad económica o presta sus servicios de manera personal y por su cuenta. El segundo es el *trabajador independiente con contrato de prestación de servicios superior a un mes*, una persona natural vinculada a través de un contrato en forma de prestación de servicios a cualquier entidad o institución de carácter público o privada; dicho contrato puede ser comercial, administrativo o civil. Y el tercero es el *trabajador independiente voluntario*, que aporta y se afilia al sistema general de riesgos por decisión propia; según la resolución, “es quien aportará por períodos mensuales completos y pagará mes vencido a los sistemas generales de salud, pensión y riesgos laborales” (2016).

Para el mercado laboral de Colombia, la informalidad es quizá una de las mayores problemáticas, teniendo en cuenta que las políticas laborales en el país han fomentado condiciones de contratación que, en su mayoría, no brindan estabilidad al trabajador para lograr el cumplimiento de sus objetivos y metas a nivel personal y familiar –adquisición de vivienda, vehículo o educación, entre otros.

Según la *Gran encuesta integrada de hogares – GEIH* del mercado laboral en Colombia, con cifras a noviembre de 2019, los indicadores de empleo muestran una tasa de ocupación (TO) nacional decreciente y un deterioro de la demanda laboral, principalmente en los sectores de menor calidad, es decir, los de los trabajadores no asalariados e informales. La TO asalariada

urbana se mantiene estable, mientras que la no asalariada continúa acumulando variaciones negativas (Colombia, Archivo Nacional de Datos, ANDA, 2019) [Figura 5].

Figura 5. Colombia. Tasa de ocupación asalariada y no asalariada – 23 ciudades (noviembre 2009-noviembre 2019)



Fuente: Colombia, Archivo Nacional de Datos, ANDA (2019).

La tasa de ocupación (TO) formal urbana (Colombia, Archivo Nacional de Datos, ANDA, 2019) presentó un comportamiento relativamente estable con respecto a períodos anteriores, mientras que su contraparte, la TO informal, mostró importantes variaciones negativas a lo largo de los últimos meses en ese mismo dominio [Figura 6].

Figura 6. Colombia. Tasa de ocupación formal e informal – 23 ciudades (octubre 2009-octubre 2019)



Fuente: Colombia, Archivo Nacional de Datos, ANDA (2019).

El concepto de *actividad económica informal* o *trabajo informal* es aplicado a todo aquel individuo que en su labor económica no posee las formalidades laborales contenidas dentro del código laboral de cada país. En Colombia, por ejemplo, un trabajador informal puede ser considerado como aquel para el cual su labor y su trabajo no se ven remunerados exclusivamente en beneficio económico, sino también en seguridad social, pensión, ahorro, estabilidad laboral e ingreso certificable. Para 2019, el total de la población ocupada por cuenta propia fue de aproximadamente 9.450.000 personas (Colombia, Archivo Nacional de Datos, ANDA, 2019).

Una de las características del mercado laboral de los países en desarrollo consiste en la existencia del sector informal, dada la poca capacidad del sector actual para absorber una oferta de trabajo en constante crecimiento. Desde esta perspectiva, el sector informal constituye hoy día la principal fuente de empleo e ingreso para millones de personas, que están obligadas a desarrollar actividades de pequeña escala, baja productividad e intensivas en mano de obra. La informalidad en el mercado laboral se manifiesta en las condiciones precarias de empleo, tales como la escasa posibilidad de acceso al sistema de seguridad social, poca estabilidad, la ausencia de un contrato escrito, condiciones físicas inadecuadas y bajos ingresos.

Al cierre de diciembre de 2019, la cartera de clientes independientes de la Entidad correspondía al 29,36 % –la segunda luego de la de los asalariados– del total de la cartera de consumo e hipotecaria (COP 3.015.478 millones), con un indicador de cartera vencida (ICV) del 7,8 %, superior a los cierres de los dos años anteriores (diciembre de 2016-diciembre de 2018) y una participación del saldo vencido del 43,8 % [Tabla 4].

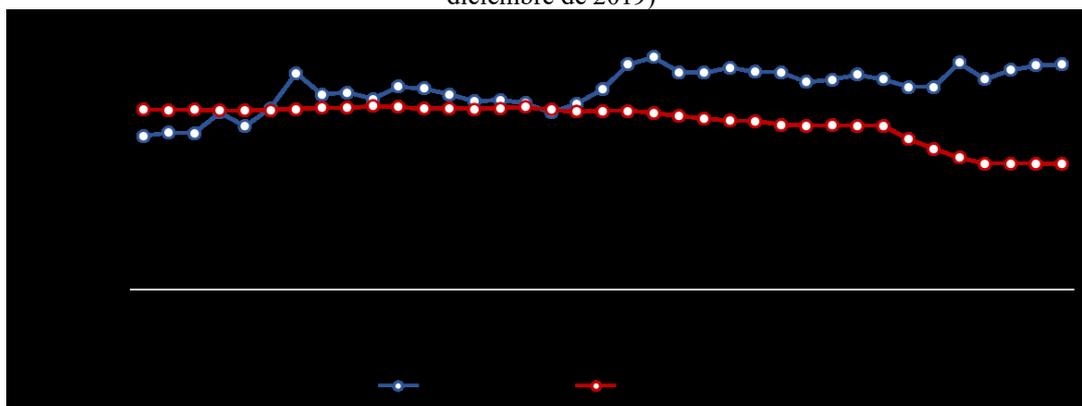
Tabla 4. La Entidad. Cartera de consumo e hipotecaria por actividad económica (diciembre de 2016-diciembre de 2019)

Actividad Económica	dic-19										dic-18		dic-17		dic-16	
	Cartera Bruta				Cartera Vencida				ICV		ICV		ICV			
ASALARIADO	72,722	71.46%	1,758,604	58.32%	4,454	62.03%	79,484	50.14%	6.1%	4.5%	5.9%	4.9%	6.0%	5.0%	6.7%	5.3%
INDEPENDIENTE	27,280	14.84%	885,315	29.36%	2,018	28.11%	69,417	43.79%	7.4%	7.8%	7.5%	7.5%	6.7%	6.4%	6.7%	5.8%
PENSIONADO	15,230	1.89%	362,711	12.03%	344	4.79%	9,195	5.80%	2.3%	2.5%	1.8%	2.1%	2.0%	2.3%	2.1%	1.9%
ESTUDIANTE	1,935	10.63%	4,394	0.15%	305	4.25%	209	0.13%	15.8%	4.8%	2.4%	4.6%	4.4%	4.7%	3.9%	3.3%
NO DEFINIDO	1,057	1.18%	4,455	0.15%	59	0.82%	212	0.13%	5.6%	4.8%	4.4%	6.7%	6.0%	7.7%	5.8%	7.9%
<b>Total</b>	<b>118,224</b>	<b>100%</b>	<b>3,015,478</b>	<b>100%</b>	<b>7,180</b>	<b>100%</b>	<b>158,518</b>	<b>100%</b>	<b>6.1%</b>	<b>5.3%</b>	<b>5.7%</b>	<b>5.4%</b>	<b>5.7%</b>	<b>5.2%</b>	<b>6.2%</b>	<b>5.2%</b>

Fuente: elaboración de las autoras a partir de datos de la entidad estudiada (archivo personal de las autoras).

En relación con su participación en los últimos tres años evaluados (diciembre de 2016-diciembre de 2019), el indicador de cartera vencida (ICV) del segmento mantiene una tendencia creciente tanto en el volumen de clientes como en el de cartera; por su parte, la participación de esta última presenta un comportamiento más estable, aunque decreciente entre mayo y septiembre de 2019 [Figura 7].

Figura 7. La Entidad. Evolución de la participación del perfil del cliente independiente (diciembre de 2016-diciembre de 2019)

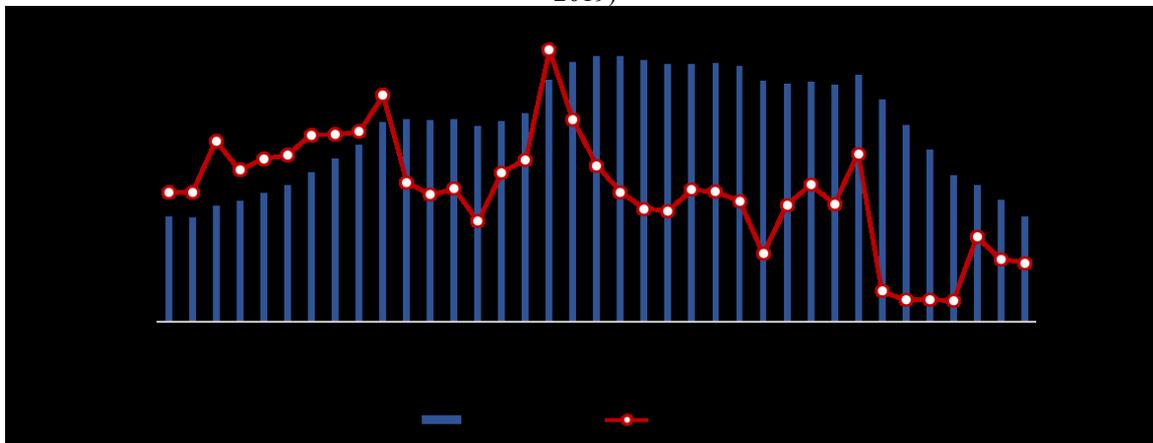


Fuente: elaboración de las autoras a partir de datos de la entidad estudiada (archivo personal de las autoras).

La Figura 8 y la Figura 9 muestran la evolución del saldo de cartera y de la cartera vencida a lo largo de los últimos tres años evaluados (diciembre de 2016-diciembre de 2019) para el perfil de clientes independientes de la Entidad. Al cierre de diciembre de 2019, ambos rubros presentaron decrecimientos respecto al mes anterior, con el segundo en mayor proporción.

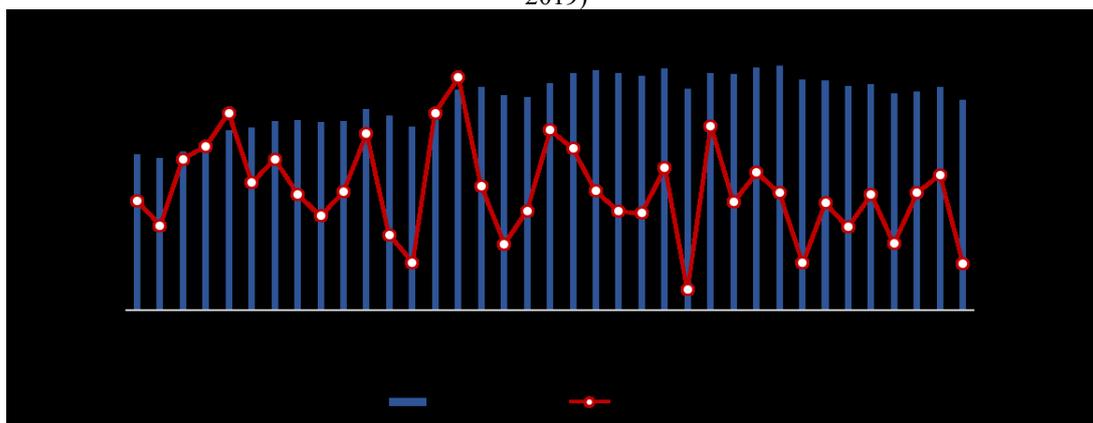
Usando estos datos, la Entidad quiere abordar este importante segmento de manera que pueda gestionar adecuadamente su riesgo y maximizar así el beneficio al minimizar los riesgos alineados a sus políticas de RC y al apetito de riesgo definido por ella.

Figura 8. La Entidad. Evolución del saldo de cartera del cliente independiente (diciembre de 2016-diciembre de 2019)



Fuente: elaboración de las autoras a partir de datos de la entidad estudiada (archivo personal de las autoras).

Figura 9. La Entidad. Evolución de la cartera vencida del cliente independiente (diciembre de 2016-diciembre de 2019)



Fuente: elaboración de las autoras a partir de datos de la entidad estudiada (archivo personal de las autoras).

## 4. Metodología

### 4.1 Metodología y técnica de recolección de datos

El presente trabajo aplica el método sintético, en razón de que integra la base de datos de la Entidad con su área de seguimiento y control del RC para el planteamiento y el desarrollo del modelo, a fin de que las variables agrupadas se puedan estudiar en conjunto. Paralelamente, tiene un enfoque mixto, es decir, cualitativo y cuantitativo, dado que mezcla el conocimiento y la experticia de dicha área con cálculos numéricos para medir las variables y hacer estadísticas que arrojen resultados que permitan construir una realidad por medio de pronósticos.

El modelo de seguimiento se aplicará a todos los deudores independientes con cartera activa en la Entidad, buscando robustecer la etapa de seguimiento y control del SARC en ella y empleando técnicas de minería de datos: por esta razón, se puede clasificar el método como un modelo híbrido, pues combina distintas técnicas –el *clustering* y los árboles de decisión– que se pueden analizar de forma separada o conjunta de agrupamiento y clasificación.

#### 4.1.1 Tipo de investigación

El modelo se desarrolló a través de una investigación descriptiva, al integrar diversas variables que arrojan un resultado para su respectivo análisis; por consiguiente, no es necesario explicar la relación entre ellas.

#### 4.1.2 Definición de los parámetros y las fuentes

Las fuentes de este trabajo son primarias y secundarias, en consideración a que se recurre a la Entidad para obtener los datos de primera mano. Cabe resaltar que la información entregada es de carácter confidencial; sin embargo, a fin de estudiar a profundidad alguna falencia o problemática que pudiera tener, se realiza un trabajo de recolección de información directa dando observaciones, mediciones y cálculos con los datos obtenidos.

Paralelamente se realiza la consulta de fuentes externas o secundarias que permitan servir de guía acerca del estado del arte del tema tanto en el ámbito local como el internacional, con el propósito de nutrir el marco teórico. Así, fueron consultados artículos, tesis de grado, informes, páginas electrónicas y otras investigaciones.

#### *Muestra y selección del periodo de estudio*

Para seleccionar el período de estudio se aplica el criterio de que la información más reciente del cliente prima sobre su comportamiento de pago; de este modo se toma la base de datos de los clientes independientes de la Entidad con cartera activa a corte de diciembre de 2019. En relación con la longitud del período de estudio, se utiliza la información de los últimos doce meses de comportamiento (mora) interno, sumado a variables cualitativas y cuantitativas al corte. Y en lo concerniente a la información externa frente al sector, se dispone de las alertas a diciembre de 2019.

### *Herramienta para la selección de variables y construcción del modelo*

En el desarrollo de esta etapa se evalúa la información hasta alcanzar la propuesta final del modelo. Para ello se utiliza la herramienta SPSS Modeler (International Business Machines Corporation, IBM, s. f.), una plataforma de análisis predictivo que impulsa el modelado y permite construir modelos predictivos precisos de forma rápida y sencilla. El SPSS Modeler puede considerarse como una aplicación avanzada en minería de datos que hace posible llevar a cabo toma de decisiones, estimaciones, predicciones y pronósticos. Entre las soluciones que ofrece esta herramienta se incluyen tecnologías de aprendizaje automático y modelado que se ajustan en función de los problemas que se pretendan resolver. Por ende, pone a disposición de los usuarios los siguientes módulos:

- Reglas de asociación: sirven para unir un conjunto de condiciones o requisitos con una conclusión.
- Modelos de agrupación o segmentación: centrados en identificar grupos de registros similares y etiquetar los registros al grupo que pertenecen. Son métodos de *clustering* que incluyen los modelos K-Medias, Kohonen, Bietápico, y TwoStep.
- Métodos predictivos de modelado: en ellos se incluyen las redes neuronales, los árboles de decisión y los modelos estadísticos como las series temporales, los logísticos y los lineales, entre otros.
- Modelos de *screening*: este módulo es empleado para localizar registros y campos que puedan ser de mayor interés en el modelado e identificar valores extraños que no se ajusten con los patrones conocidos. Entre ellos están los modelos de detección de anomalías y los de selección de características (International Business Machines Corporation, IBM, s. f.).

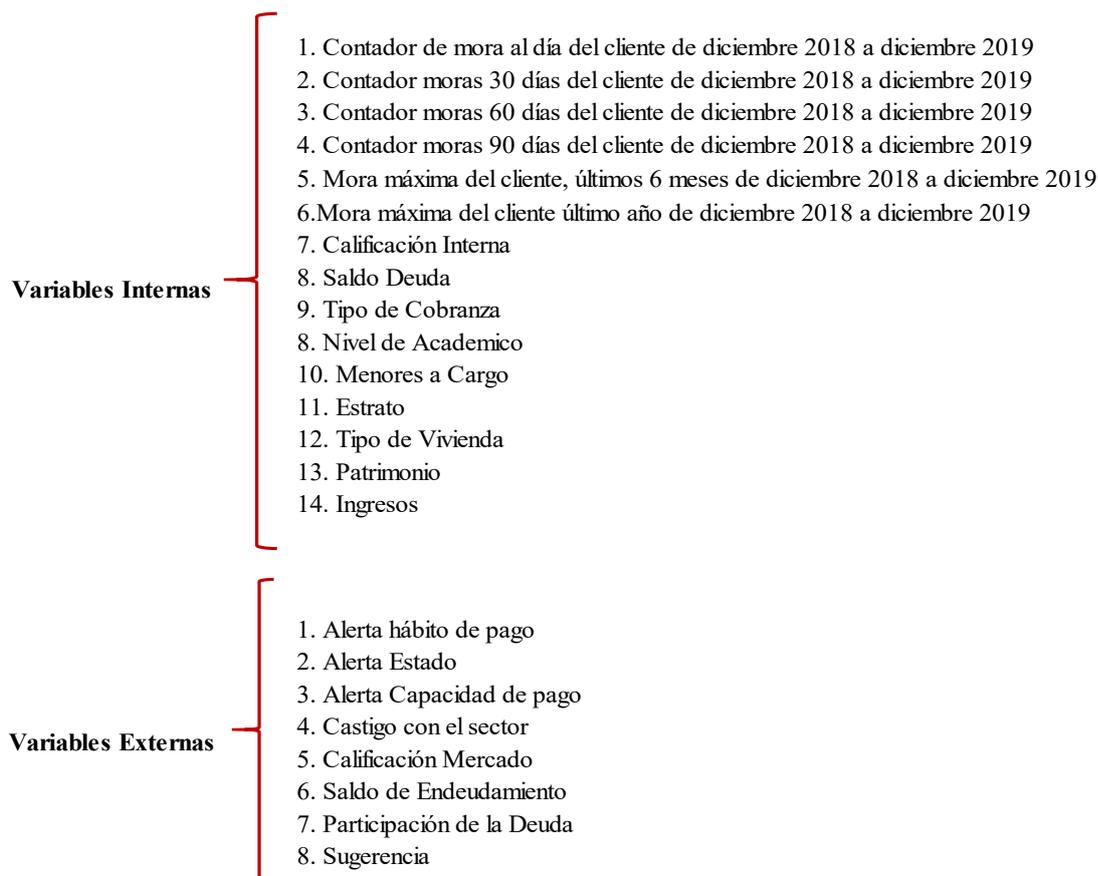
Teniendo en cuenta las bondades de la herramienta, se trae la información de la base de datos de la entidad y se construye la ruta a partir de dicha información, generando las condiciones iniciales para que la herramienta apoye la construcción y procesamiento del modelo.

### *Procesamiento de la información*

Incluye los siguientes pasos:

- a) Extracción de datos y calidad de la información. Inicialmente se realiza una minería de datos para garantizar su calidad y poder tener así resultados más precisos y eficaces. Siendo una de las autoras de este trabajo la analista de seguimiento y control en la gerencia de riesgos financieros de la Entidad, y de la mano con la jefe del área, se obtiene el acceso a su base de datos –bajo los acuerdos de confidencialidad entre las partes– para obtener la información de 27.280 clientes independientes al corte de diciembre de 2019 [Figura 10].

Figura 10. La Entidad. Variables internas y externas – base de datos



Fuente: elaboración de las autoras.

b) Propuesta de indicadores. Basados en las necesidades de la Entidad y en su visto bueno, se concretan los siguientes indicadores:

1) Comportamiento interno a corto plazo

Muestra el comportamiento que registró el cliente en los últimos seis meses [Tabla 5].<sup>4</sup>

Tabla 5. La Entidad. Comportamiento interno a corto plazo – rango días (últimos seis meses)

Rango [días]	Clasificación
0 – 30	Excelente
31 – 60	Aceptable
61 – 90	Regular
> 90	Malo

Fuente: elaboración de las autoras.

En el caso de aquellos clientes que en el mes de análisis no poseen comportamiento crediticio con la Entidad porque es su primer crédito, se clasifican como nuevos en el indicador y se les da una clasificación de Excelente.

<sup>4</sup> Cabe recordar que en esta y algunas de las tablas siguientes, el período de análisis cubre la información hasta diciembre de 2019.

## 2) Comportamiento interno a largo plazo

Busca identificar el número de veces que el cliente registra una mora en los últimos 12 meses, por lo que inicialmente se realiza la construcción de variables de conteo de moras parciales [Tabla 6].

Tabla 6. La Entidad. Comportamiento interno a largo plazo (12 meses)

Rango [días]	Veces	Clasificación
0 - 30	0	Excelente
	(0 - 2]	Excelente
	(2 - 5]	Aceptable
	(5 - 7]	Regular
	> 7	Malo
30 - 60	0	Excelente
	(0 - 2]	Excelente
	(2 - 5]	Aceptable
	(5 - 7]	Regular
	> 7	Malo
60 - 90	0	Excelente
	(0 - 2]	Excelente
	(2 - 5]	Aceptable
	(5 - 7]	Regular
	> 7	Malo
> 90	0	Excelente
	(0 - 2]	Excelente
	(2 - 5]	Aceptable
	(5 - 7]	Regular
	> 7	Malo

Fuente: elaboración de las autoras.

En el caso de aquellos clientes que en el mes de análisis no poseen comportamiento crediticio con la Entidad porque es su primer crédito, se clasifican como nuevos en el indicador y se les da una clasificación de Excelente.

## 3) Calificación interna a corto plazo

Busca identificar el nivel de riesgo que ha presentado el cliente con la Entidad en los últimos seis meses, tomando la máxima calificación en esta ventana de tiempo [Tabla 7].

Tabla 7. La Entidad. Calificación interna a corto plazo (últimos seis meses)

Calificación	Clasificación
A	Excelente
B	Aceptable
C	Regular
D	Malo
E	Malo

Fuente: elaboración de las autoras.

## 4) Calificación interna a largo plazo

Busca identificar el nivel de riesgo que ha presentado el cliente con la Entidad en los últimos 12 meses, tomando la máxima calificación en esta ventana de tiempo [Tabla 8].

Tabla 8. La Entidad. Calificación interna a largo plazo (últimos 12 meses)

Calificación	Puntaje
A	Excelente
B	Aceptable
C	Regular
D	Malo
E	Malo

Fuente: elaboración de las autoras.

### 5) Rango de porcentaje amortizado

Corresponde al porcentaje de desembolso que ya ha sido abonado a la Entidad y permite observar la forma en que el cliente reembolsa paulatinamente el capital prestado [Tabla 9].

$$1 - \frac{\text{Saldo}}{\text{Desembolso}}$$

Tabla 9. La Entidad. Rango del porcentaje amortizado

Rango amortizado [%]	Puntaje
[0 – 30]	Malo
(30 – 50]	Regular
(50 – 70]	Aceptable
> 70	Excelente

Fuente: elaboración de las autoras.

En el caso de aquellos clientes que en el mes de análisis no poseen comportamiento crediticio con la entidad porque es su primer crédito, se clasifican como nuevos en el indicador y se les da una clasificación de Excelente.

### 6) Tipo de cobranza

Busca identificar el estado del cliente respecto al sistema de cobro de la Entidad [Tabla 10].

Tabla 10. La Entidad. Tipo de cobranza

Tipo de cobranza	Clasificación
Administrativa	Excelente
Persuasiva	Aceptable
Pre-jurídica	Regular
Jurídica	Malo

Fuente: elaboración de las autoras.

### 7) Porcentaje participación de la deuda con el sector

Detalla por rango el porcentaje de participación de la deuda del cliente con la Entidad frente el resto del sector. Es decir, se puede identificar al corte cuánto debe el cliente a la Entidad de todo lo que tiene prestado en el sector [Tabla 11].

Tabla 11. La Entidad. Porcentaje de participación de la deuda con el sector

Rango de participación de la deuda [%]	Clasificación
[0 – 30]	Malo
(30 – 50]	Regular
(50 – 70]	Aceptable
> 70	Excelente

Fuente: elaboración de las autoras.

En el caso de aquellos clientes que en el mes de análisis no poseen comportamiento crediticio con el sector, se clasifican como nuevos en el indicador y se les da una clasificación de Excelente, pues se asume que apenas adquirieron la deuda con la Entidad y por eso no tienen reportadas otras deudas.

#### 8) Número de alertas en el sector

Busca identificar el comportamiento del cliente fuera de la Entidad respecto al hábito de pago, su estado y la capacidad de pago [Tabla 12].

Tabla 12. Número de alertas en el sector

Número de alertas	Clasificación
0	Excelente
1	Aceptable
2	Regular
3	Malo

Fuente: elaboración de las autoras.

#### 9) Calificación interna frente a calificación externa

Busca comparar la calificación actual que tiene el cliente internamente contra la calificación en el sector. La escala es la siguiente [Tabla 13]:

Tabla 13. Calificación interna de la Entidad frente a la calificación externa del sector

Calificación	Clasificación
Calificación interna peor a la del sector	Excelente
Calificación interna igual a la del sector	Aceptable
Calificación interna mejor que en la del sector en una escala	Regular
Calificación interna mejor que la del sector	Malo

Fuente: elaboración de las autoras.

#### 10) Porcentaje de participación del saldo por la calidad de la cartera con el sector

Detalla la distribución del saldo por la calidad de la cartera que presenta el cliente con el sector [Tabla 14].

Tabla 14. Porcentaje de participación del saldo de calidad con el sector

Rango de participación del saldo de la calidad de la cartera con el sector [ %]	Puntaje
[0 – 30]	Excelente
(30 – 50]	Aceptable
(50 – 70]	Regular
> 70	Malo

Fuente: elaboración de las autoras.

#### 11) Rango indicador de cartera vencida interno

Corresponde al porcentaje total de la cartera del cliente que se encuentra vencida (mora mayor a 30 días) y muestra la realidad sobre la exposición que tiene la Entidad ante el RC [Tabla 15].

$$\frac{\text{Total cartera vencida}}{\text{Total cartera}}$$

Tabla 15. La Entidad. Indicador de cartera vencida interno

Rango ICV [ %]	Clasificación
[0 – 15]	Excelente
(15 – 30]	Aceptable
(30 – 50]	Regular
> 50	Malo

Fuente: elaboración de las autoras.

## 12) Nivel académico

Muestra el nivel de estudios que tiene el cliente tomando como partida la información de la base de datos interna [Tabla 16].

Tabla 16. La Entidad. Nivel académico

Nivel académico	Clasificación
Primaria	Malo
Bachiller	Regular
Profesional	Excelente
Especialista	Excelente
Magíster	Excelente
Doctorado	Excelente
Técnico	Aceptable
Tecnólogo	Aceptable
No definido	Malo
No diligenciado	Malo
Otros	Malo

Fuente: elaboración de las autoras.

## 13) Menores a cargo

Muestra si el cliente tiene personas a cargo, de manera que se tenga presente el factor de riesgo de un impago futuro [Tabla 17].

Tabla 17. La Entidad. Menores a cargo

Menores a cargo	Clasificación
Sí	Malo
No	Excelente

Fuente: elaboración de las autoras.

## 14) Estrato

Indicador sociodemográfico que es una referencia del estado actual donde reside el cliente [Tabla 18].

Tabla 18. La Entidad. Estrato

Estrato	Clasificación
1	Malo
2	Malo
3	Regular
4	Aceptable
5	Excelente
6	Excelente
No definido	Malo

Fuente: elaboración de las autoras.

## 15) Tipo de vivienda

Muestra el tipo de vivienda en la que actualmente reside el cliente [Tabla 19].

Tabla 19. La Entidad. Tipo de vivienda

Tipo de vivienda	Clasificación
Propia	Excelente
Familiar	Aceptable
Alquiler	Regular
No definido	Malo

Fuente: elaboración de las autoras.

## 16) Patrimonio

Refleja el respaldo con el que cuenta el cliente actualmente. En caso de no tener un valor mayor a cero (0) en estas variables, se clasifica como Malo [Tabla 20].

Tabla 20. La Entidad. Patrimonio

Patrimonio (bien)	Clasificación
Vivienda	Excelente
Vehículo	Excelente
Otros bienes	Excelente
Vivienda - Hipoteca	Aceptable
Vehículo pignorado	Aceptable
Otros bienes pignorados	Aceptable
No posee patrimonio	Malo

Fuente: elaboración de las autoras.

## 17) Rango de ingresos

Permite conocer el estatus de la principal fuente de pago de la responsabilidad interna que tiene el cliente con la Entidad [Tabla 21].

Tabla 21. La Entidad. Rango de ingresos

Rango de ingresos [COP millones]	Clasificación
0 – 2,5	Malo
2,5 – 5	Regular
5 – 10	Aceptable
> 10	Excelente

Fuente: elaboración de las autoras.

## 18) Rango de mora actual

Muestra el comportamiento actual del cliente dentro de la Entidad. Permite observar la fe de pago de la obligación [Tabla 22].

Tabla 22. La Entidad. Rango de mora actual

Rango de mora actual [días]	Clasificación
[0 – 30]	Excelente
(30 – 60]	Aceptable
(60 – 90]	Regular
> 90	Malo

Fuente: elaboración de las autoras.

c) Identificación de factores de riesgo. Este paso incluye la reclasificación en grupos de mayor riesgo de todos aquellos clientes que posean otros factores de riesgo diferentes a mora o alertas internas propias de la Entidad y externas tomadas del comportamiento en el sector. La Entidad entrega trimestralmente un archivo Excel a Datacrédito-Experian (<https://www.datacredito.com.co/>) para el proceso de recalificación de cartera.

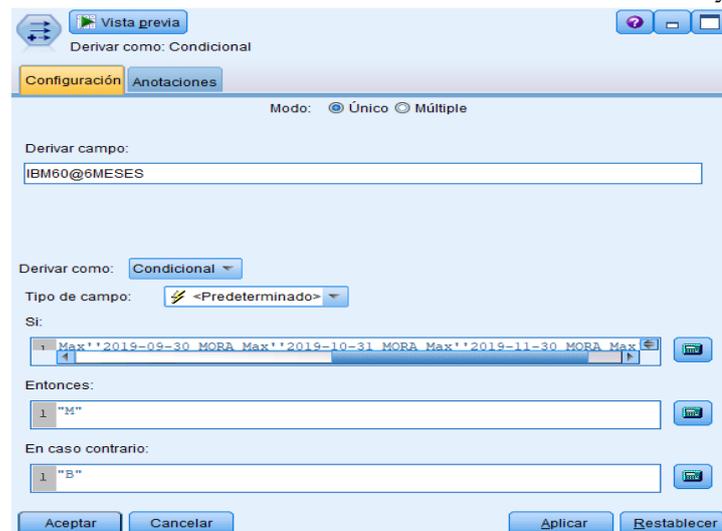
Las alertas consideradas son dos:

- Interna: si el cliente declara o no renta
- Externa: si el cliente tiene un castigo vigente en el sector

En ese sentido se propone que si el cliente tiene activo este factor de riesgo (1 = factor de riesgo; 0 = no factor de riesgo), la Entidad está expuesta a un mayor riesgo; por tanto, se incluye dentro del modelo.

d) Definición de incumplimiento (*default*). En este paso se define el *default* como la altura de mora donde existe una mayor probabilidad de seguirse deteriorando; normalmente para su definición se considera como período 12 meses [Figura 11].

Figura 11. La Entidad. Ventana del SPSS Modeler. Indicador de clientes buenos y malos



Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

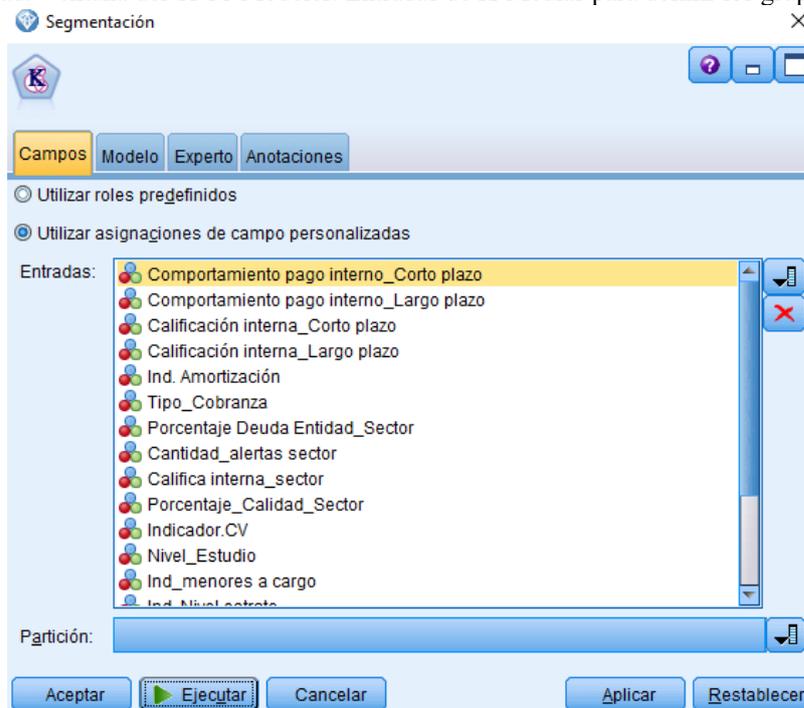
En razón de que se ha encontrado que 60 días de mora definen el incumplimiento (*default*) para una persona natural en la Entidad, se calcula el indicador Buenos y Malos a seis meses.

e) Procedimiento para definir los grupos de riesgo. Se procede a diseñar el modelo con la combinación de las técnicas estadísticas de agrupación por conglomerados (clústeres) y de los árboles de decisión, ajustada al propósito que se pretende: reclasificar a los clientes de la cartera en grupos de riesgo para su seguimiento.

En primer lugar, se utiliza el módulo de agrupación K-Medias, que permite obtener grupos homogéneos entre sí y diferentes entre ellos, con la aplicación del SPSS Modeler (International Business Machines Corporation, IBM, s. f.), donde las entradas son los

indicadores propuestos, incluyendo el factor de riesgo identificado [Figura 12], definiendo a priori 5 grupos de conglomerados para la salida de acuerdo con la Entidad.

Figura 12. La Entidad. Ventana del SPSS Modeler. Entradas de K-Medias para definir los grupos de riesgo

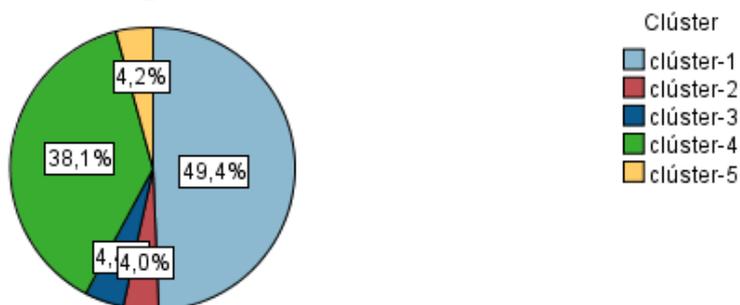


Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

Específicamente, el nodo K-Medias agrupa conjuntos de datos en grupos distintos o clústeres; el método define un número fijo de estos, donde de forma iterativa les asigna registros y ajusta sus centros hasta que el modelo no pueda ser mejorado. En lugar de intentar predecir un resultado, los modelos K-Medias utilizan un proceso conocido como “aprendizaje no supervisado” para revelar los patrones del conjunto de los campos de entrada.

Posteriormente se procede a ejecutar el modelo de segmentación, donde se obtienen cinco clústeres definidos a priori de acuerdo con la base de los 27.280 clientes independientes, donde los clústeres 1 y 4 tienen la mayor parte de la cartera independiente (23.846 clientes – 87.5%) [Figura 13 y Tabla 23].

Figura 13. La Entidad. Tamaños de los clústeres



Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

Tabla 23. La Entidad. Clústeres – Segmentación de los clientes independientes

Tamaño	Clúster 1	Clúster 2	Clúster 3	Clúster 4	Clúster 5
	<b>49.4% (13.469)</b>	<b>4.0% (1.097)</b>	<b>4.4% (1.190)</b>	<b>38.1% (10.387)</b>	<b>4.2% (1.137)</b>
Entradas	Calif_interna_corto plazo Excelente (99.7%)	Calif_interna_corto plazo Malo (97.2%)	Calif_interna_corto plazo Aceptable (96.4%)	Calif_interna_corto plazo Excelente (99.5%)	Calif_interna_corto plazo Regular (88.2%)
	Calif_interna_largo plazo Excelente (97.6%)	Calif_interna_largo plazo Malo (98.4%)	Calif_interna_largo plazo Aceptable (93.4%)	Calif_interna_largo plazo Excelente (96.7%)	Calif_interna_largo plazo Regular (84.8%)
	Calificación interna vs sector Aceptable (70.9%)	Calificación interna vs sector Excelente (83.6%)	Calificación interna vs sector Excelente (50.3%)	Calificación interna vs sector Excelente (54.3%)	Calificación interna vs sector Excelente (72.6%)
	Cantidad Alertas_Sector Aceptable (51.2%)	Cantidad Alertas_Sector Malo (37.2%)	Cantidad Alertas_Sector Aceptable (43.5%)	Cantidad Alertas_Sector Excelente (80.0%)	Cantidad Alertas_Sector Excelente (29.6%)
	Comportamiento Pago interno_CP Excelente (99.8%)	Comportamiento Pago interno_CP Malo (95.0%)	Comportamiento Pago interno_CP Aceptable (88.1%)	Comportamiento Pago interno_CP Excelente (99.7%)	Comportamiento Pago interno_CP Regular (60.9%)
	Comportamiento Pago interno_LP Malo (92.9%)	Comportamiento Pago interno_LP Excelente (42.2%)	Comportamiento Pago interno_LP Malo (88.8%)	Comportamiento Pago interno_LP Malo (84.9%)	Comportamiento Pago interno_LP Malo (49.9%)
	FACTOR_RIESGO Media 0.20	FACTOR_RIESGO Media 0.78	FACTOR_RIESGO Media 0.54	FACTOR_RIESGO Media 0.87	FACTOR_RIESGO Media 0.58
	Tipo_Vivienda Excelente (70.4%)	Tipo_Vivienda Aceptable (54.0%)	Tipo_Vivienda Aceptable (53.9%)	Tipo_Vivienda Aceptable (72.0%)	Tipo_Vivienda Aceptable (51.3%)
	Nivel_Estrato Excelente (47.7%)	Nivel_Estrato Aceptable (32.3%)	Nivel_Estrato Excelente (36.2%)	Nivel_Estrato Regular (36.7%)	Nivel_Estrato Excelente (34.5%)
	Indicador_CV Excelente (99.3%)	Indicador_CV Malo (81.7%)	Indicador_CV Excelente (71.8%)	Indicador_CV Excelente (98.3%)	Indicador_CV Excelente (58.5%)
	Patrimonio Excelente (94.6%)	Patrimonio Excelente (75.7%)	Patrimonio Excelente (75.5%)	Patrimonio Excelente (53.7%)	Patrimonio Excelente (80.9%)
	%Calidad_Sector Excelente (96.2%)	%Calidad_Sector Excelente (54.3%)	%Calidad_Sector Excelente (78.4%)	%Calidad_Sector Excelente (96.9%)	%Calidad_Sector Excelente (60.2%)
	%Deuda Malo (90.0%)	%Deuda Malo (77.8%)	%Deuda Malo (86.0%)	%Deuda Malo (53.0%)	%Deuda Malo (87.9%)
	Rango_Ingresos Aceptable (38.3%)	Rango_Ingresos Regular (28.7%)	Rango_Ingresos Regular (29.7%)	Rango_Ingresos Regular (40.6%)	Rango_Ingresos Aceptable (31.7%)
	Rango_Mora Actual Excelente (99.2%)	Rango_Mora Actual Malo (72.7%)	Rango_Mora Actual Excelente (71.5%)	Rango_Mora Actual Excelente (98.3%)	Rango_Mora Actual Excelente (57.2%)
	Tipo_Cobranza Excelente (82.6%)	Tipo_Cobranza Malo (76.3%)	Tipo_Cobranza Excelente (54.7%)	Tipo_Cobranza Excelente (83.7%)	Tipo_Cobranza Excelente (48.9%)
	Nivel_Académico Excelente (84.5%)	Nivel_Académico Excelente (74.5%)	Nivel_Académico Excelente (79.1%)	Nivel_Académico Excelente (69.7%)	Nivel_Académico Excelente (80.6%)
	Menores a Cargo Excelente (69.6%)	Menores a Cargo Excelente (73.5%)	Menores a Cargo Excelente (70.7%)	Menores a Cargo Excelente (81.3%)	Menores a Cargo Excelente (71.9%)
%Amortización Malo (58.1%)	%Amortización Malo (60.4%)	%Amortización Malo (58.0%)	%Amortización Malo (53.3%)	%Amortización Malo (54.1%)	
<b>% Bad Rate</b>	<b>0.025%</b>	<b>98.359%</b>	<b>2.941%</b>	<b>0.025%</b>	<b>62.445%</b>

Fuente: elaboración de las autoras a partir de datos de la entidad estudiada (archivo personal de las autoras).

Asimismo, se muestra la importancia del predictor de las variables de entrada, que sirve para identificar aquellas que más influyen en cada conglomerado y reconocer los perfiles de dichos clientes a la Entidad, según el riesgo asociado [Tabla 24].

Tabla 24. Importancia del predictor de las variables de entrada

Importancia de entrada (predictor)  
 ■ 1,0 ■ 0,8 ■ 0,6 ■ 0,4 ■ 0,2 ■ 0,0

Variable	Importancia Predictor
Calif_interna_corto plazo	1.00
Calif_interna_largo plazo	1.00
Calificación interna vs sector	1.00
Cantidad Alertas_Sector	1.00
Comportamiento Pago interno_CP	1.00
Comportamiento Pago interno_LP	1.00
FACTOR_RIESGO	1.00
Tipo_Vivienda	1.00
Nivel_Estrato	1.00
Indicador_CV	1.00
Patrimonio	1.00
%Calidad_Sector	1.00
%Deuda	1.00
Rango_Ingresos	1.00
Rango_Mora Actual	1.00
Tipo_Cobranza	1.00
Nivel_Académico	0.55
Menores a Cargo	0.3
%Amortización	0.05

Fuente: elaboración de las autoras.

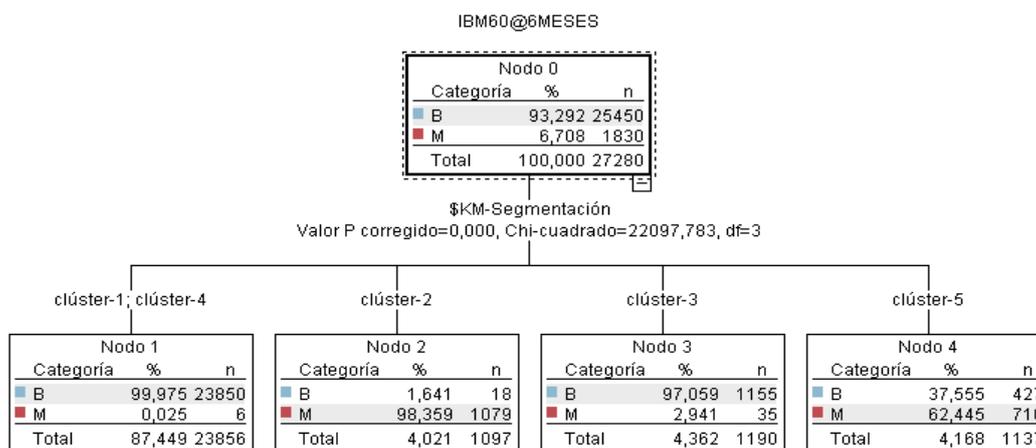
- f) Evaluación del modelo. A cada uno de los clústeres se le evalúa su riesgo usando árboles de decisión CHAID, ya que con esta técnica se logra abordar la solución a los problemas de predicción, clasificación y segmentación, dándole a cada conglomerado un riesgo (% Bad Rate) según la definición de *default* o incumplimiento, con un indicador de buenos y malos (IBM60) a seis meses, todo ello con el fin de respaldar y examinar los resultados y

determinar visualmente cómo fluye el modelo, para así asignar la estrategia o nivel de riesgo para el seguimiento.

En cuanto al SPSS Modeler, existen diferentes tipos de árbol –v. § 3.2.1, Modelos de medición del riesgo de crédito, RC, *Árboles de decisión*–. Para el presente trabajo, el modelo KM-Segmentación se evalúa con el árbol de decisión CHAID, que consiste en un rápido algoritmo de árbol estadístico y multidireccional que explora datos de forma rápida y eficaz y crea segmentos y perfiles con respecto al resultado deseado a través de interacciones mediante Chi-cuadrado.

Este modelo elige la variable independiente (predictora) –en este caso, el modelo KM-Segmentación– que presenta la interacción más fuerte con la variable dependiente –el IBM60 a seis meses–. Si llegare a encontrar categorías de cada predictor que no son significativamente distintas respecto a la variable dependiente, estas se funden; es el caso de los clústeres 1 y 4 [Figura 14].

Figura 14. La Entidad. Diagrama de árbol para el modelo de nivel de riesgo de la cartera independiente



Fuente: elaboración de las autoras.

En el nodo 0 se describe la variable dependiente: Buenos (B) y Malos (M) de la cartera de clientes independientes. Seguidamente se observa que la variable dependiente se ramifica en cuatro nodos pertenecientes a los clústeres obtenidos en el modelo.

En relación con la Figura 14, algunas conclusiones del árbol son las siguientes:

- Los clústeres 1 y 4 no tiene variables significativamente distintas respecto al indicador Buenos y Malos a seis meses; por tanto, se funden en un solo nodo. Esto indica que posiblemente las características o el perfil de este grupo de clientes son similares.
- Estos mismos (1 y 4) son los clústeres con la tasa de riesgo Malos más baja: 0,025 %. Así, se puede clasificar a este grupo como el de menor riesgo.
- El clúster 3 tiene una tasa de Malos levemente más alta que los clústeres 1 y 4, por lo que la Entidad define clasificar a este grupo como los de riesgo moderado.
- El clúster 2 es el de mayor clasificación de clientes Malos: 98,359 %; por tanto, permite inducir que son clientes con características o un perfil de mayor riesgo.

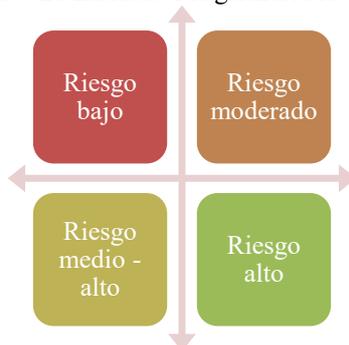
- El clúster 5 tiene una tasa de riesgo de Malos del 62,445 %, lo que muestra un nivel de riesgo intermedio levemente menor al Alto.

En síntesis, dicho análisis es la partida para la interpretación de los resultados.

## 5. Resultados

En esta sección se lleva a cabo la interpretación del modelo de seguimiento del RC propuesto para la cartera de clientes independientes de la entidad estudiada (la Entidad), realizando una reclasificación en cuatro grupos de riesgo a partir de los conglomerados generados [Figura 15 y Figura 16].

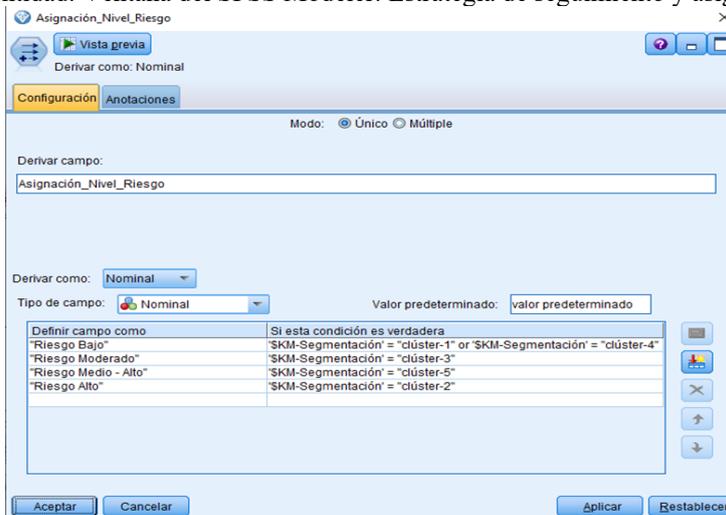
Figura 15. La Entidad. Conglomerados de riesgo



Fuente: elaboración de las autoras.

- 1) Riesgo bajo (clústeres 1 y 4). Representa un nivel de incertidumbre bajo que no potencializaría grandes pérdidas; sin embargo, debe contar con un monitoreo para evitar incrementos en el nivel de riesgo.
- 2) Riesgo moderado (clúster 3). Este grupo de clientes debe ser observado y tratado para que en un mediano plazo no genere grandes pérdidas.
- 3) Riesgo medio-alto (clúster 5). Este grupo de clientes debe ser monitoreado para que no genere grandes pérdidas, pues sus factores de riesgo son mayores.
- 4) Riesgo alto (clúster 2). Como su nombre lo indica, abarca un segmento de clientes con nivel de incertidumbre muy alta que podría materializar grandes pérdidas.

Figura 16. La Entidad. Ventana del SPSS Modeler. Estrategia de seguimiento y asignación de riesgo



Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

Las estrategias de riesgo expuestas se basan en las características del clúster según la evaluación del árbol de decisiones, donde se le asigna una tasa de %Bad Rates a cada conglomerado. Dado que el riesgo es una incertidumbre ante la duda de que algo ocurra, en otras palabras, la probabilidad de que una amenaza se convierta en desastre, se trata de gestionarlo o reducirlo a través de acciones que minimicen su impacto.

Por esta razón, según el nivel de impacto en las provisiones que puedan generar, los riesgos se categorizan de menor a mayor. Para el modelo interno de seguimiento a la cartera de los clientes independientes de la Entidad se trabaja con riesgos intermedios, en razón de que tener pocas segmentaciones entre los extremos causa una brecha amplia entre los segmentos que podría categorizar a clientes muy cercanos –pero con características diferentes– en un solo segmento y con un mismo tratamiento.

De la mano del análisis del árbol de decisiones se obtiene la reclasificación de los clústeres por nivel de riesgo, que permite obtener el resultado del modelo para el seguimiento de la cartera de los clientes independientes de la Entidad y apoyar la toma de acciones preventivas frente un posible deterioro e impacto en los resultados de la cartera. Así se obtiene el archivo resultante con el detalle cliente a cliente y su estrategia de seguimiento.

Los resultados de análisis del modelo señalan que este clasifica correctamente el 98,2 % de los clientes con tan solo el 1,8 % de error [Figura 17]. Con respecto a las métricas de evaluación se observa su efectividad y ajuste, ya que estas son cercanas a 1 [Figura 18]; el 0,993 del indicador AUC muestra que las predicciones son correctas en el 99,3 %. Este indicador, que mide la exactitud y la calidad de las predicciones, oscila entre 0 y 1. Así, se puede inferir que el modelo clasifica correctamente a los clientes de la Entidad dentro de la segmentación obtenida.

Figura 17. La Entidad. Clasificación del modelo

<b>Correctos</b>	26.790	98,2%
<b>Erróneos</b>	490	1,8%
<b>Total</b>	27.280	

Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

Figura 18. La Entidad. Métricas de evaluación

<b>Modelo</b>	<b>AUC</b>	<b>Gini</b>
<b>\$R-IBM60@6MESES</b>	0,993	0,987

Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

Finalmente, la matriz de coincidencias permite conocer, observando la diagonal de dicha matriz, qué tan efectivo es el modelo en el momento de clasificar y calificar [Figura 19].

Figura 19. La Entidad. Matriz de coincidencias

	<b>B</b>	<b>M</b>
<b>B</b>	25.002	448
<b>M</b>	42	1.788

Fuente: elaboración de las autoras a partir de International Business Machines Corporation, IBM (s. f.).

## 6. Conclusiones y recomendaciones

### 6.1 Conclusiones

El estudio realizado en esta investigación ha permitido desarrollar un modelo híbrido *clustering*-árboles de decisión para una mejor evaluación y desempeño en torno a la predicción y calificación, como apoyo a la gestión del RC. Los resultados del análisis del modelo permiten a la Entidad confiar en la asignación de la estrategia de seguimiento, ya que este clasifica correctamente a los clientes buenos y malos.

La gestión del RC en las entidades financieras se convierte en una solución importante para medir y monitorear el riesgo de una cartera bancaria que apoya la maximización de los rendimientos y mantiene la exposición al riesgo dentro de parámetros aceptables para ellas.

Al ser una nueva propuesta de modelo –la Entidad no cuenta con uno para el seguimiento a este segmento de la cartera–, los indicadores son asignados en conjunto con el panel de expertos en el área y conocimientos en la gestión de riesgos. De este modo, el modelo se podrá calibrar y ajustar según las necesidades y alertas que evidencie la Entidad, puesto que sus colaboradores son personas idóneas, con el conocimiento y la experiencia del tema, que pueden darle la trascendencia y la oportunidad de aplicación para una gestión adecuada de la gestión del riesgo de la cartera, como un apoyo para presupuestar pérdidas esperadas, pérdidas no esperadas y provisiones, entre otras, como una herramienta de ayuda a su gestión financiera.

Tanto el procesamiento de los datos como la construcción del modelo se realizaron con la herramienta SPSS Modeler (International Business Machines, IBM, s. f.), una aplicación de software de análisis de texto y minería de datos muy utilizada para crear modelos predictivos y realizar otras tareas de analíticas.

El modelo cumple con la normatividad vigente en materia de RC y a la par se ajusta a las particularidades que tiene la cartera de banca personal, específicamente la de los clientes independientes de la Entidad, teniendo en cuenta que esta no contaba con un seguimiento a este perfil en específico y lo realizaba dentro de un modelo general: la cartera de asalariado.

Este sistema permite arrojar alertas tempranas, segmentarlas según su impacto y plantear planes de acción oportunos direccionados a disminuir los posibles incumplimientos en la cartera, sobre todo en un segmento considerado de riesgo debido a su perfil. Así, la Entidad puede proponer políticas más robustas sobre este segmento para el monitoreo y vigilancia que le aporten no solo a su esquema de riesgos, sino a todas las áreas involucradas en el proceso de carrea de una u otra manera.

La propuesta del modelo ayuda a profundizar y mejorar la etapa de seguimiento y control que posee actualmente la Entidad, proporcionando una herramienta al área encargada en la gestión del riesgo de crédito con la categorización de los clientes al asignarles un nivel de riesgo,

retroalimentando las etapas de otorgamiento y recuperación y engranando el proceso hacia un mismo resultado: maximizar su beneficio.

Por normativa de la Superintendencia, los modelos requieren de una revisión semestral –en mayo y noviembre–, a fin de mejorar y calibrar su asertividad. Por tanto, se propone inicialmente una revisión trimestral de la mano de los modelos internos de recalificación trimestral de la cartera Persona natural y Persona jurídica, ya que es la periodicidad con la que cuenta la Entidad para validar los modelos y actualizar la información externa.

Es importante que la Entidad cuente con información sectorial proveniente de las centrales de riesgo –en este caso Datacrédito-Experian–, ya que permite una visión más amplia del comportamiento del cliente.

## 6.2 Recomendaciones

Crear planes de trabajo periódicos con el personal encargado del área de seguimiento y control de la cartera para hacerle un monitoreo a las actividades y tener retroalimentaciones.

Como es un modelo nuevo para el monitoreo y seguimiento del segmento independiente de la cartera de la Entidad, se recomienda probar el modelo por tres meses, en los que se le realicen los ajustes necesarios y se calibre según sus necesidades, de modo que permita retroalimentar otras áreas de RC como las de otorgamiento y análisis de riesgo de cartera.

Divulgar la información obtenida a través de medios internos como informes de gestión y el comité de riesgos o seguimiento al perfil independiente para que complementen el modelo que permita a otras áreas tomar decisiones y realizar retroalimentaciones, dejando a discreción de la Entidad si se realizarán con la misma periodicidad con la que se correrá el modelo o si se cambiará la regularidad para su emisión y divulgación.

Complementar el modelo con un análisis del impacto económico sobre las provisiones, en el que se plantee que, según el nivel de riesgo, se deteriore la calificación del cliente en un grado y así, por medio del simulador interno de la Entidad, realizar el cálculo de los modelos MRCO y MRCH para conocer el impacto en el PyG (la cuenta de pérdidas y ganancias) de los clientes.

## Referencias

- Ammann, M. (2001). *Credit risk valuation. Methods, models and applications*. Ciudad de Nueva York, NY: Springer Finance.
- Benjumea Velásquez, A. (2013). *Diseño de un modelo predictivo de seguimiento de riesgo de crédito para la cartera comercial, para una entidad financiera del Valle de Aburrá* [tesis de especialización, Universidad de Medellín]. Disponible en <https://repository.udem.edu.co/bitstream/handle/11407/111/Dise%C3%B1o%20de%20un%20modelo%20predictivo%20de%20seguimiento%20de%20riesgo%20de%20cr%C3%A9dito%20para%20la%20cartera%20comercial%20C%20para%20una%20entidad%20financiera%20del%20Valle%20de%20Aburra.pdf?sequence=1&isAllowed=y>
- Berlanga Silvente, V., Rubio Hurtado, M. J. y Vilà Baños, R. (2013). Cómo aplicar árboles de decisión en SPSS. *REIRE. Revista d'Innovació i Recerca en Educació*, 6(1), 65-79. Disponible en <https://revistes.ub.edu/index.php/REIRE/article/viewFile/5155/7229>
- Bravo Román, C., Maldonado, S. y Weber, R. (2009). Seguimiento en modelos de regresión logística. *Ingeniería Industrial*, 8(2), 31-44. Disponible en <http://revistas.ubiobio.cl/index.php/RI/article/view/73/3312>
- Breiman, L., Friedman, J., Stone, C. J. y Olshen, R. A. (1984). *Árboles de clasificación y regresión*. Boca Ratón, FL: CRC Press.
- Caballo Trebot, Á. (2013). *Medición de riesgo de crédito: desarrollo de una nueva herramienta*. Madrid: Universidad Pontificia Comillas.
- Caicedo Cerezo, E., Claramunt Bielsa, M. M. y Casanovas Ramón, M. (2011). Medición del riesgo de crédito mediante modelos estructurales: una aplicación al mercado colombiano. *Cuadernos de Administración*, 24(42), 73-100. Disponible en <http://www.scielo.org.co/pdf/cadm/v24n42/v24n42a04.pdf>
- Chorafas, D. N. (2000). *Managing credit risk*. Londres: Euromoney Books.
- Colombia Ministerio de Salud y Protección Social (2016). *Resolución 5858* [en línea]. Disponible en <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/resolucion-5858-de-2016.pdf>
- Colombia, Archivo Nacional de Datos, ANDA (2019). *Gran encuesta integrada de hogares – GEIH* [en línea]. Disponible en <http://microdatos.dane.gov.co/index.php/catalog/599/study-description>
- Colombia, Congreso de la República (1923). *Ley 45*. Bogotá: Diario Oficial 19.137. Disponible en <http://www.suin-juriscol.gov.co/viewDocument.asp?id=1599516>
- Colombia, Congreso de la República (1979). *Ley 32* [en línea]. Disponible en <https://www.javeriana.edu.co/personales/hbermude/leycontable/contadores/1979-ley-32.pdf>
- Colombia, Congreso de la República (2005). *Ley 964*. Bogotá: Diario Oficial 45.963. Disponible en [http://www.secretariasenado.gov.co/senado/basedoc/ley\\_0964\\_2005.html](http://www.secretariasenado.gov.co/senado/basedoc/ley_0964_2005.html)

- Colombia, Presidencia de la República (2005). *Decreto 4327*. Bogotá: Diario Oficial 46.104. Disponible en [http://www.secretariassenado.gov.co/senado/basedoc/decreto\\_4327\\_2005.html](http://www.secretariassenado.gov.co/senado/basedoc/decreto_4327_2005.html)
- Colombia, Superintendencia Financiera (1995). *Circular básica contable y financiera (Circular externa 100 de 1995)* [en línea]. Disponible en <https://www.superfinanciera.gov.co/inicio/normativa/normativa-general/circular-basica-contable-y-financiera-circular-externa--de---15466>; y <https://legislacion.vlex.com.co/vid/circular-externa-basica-contable-398711989>
- De Andrés Suárez, J. (2000). *Técnicas de inteligencia artificial aplicadas al análisis de la solvencia empresarial*. Universidad de Oviedo, Facultad de Ciencias Económicas, Documento de Trabajo 209. Disponible por descarga en <https://dialnet.unirioja.es/servlet/articulo?codigo=1252835>
- De la Fuente Fernández, S. (2011). *Análisis conglomerados*. Madrid: Universidad Autónoma de Madrid, Facultad de Ciencias Económicas y Empresariales.
- Doldán Tié, F. R. y Rodríguez López, M. (2002). *La gestión del riesgo de crédito: métodos y modelos de predicción de la insolvencia empresarial*. Madrid: Asociación Española de Contabilidad y Administración de Empresas.
- Economipedia (s. f.). *Comité de Basilea* [en línea]. Disponible en <https://economipedia.com/definiciones/comite-de-basilea.html>
- Elizondo, A. y Altman, E. I. (2004). *Medición integral del riesgo de crédito*. México: Limusa.
- Elizondo, A. y López Romero, C. (1999). El riesgo de crédito en México: una evaluación de modelos recientes para cuantificarlo. *Gaceta de Economía*, 4(8), 51-74.
- Flórez López, R. (2007). Análisis de los determinantes del riesgo de crédito en presencia de carteras de bajo incumplimiento. Una nueva propuesta de aplicación. *Revista Europea de Dirección y Economía de la Empresa*, 16(2), 71-92. Disponible por descarga en <https://dialnet.unirioja.es/servlet/articulo?codigo=2336244>
- García Baena, R., González Mosquera, L. y Oroz García, M. (2005). Aspectos críticos en la implantación y validación de modelos internos de riesgo de crédito. *Estabilidad financiera*, 9, 29-57, noviembre. Disponible en <https://www.bde.es/f/webbde/Secciones/Publicaciones/InformesBoletinesRevistas/RevistaEstabilidadFinanciera/05/Fich/estfin09.pdf>
- García Cruz, G. A. (2011). Determinantes macro y efectos locales de la informalidad laboral en Colombia. *Sociedad y Economía*, 21, 69-98. Disponible en <http://www.scielo.org.co/pdf/soec/n21/n21a04.pdf>
- Henaó Jassan, R. (2013). *Scoring de seguimiento para el cálculo de pérdidas esperadas y capital económico de una entidad colombiana* [tesis de maestría, Universidad EAFIT]. Disponible en [https://repository.eafit.edu.co/xmlui/bitstream/handle/10784/1446/HenaóJassan\\_Rodrigo\\_2013.pdf;jsessionid=F30498C69F83AEAEF4B0F606D17545FF?sequence=1](https://repository.eafit.edu.co/xmlui/bitstream/handle/10784/1446/HenaóJassan_Rodrigo_2013.pdf;jsessionid=F30498C69F83AEAEF4B0F606D17545FF?sequence=1)
- Hernández Ramírez, M. (2014). Modelo financiero para la detección de quiebras con el uso de análisis discriminante múltiple. *InterSedes: Revista de las Sedes Regionales*, 15(32), 4-19. Disponible en <https://www.redalyc.org/pdf/666/66633023001.pdf>
- International Business Machines Corporation, IBM (s. f.). IBM SPSS Modeler. Sitio web <https://www.ibm.com/products/spss-modeler>

- Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29(2), 119-127, doi:10.2307/2986296 Disponible en <https://www4.stat.ncsu.edu/~dickey/analytics/datamine/Reference%20Papers/kass80.pdf>
- Mermelstein, D. A. (2006). *Defaults en carteras hipotecarias, macroeconomía y arreglos institucionales: más allá de los modelos de Credit-Scoring tradicionales*. Munich Personal RePEc Archive (MPRA) Paper. Disponible en [http://bibliotecadigital.econ.uba.ar/download/docin/docin\\_cma\\_005.pdf](http://bibliotecadigital.econ.uba.ar/download/docin/docin_cma_005.pdf)
- Mesa González, H. (2003). *La superintendencia bancaria: origen y antecedentes* [en línea]. Disponible en <https://www.superfinanciera.gov.co/SFCant/ComunicadosyPublicaciones/80web/archivos/Humbertomesa.pdf>
- Pérez López, C. y Santín González, D. (2007). *Minería de datos. Técnicas y herramientas*. Madrid: Paraninfo.
- Pra Martos, I., Ríos, A., Arguedas Sanz, R. y Casals Carro, J. (2010). *Gestión y control del riesgo de crédito en modelos avanzados*. Madrid: Ediciones Académicas.
- Rankia (s. f.). *¿Qué es el SARC y para qué sirve?* [en línea]. Disponible en <https://www.rankia.co/blog/mejores-creditos-y-prestamos-colombia/4104690-que-sarc-para-sirve>
- RiskMetrics Group y J. P. Morgan & Co. (1997/2007). *CreditMetrics. Technical Document*. Ciudad de Nueva York, NY: RiskMetrics Group.
- Saavedra García, M. L. y Saavedra García, M. J. (2010). Modelos para medir el riesgo de crédito de la banca. *Cuadernos de Administración*, 23(40), 295-319, enero-junio. Disponible en <http://www.scielo.org.co/pdf/cadm/v23n40/v23n40a13.pdf>
- Saunders, A. y Allen, L. (1999). *Credit risk measurement: New approaches to value at risk and other paradigms*. Hoboken, NJ: John Willey & Sons.
- Sistema Económico Latinoamericano y del Caribe, SELA (2018). *Con 61,3 %, Colombia es el primer país del mundo con empleados autónomos* [en línea, 11 y 12 de julio]. Disponible en <http://www.sela.org/es/prensa/servicio-informativo/20180711/si/26939/colombiaemprend>
- Spuchl'áková, E., Valášk'ová, K. y Adamko, P. (2015). The credit risk and its measurement, hedging and monitoring. *Procedia Economics and Finance*, 24, 675-681. [https://doi.org/10.1016/S2212-5671\(15\)00671-1](https://doi.org/10.1016/S2212-5671(15)00671-1)
- Trigo Martínez, E. (2009). *Análisis y medición del riesgo de crédito en carteras de activos financieros ilíquidos emitidos por empresas* [tesis doctoral, Universidad de Málaga]. Disponible en <https://riuma.uma.es/xmlui/bitstream/handle/10630/4068/An%c3%a1lisis%20y%20medici%c3%b3n%20del%20riesgo%20de%20cr%c3%a9dito%20en%20carteras.pdf?sequence=1&isAllowed=y>