



**Colegio de Estudios
Superiores de Administración**

Revolucionando el Acceso al Crédito: Cómo el Scoring Alternativo puede Impulsar las PYMES en Colombia

Felipe Jaramillo Peña & Enrique Colimodio

**Pregrado en Administración de empresas
Colegio de Estudios Superiores de Administración (CESA)
Bogotá, Colombia
2025**

**Revolucionando el Acceso al Crédito: Cómo el Scoring Alternativo puede Impulsar las PYMES
en Colombia**

Felipe Jaramillo Peña & Enrique Colimodio

Director de trabajo de grado: Guillermo Arevalo

Pregrado en Administración de empresas

Colegio de Estudios Superiores de Administración (CESA)

Bogotá, Colombia

2025

Tabla de contenido

Resumen.....	4
1. Introducción.....	5
1.1 Planteamiento del problema.....	6
1.2 Justificación de la investigación.....	8
1.4 Objetivo general.....	9
1.5 Objetivos específicos.....	10
2. Revisión de literatura.....	11
2.1 Open data.....	14
2.2 Machine learning:.....	18
2.3 Estado del arte:.....	20
2.4 Hipótesis:.....	23
3. Marco teórico:.....	24
3.1 Conceptualización de la inclusión financiera: de la teoría a la práctica.....	24
3.2 Barreras estructurales: un análisis multidimensional.....	25
3.4 Conceptualización del scoring alternativo y su relevancia.....	26
3.5 Evidencia empírica y validación de modelos.....	26
3.6 Desafíos en contextos de alta informalidad.....	27
3.7 Tecnologías habilitadoras: IA, blockchain y banca abierta.....	27
3.8 Fintechs en Colombia.....	28
3.9 Lecciones globales: casos de éxito y adaptabilidad.....	28
3.10 Políticas públicas: logros, limitaciones y recomendaciones.....	29
4. Metodología de la investigación:.....	30
4.1 Tipo y diseño de la investigación.....	30
4.2 Población y muestra.....	30
4.3 Instrumentos de recolección.....	31
4.4 Procedimientos.....	31
5. Análisis de los resultados:.....	32
5.1 Limitaciones de la investigación.....	34
6. Discusión.....	35
7. Conclusión.....	38
8. Recomendaciones a futuras investigaciones.....	39
Referencias.....	41

Indice de Anexos

Anexo:	50
Investigación (Encuesta) Anexo 1	50
Anexo 2	53
Anexo 3	53
Anexo 4	54
Anexo 5	54
Anexo 6	55

Resumen

La presente investigación analiza cómo en Colombia se puede implementar sistemas de scoring alternativos basados en open data y machine learning para mejorar el acceso al crédito de las pequeñas y medianas empresas (PYMES) del sector comercial.

El nace del reconocimiento de que los modelos tradicionales de evaluación crediticia centrados en historial financiero, garantías y documentación excluyen sistemáticamente a muchas PYMES, poniendo barreras al acceso a estos créditos afectando especialmente aquellas sin historial crediticio formal, ubicadas en zonas rurales o que operan bajo esquemas de informalidad. Esta exclusión limita su capacidad de crecer, innovar y formalizarse, perpetuando brechas estructurales y regionales.

La tesis propone que el uso de datos abiertos como registros mercantiles, patrones de consumo digital, pagos a proveedores e interacciones en redes sociales (Big Data) junto con algoritmos de machine learning, permite construir modelos predictivos más inclusivos, precisos y contextuales. Estos modelos, cuando se desarrollan bajo principios de ética, transparencia, gobernanza algorítmica y protección de datos, pueden ayudar con la toma de decisiones crediticias más equitativas y fomentar la inclusión financiera.

Mediante un enfoque metodológico mixto, la investigación encuestó a 98 PYMES formales del sector comercio en Colombia, revelando hallazgos clave sobre su acceso al crédito.

Asimismo, se comparan dos rutas para implementar el scoring alternativo: (1) mediante los bancos tradicionales y (2) a través de agencias independientes especializadas. La tesis concluye que, en el contexto colombiano, las agencias externas ofrecen mayor flexibilidad, innovación y neutralidad, aunque enfrentan desafíos regulatorios y de confianza. La cooperación con el sistema bancario y un marco regulatorio sólido van a ser claves para su éxito en corto, mediano y largo plazo.

Finalmente, el estudio evidencia que el scoring alternativo no solo es viable, sino necesario. Su adecuada implementación puede transformar la lógica de exclusión documental del sistema financiero colombiano hacia un enfoque más integral, dinámico y alineado con los desafíos del siglo XXI.

1. Introducción

La inclusión financiera es un componente fundamental para el desarrollo económico sostenible y la reducción de la pobreza en todo el mundo. Permitir que individuos y empresas accedan a servicios financieros formales es esencial para fomentar el crecimiento económico, impulsar la innovación y promover la equidad social y tiende a ser un tema de mayor importancia en economías emergentes como lo es Colombia. En Colombia, las pequeñas y medianas empresas lo que se conoce como PYMES representan el 99.5% del total de las empresas del país y contribuyen con aproximadamente el 40% del PIB. Además, generan cerca del 80% del empleo formal, lo que las convierte en motores clave para el bienestar socioeconómico de la nación (BBVA Research 2024, febrero).

A pesar de su importancia, las PYMES en Colombia enfrentan desafíos significativos para acceder a servicios financieros formales, especialmente en lo que respecta al crédito bancario. Estudios recientes indican que solo el 78% de las empresas medianas tienen acceso a financiamiento formal, 64% de las pequeñas y el 16% de las micro a comparación con las grandes empresas que el 84% tienen acceso a crédito formal, lo que limita su capacidad para invertir, crecer y competir en mercados nacionales e internacionales. Esta brecha de financiamiento no solo obstaculiza el desarrollo empresarial, sino que también impide que la economía colombiana alcance su máximo potencial de crecimiento y desarrollo inclusivo (BBVA Research 2024, febrero).

En este escenario, la implementación de sistemas con scoring alternativo con open data surge como una solución innovadora para abordar las limitaciones actuales en la evaluación del riesgo crediticio de las PYMES. El Open Data, definido como conjuntos de datos accesibles públicamente y reutilizables, ofrece una amplia gama de información que puede complementar y enriquecer los modelos tradicionales de evaluación crediticia (MinTIC, N/A). Al integrar datos provenientes de diversas fuentes como registros comerciales, transacciones digitales, comportamiento en redes sociales y datos macroeconómicos, los bancos pueden obtener una perspectiva más completa del perfil crediticio de las PYMES, facilitando decisiones de crédito más informadas abriendo la puerta para PYMES que no cuentan con un historial crediticio tradicional (Banco Mundial, 2017).

La adopción de estas metodologías no solo tiene el potencial de mejorar la inclusión financiera de las PYMES, sino que también puede contribuir a la modernización y eficiencia del sector financiero colombiano. Al aprovechar tecnologías avanzadas como el análisis de big data y la inteligencia artificial, los bancos pueden optimizar sus procesos de evaluación de riesgo de esta misma manera pueden reducir costos operativos y desarrollar productos financieros más adaptados a las necesidades específicas de las PYMES que hoy en día son pocas las existentes en el mercado. De esta manera esta investigación se centra en investigar cómo los bancos en Colombia pueden implementar sistemas de scoring alternativo utilizando open data para aumentar la inclusión financiera de las PYMES.

1.1 Planteamiento del problema

El acceso limitado al financiamiento formal es uno de los obstáculos más persistentes y desafiantes que enfrentan los colombianos en el momento que toman la iniciativa de emprender en Colombia. Según el informe "Financiamiento de las PYMES en América Latina" del Banco Interamericano de Desarrollo (BID, 2023) existe una brecha de financiamiento estimada en más de 50 mil millones de dólares para las PYMES en latino america, lo que demuestra un acceso muy limitado a recursos financieros necesarios para sus operaciones y expansión. Esta carencia de financiamiento se traduce en limitaciones para invertir en capital humano, tecnología e innovación, elementos de gran importancia para aumentar la productividad y competitividad de las empresas tanto a nivel nacional como internacional (OECD, 2024).

Una de las principales causas de este problema es la dependencia de los bancos en modelos tradicionales de evaluación de riesgo crediticio que se basan principalmente en información financiera histórica, como estados financieros auditados, historial crediticio previo y garantías colaterales (Banco de la República, 2023). Muchas PYMES, especialmente las microempresas y aquellas ubicadas en regiones rurales o menos desarrolladas, carecen de esta información debido a niveles altos de informalidad, prácticas contables deficientes o simplemente por ser empresas de reciente creación sin un historial financiero. Esta falta de información dificulta la evaluación precisa del riesgo crediticio, llevando a los bancos a considerar a estas empresas como inversiones de alto riesgo y, en

consecuencia, restringiendo el acceso al crédito o imponiendo condiciones poco favorables para las empresas en el momento de pedir un crédito.

Otra causa significativa es la insuficiente infraestructura financiera y la baja penetración de servicios bancarios en ciertas regiones del país. Según el Global Findex Database del Banco Mundial (2022), aproximadamente el 45% de los adultos colombianos no tienen acceso a una cuenta bancaria formal, lo que indica una inclusión financiera general limitada que afecta tanto a individuos como a empresas. Esta situación se ve agravada por la falta de competencia en el sector financiero y la ausencia de productos financieros adaptados a las necesidades específicas de las PYMES, especialmente en términos de montos, plazos y condiciones de financiamiento (Banca de las Oportunidades, 2024).

Las consecuencias de este acceso limitado al financiamiento son incontables teniendo en cuenta el efecto dominó que genera en la economía. Económicamente, limita la capacidad de las PYMES para contribuir al crecimiento económico teniendo en cuenta la participación en el empleo que tiene en Colombia. Socialmente, perpetúa las desigualdades regionales y sectoriales, ya que las empresas en áreas urbanas y sectores más desarrollados tienen un acceso relativamente mejor al crédito que aquellas en áreas rurales o sectores emergentes. Además, esta situación obliga a muchas PYMES a recurrir a fuentes de financiamiento informales, que a menudo implican tasas de interés más altas y condiciones menos favorables, aumentando su vulnerabilidad financiera y riesgo de insolvencia (La República, 2023).

La problemática que se pretende abordar con esta investigación es la ausencia de metodologías inclusivas y eficaces para evaluar el riesgo crediticio de las PYMES que carecen de información financiera tradicional. Los sistemas de scoring alternativo con open data ofrecen una oportunidad para superar dichas limitaciones al utilizar datos no tradicionales pero relevantes, como transacciones digitales, interacciones en redes sociales, evaluaciones de proveedores y clientes, y datos sectoriales y macroeconómicos (Asian Development Bank, 2019). Estas metodologías permiten una evaluación más amplia del riesgo crediticio, habilitando mayor acceso al financiamiento para PYMES y promoviendo una mayor inclusión financiera en el país.

Adicionalmente, la implementación de estos sistemas requiere considerar y adaptar este modelo al entorno regulatorio colombiano. Es necesario garantizar la protección de datos y la privacidad de la información, así como asegurar que los modelos de evaluación sean transparentes, justos y libres de sesgos que puedan perpetuar desigualdades existentes. Por lo tanto, esta investigación busca explorar no sólo la viabilidad técnica de implementar sistemas de scoring alternativo utilizando open data, sino también los marcos regulatorios y éticos necesarios para su adopción exitosa y sostenible en el contexto de la coyuntura económica colombiana.

1.2 Justificación de la investigación

La investigación sobre la implementación de sistemas de scoring alternativo utilizando open data para mejorar la inclusión financiera de las PYMES en Colombia es altamente relevante tanto para el sector financiero como para el desarrollo económico y social del país. En primer lugar, desde el ángulo del sector bancario, la adopción de metodologías de evaluación de riesgo más sofisticadas y adaptativas puede significar un aumento de clientes y en nuevas oportunidades de negocio el cual conlleva a una mayor salud financiera en términos de ingresos. Estudios han demostrado que el uso de datos alternativos puede mejorar significativamente la precisión de los modelos de riesgo crediticio, reduciendo las tasas de morosidad y aumentando la rentabilidad de las carteras de crédito (radwaj, A., Narayanan, V. G., & Rao, V. R. 2022). Además, permite a las instituciones financieras diversificar sus productos y servicios, adaptándose mejor a las necesidades específicas de diferentes segmentos de PYMEs y aumentando su competitividad en un mercado cada vez más dinámico y tecnológico (Salvioni, D. M., & Gennari, F. 2023).

Desde una perspectiva económica más amplia, mejorar el acceso al financiamiento formal para las PYMES tiene efectos positivos sobre el crecimiento económico, la generación de empleo y la reducción de la desigualdad (Banco Mundial 2022). Además, el acceso adecuado al financiamiento permite a las PYMES invertir en innovación, mejorar su productividad y competir más efectivamente en mercados internacionales, contribuyendo a la expansión de la economía colombiana.

Socialmente, la inclusión financiera de las PYMES contribuye a la reducción de la pobreza y a la promoción de la equidad. Muchas PYMES operan en sectores y regiones que son esenciales para el crecimiento económico de poblaciones vulnerables el cual incluye áreas rurales. Al facilitar el acceso al crédito, estas empresas pueden expandirse y crear más oportunidades de empleo, mejorando los ingresos y las condiciones de vida de un amplio sector de la población. Además, fomenta la formalización de la economía, lo que se traduce en un aumento de recaudos tributarios que contribuye a mayor provisión de servicios públicos.

Desde el lado académico, esta investigación puede aportar conocimientos y análisis valiosos sobre la aplicación práctica de tecnologías y metodologías emergentes en el sector financiero. Si bien existen estudios previos sobre el uso de open data y scoring alternativo en contextos globales, hay poco interés académico de investigaciones enfocadas específicamente en el contexto colombiano, considerando sus particularidades culturales, económicas y regulatorias. Los hallazgos de esta investigación pueden usarse como base para futuras investigaciones y contribuir al desarrollo de políticas públicas más efectivas y favorables a dicha práctica en el ámbito bancario.

En resumen, esta investigación es de alta relevancia, ya que aborda una problemática crítica para el desarrollo económico y social de Colombia, proponiendo investigar una solución innovadora que puede tener un impacto significativo en la inclusión financiera de las PYMES y, por ende, en el crecimiento económico del país.

1.3 Pregunta de investigación

¿Cómo implementar sistemas de scoring alternativo en Colombia utilizando open data para mejorar el acceso al crédito de las PYMES del sector comercio?

1.4 Objetivo general

Investigar cómo en Colombia se puede implementar sistemas de scoring alternativos basados en open data y machine learning para mejorar el acceso al crédito de las PYMES del sector comercial. Este objetivo incluye evaluar cómo estas tecnologías pueden superar las limitaciones de los modelos

tradicionales de evaluación crediticia, caracterizados por su dependencia de información financiera histórica, y permitir una evaluación más inclusiva del riesgo crediticio. Asimismo, busca fomentar la inclusión financiera al proporcionar a las PYMES mayores oportunidades de financiamiento, lo que impulsará su crecimiento, y este fortalecerá su contribución al desarrollo económico nacional y promoverá la equidad social, especialmente en contextos de alta informalidad y exclusión financiera.

1.5 Objetivos específicos

1. Analizar las principales barreras de acceso al crédito que enfrentan las PYMES del sector comercial en Colombia bajo los modelos tradicionales de evaluación crediticia

Este objetivo busca realizar un análisis de las limitaciones y desafíos que enfrentan las pequeñas y medianas empresas (PYMES) del sector comercial al intentar acceder a financiamiento formal en Colombia. Se pondrá especial énfasis en las características de los modelos tradicionales de evaluación crediticia, los cuales se basan predominantemente en historiales financieros formales, estados financieros auditados y garantías colaterales.

El análisis incluirá la identificación de factores estructurales y contextuales que perpetúan la exclusión financiera, tales como la informalidad empresarial, las prácticas contables deficientes y la falta de información histórica sobre solvencia crediticia. Además, se investigarán las percepciones de riesgo de los bancos hacia estas empresas, particularmente aquellas ubicadas en regiones rurales o sectores emergentes, donde las condiciones del mercado aumentan la percepción de incertidumbre.

2. Examinar el comportamiento de sistemas de scoring alternativo directamente en los bancos tradicionales frente a la creación de agencias especializadas independientes

Las agencias especializadas, en este contexto, se definen como empresas con un enfoque exclusivo en el scoring alternativo. Se trata de empresas privadas cuyo objetivo es evaluar el riesgo de las PYMES mediante este tipo de scoring, de modo que los resultados sean validados por los bancos. De esta

manera, se fomenta una alternativa para la concesión de créditos a PYMES a través de una evaluación externa al banco.

Este objetivo busca comparar dos enfoques clave para la implementación de sistemas de scoring alternativo: por un lado, la adopción por parte de los bancos tradicionales, analizando su capacidad para integrar tecnologías como machine learning y open data dentro de sus estructuras operativas existentes, y por otro, la creación de agencias independientes especializadas, evaluando su flexibilidad, neutralidad y potencial para reducir costos y barreras de acceso al crédito. El análisis se centrará en determinar cuál de estos dos modelos es más eficaz para ampliar el acceso al financiamiento de las PYMES, especialmente aquellas en sectores informales o regiones rurales, que históricamente han enfrentado mayores dificultades para acceder al crédito formal.

Se estudiarán casos exitosos de implementación de open data en otros contextos internacionales, identificando buenas prácticas y lecciones aprendidas que puedan adaptarse al entorno colombiano. Este objetivo permitirá entender si es más eficaz implementar los modelos de scoring alternativo dentro de los bancos, aprovechando su infraestructura y experiencia en gestión de riesgos, o si es preferible crear agencias independientes que puedan ofrecer evaluaciones crediticias más especializadas y neutrales. Además, se considerarán aspectos como la protección de datos, la transparencia en la toma de decisiones y la capacidad para generar confianza entre las PYMES y las instituciones financieras.

2. Revisión de literatura

En esta investigación se utilizarán las variables open data y machine learning para evaluar cómo los bancos en Colombia pueden implementar un sistema de scoring alternativo para mejorar la inclusión financiera de las PYMES en el sector comercial. Open data ofrece una oportunidad valiosa al proporcionar conjuntos de datos accesibles que pueden enriquecer los modelos tradicionales de evaluación crediticia, mientras que machine learning permite procesar y analizar estos datos para generar predicciones más precisas y eficientes. Estas dos tecnologías, al ser combinadas, tienen el

potencial de transformar los métodos de evaluación de riesgo crediticio, permitiendo a los bancos tomar decisiones más informadas y justas. El enfoque en open data reduce la dependencia de los historiales financieros formales, brindando mayor acceso al crédito a aquellas PYMES que tradicionalmente han sido excluidas del sistema financiero.

La utilización de open data como fuente de información alternativa, junto con la capacidad de machine learning para aprender y mejorar a partir de grandes volúmenes de datos, puede conducir a la creación de modelos de scoring crediticio más inclusivos. Esta investigación tiene como objetivo explorar cómo estas tecnologías pueden mejorar el acceso al crédito para las PYMES en el sector comercial de Colombia, permitiendo un mayor desarrollo económico y reduciendo las barreras que estas empresas enfrentan para acceder a financiamiento.

Open data se refiere a los conjuntos de datos que son accesibles públicamente y pueden ser utilizados y reutilizados sin restricciones legales o financieras. Estos datos pueden provenir de múltiples fuentes, como gobiernos, empresas privadas, y registros comerciales, y suelen incluir información no estructurada como datos sobre transacciones, comportamiento en redes sociales, y datos económicos (Kitchin, 2014). En el contexto de los servicios financieros, open data se puede utilizar para complementar las fuentes tradicionales de información, brindando a los bancos una visión más completa y diversa del comportamiento crediticio de las PYMES, especialmente aquellas que no cuentan con un historial crediticio formal (Narayan et al., 2020).

El uso de open data en la evaluación del riesgo crediticio ha sido explorado en varios estudios recientes. Asobancaria en el 2022 demostraron que la incorporación de datos abiertos permite mejorar la precisión de los modelos de scoring crediticio al proporcionar información adicional que no se encuentra en los registros financieros tradicionales. Esto es particularmente importante para las PYMES, ya que su acceso al crédito suele estar limitado por la falta de historial financiero. Además, Bizer, Heath y Berners-Lee (2009) señalaron que el uso de open data fomenta la transparencia y la

competencia en el sector financiero, lo que facilita una evaluación de riesgos más inclusiva y equitativa.

La medición de open data en esta investigación se basará en la capacidad de los bancos para integrar estos datos no estructurados en sus modelos de evaluación de riesgo crediticio y en cómo esto afecta las tasas de aprobación de crédito para las PYMES.

Machine learning es una subdisciplina de la inteligencia artificial que permite a los sistemas aprender de los datos, identificar patrones y hacer predicciones sin ser programados explícitamente. En el contexto financiero, machine learning se utiliza para procesar grandes volúmenes de datos, como los proporcionados por open data, y generar modelos predictivos que mejoran la evaluación de riesgo crediticio. Los algoritmos de machine learning permiten a los bancos identificar patrones en los datos de las PYMES que no serían visibles con los métodos tradicionales de scoring, mejorando así la precisión y eficiencia en la toma de decisiones crediticias (Bhardwaj, 2020).

Varios estudios han demostrado el impacto positivo de machine learning en la evaluación crediticia. Kannan et al. (2023) encontraron que los algoritmos de machine learning pueden reducir significativamente las tasas de impago al predecir con mayor precisión el comportamiento futuro de los solicitantes de crédito. Además, Chen et al. (2020) señalaron que el uso de machine learning permite a los bancos procesar grandes volúmenes de open data para generar modelos de scoring alternativos que son más inclusivos y que ofrecen acceso a financiamiento a empresas que tradicionalmente no tienen acceso a crédito.

En términos de medición, el rendimiento de machine learning se evalúa a través de métricas como la precisión, el recall y la curva ROC (Receiver Operating Characteristic), que permiten medir la capacidad predictiva de los modelos (Baker & Wurgler, 2020). Estas métricas ayudarán a determinar cómo los modelos de machine learning basados en open data pueden mejorar la inclusión financiera de las PYMES en el sector comercial colombiano.

2.1 Open data

Definición

El open data se refiere a información que se encuentra disponible públicamente para su uso, modificación, y redistribución sin barreras legales, tecnológicas o financieras. Estos datos provienen de diversas fuentes, como gobiernos, empresas privadas, redes sociales, registros comerciales, y transacciones digitales (MinTIC, n.d.). En el sector financiero, open data se ha convertido en una herramienta poderosa para complementar los métodos tradicionales de evaluación crediticia, especialmente para actores que no cuentan con un historial financiero sólido, como muchas PYMES en Colombia. La inclusión de estos datos en los modelos crediticios permite a las instituciones financieras generar una visión más integral del riesgo crediticio y las capacidades financieras de estas empresas, lo que potencialmente mejora la precisión de las decisiones de crédito y promueve la inclusión financiera (Narayan et al., 2020).

En el caso de Colombia, open data presenta una oportunidad significativa para mejorar la inclusión financiera de las PYMES, especialmente aquellas en el sector de comercio, donde el acceso al crédito es limitado. Las instituciones financieras pueden aprovechar una variedad de datos públicos, como registros de comportamiento empresarial, interacciones en redes sociales, datos macroeconómicos y sectoriales, para evaluar de manera más precisa la capacidad crediticia de estas PYMES. Esto es especialmente relevante en un entorno donde la informalidad y la falta de historial financiero representan barreras importantes para el acceso a financiamiento (Muñoz & Vélez, 2017).

Origen teórico

El concepto de open data está basado en la teoría de sistemas abiertos, que sostiene que las organizaciones deben interactuar con su entorno para adaptarse y evolucionar. En el ámbito financiero, esta teoría se traduce en la necesidad de las instituciones de aprovechar múltiples fuentes de datos para mejorar la toma de decisiones, especialmente en la evaluación de riesgo crediticio. Al acceder a datos abiertos provenientes de diversas fuentes, los bancos pueden mejorar su capacidad

para identificar patrones de comportamiento, lo que les permite evaluar mejor el riesgo de las PYMES que carecen de un historial financiero formal (Kitchin, 2014). Esta teoría también resuena con los principios de la economía digital, donde el acceso a datos abiertos fomenta la innovación, la transparencia, y la competitividad en el mercado financiero.

Por otro lado, la teoría del conocimiento abierto propone que el acceso público a los datos fomenta una sociedad más inclusiva y equitativa, al permitir que más actores puedan utilizar esta información para tomar decisiones más informadas incluyendo a las instituciones financieras (Bizer, Heath, & Berners-Lee, 2009). En el contexto de la evaluación de riesgo crediticio, estudios como el de Asobancaria en el 2024 destacan que el uso de open data puede mitigar los sesgos en las evaluaciones tradicionales, proporcionando una herramienta más justa para evaluar la capacidad crediticia de las PYMES.

Definición operacional

En esta investigación, open data se entiende como un recurso crítico para mejorar el acceso de las PYMES al financiamiento en Colombia, particularmente en el sector comercial. Los datos abiertos permiten a los bancos y otras instituciones financieras acceder a una gama de información no tradicional, como registros comerciales, interacciones en redes sociales, y datos sectoriales, que enriquecen los modelos de evaluación crediticia. Este uso de datos alternativos ofrece una visión más completa de la situación financiera y operativa de las PYMES, lo que permite a los bancos tomar decisiones de crédito más informadas y reducir la dependencia de los métodos tradicionales basados exclusivamente en estados financieros históricos y colaterales.

El uso de open data en los modelos crediticios para PYMES se fundamenta en la capacidad de estas fuentes de datos para proporcionar información valiosa que los bancos no podrían obtener de otra manera. Por ejemplo, los datos abiertos pueden incluir:

- Registros gubernamentales: Información sobre la constitución y licencias de las PYMES, lo que permite evaluar la formalidad y legalidad de sus operaciones.

- Interacciones en redes sociales y plataformas digitales: Datos sobre la reputación y las relaciones comerciales de las PYMES con sus clientes y proveedores, proporcionando una visión indirecta de su estabilidad financiera.
- Datos macroeconómicos: Información sobre el sector comercial en el que operan las PYMES, permitiendo una evaluación más contextualizada del riesgo.

Al utilizar estos datos alternativos, las instituciones financieras pueden crear modelos de riesgo más inclusivos que benefician a las PYMES que anteriormente no podían acceder al financiamiento debido a la falta de un historial crediticio formal.

La efectividad del uso de open data en la mejora de la inclusión financiera se mide a través de varios indicadores clave que permiten evaluar el impacto de esta variable en el proceso de evaluación crediticia:

- Diversidad y volumen de fuentes de datos:** Se mide el número de fuentes de datos abiertos que se integran en los modelos de evaluación crediticia. Un mayor número y diversidad de fuentes puede generar una visión más completa de la situación financiera y operativa de las PYMES.
- Precisión de los modelos predictivos:** Un indicador fundamental es la precisión con la que los modelos de scoring alternativo predicen la probabilidad de impago. Al incluir open data, se espera que los modelos mejoren su capacidad predictiva, lo que puede medirse utilizando métricas como precisión, recall y la curva ROC (Bhadwaj , 2020).
- Tasa de aprobación de créditos:** Se evalúa si el uso de open data aumenta la tasa de aprobación de créditos a PYMES que anteriormente no califican bajo los modelos tradicionales. Este indicador refleja el impacto de la inclusión de datos alternativos en la accesibilidad al crédito (Chen et al., 2020).
- Reducción de costos operativos:** El uso de open data y la automatización de los procesos de evaluación crediticia pueden reducir los costos asociados al análisis de riesgo, ya que los bancos pueden acceder a datos relevantes sin tener que depender de métodos costosos de recolección de información (Alam, N., Gupta, L., & Zameni, A. 2025).

La inclusión financiera se evalúa a través de indicadores clave que reflejan el acceso, uso y calidad de los servicios financieros disponibles para las PYMES. Entre estos, se encuentra el porcentaje de

PYMES con cuenta bancaria, que mide el acceso básico al sistema financiero formal, y el volumen y valor de créditos otorgados, que permite evaluar el crecimiento del financiamiento para este sector. También se consideran la tasa de aprobación de créditos, que refleja las oportunidades de acceso al financiamiento, y el uso de productos financieros digitales, indicador que mide la modernización y digitalización de las operaciones empresariales. Además, se monitorea la frecuencia de transacciones realizadas por las PYMES en el sistema financiero formal, como depósitos, pagos y transferencias, así como la composición de la cartera de crédito según el tipo de empresa, lo que ayuda a identificar áreas subatendidas. Asimismo, se supervisan las tasas de interés y las condiciones de crédito, permitiendo analizar si estas han mejorado como resultado del uso de open data. Otros indicadores relevantes incluyen el índice de exclusión financiera, que mide la proporción de empresas sin acceso a servicios financieros formales, y la diversidad y accesibilidad de productos diseñados para PYMES, como microcréditos y líneas específicas. Finalmente, se evalúa el impacto económico del financiamiento, considerando el crecimiento de ingresos, generación de empleo y sostenibilidad financiera de las PYMES beneficiadas. Estos mecanismos permiten una evaluación integral de la inclusión financiera, ayudando a identificar barreras y oportunidades para promover un acceso más equitativo al sistema financiero en Colombia.

Implementación y beneficios

La implementación de open data en los modelos de scoring alternativo presenta numerosos beneficios tanto para las instituciones financieras como para las PYMES: Acceso a nuevas fuentes de información: Al utilizar datos abiertos, los bancos pueden reducir la dependencia de información financiera tradicional y crear un perfil más completo de las PYMES que carecen de historial crediticio. Esto es especialmente beneficioso para aquellas empresas que operan en el sector informal o que están en etapas tempranas de desarrollo. Mejora en la precisión del scoring: Los modelos de scoring que integran open data pueden ofrecer una evaluación más precisa del riesgo crediticio, lo que mejora la capacidad de los bancos para identificar oportunidades de financiamiento seguras y rentables (Kannan et al., 2023).

Fomento de la transparencia y la competencia: El acceso a datos abiertos impulsa la transparencia en el sector financiero, lo que promueve la competencia y permite a las PYMES acceder a mejores condiciones de financiamiento. Este proceso también fomenta la innovación en el desarrollo de productos financieros personalizados para diferentes segmentos de mercado (Baker & Wurgler, 2020).

Inclusión financiera para PYMES: Las PYMES en sectores subatendidos, como aquellas ubicadas en áreas rurales o emergentes, se benefician directamente del uso de open data, ya que esta metodología permite una evaluación de riesgo más inclusiva.

2.2 Machine learning:

El machine learning aplicado a las finanzas hace referencia a la intersección de la inteligencia artificial y la teoría financiera moderna. El padre del machine learning aplicado a sistemas algorítmicos de datos se llama Claude Shannon, el cual introdujo la base matemática para el procesamiento y análisis en 1948. Shannon introdujo conceptos como la entropía de la información, que es un concepto crucial para entender cómo los algoritmos de machine learning pueden extraer patrones de grandes volúmenes de datos (Shannon, 1948). A partir de esta teoría surgieron muchas otras, pero la que más ha impulsado este modelo al uso que buscamos aplicar con esta variable en la investigación es la de "Aprendizaje Estadístico" propuesta por Vladimir Vapnik y Alexey Chervonenkis en la década de los 60. Esta proporciona el marco matemático para entender cómo los algoritmos pueden aprender de los datos y hacer predicciones precisas, el cual se aplica hoy en día en algoritmos de machine learning aplicados en finanzas, conocidos como Máquinas de Vectores de Soporte (SVM) y las redes neuronales (Vapnik, 1999).

En virtud de la investigación, el machine learning aplicado a las finanzas se define como el uso de algoritmos y modelos computacionales que aprenden de los datos para mejorar de manera automática su rendimiento en tareas de análisis cuantitativo en situaciones específicas, en este caso aplicado en modelos financieros de riesgo crediticio. Estos modelos utilizan técnicas de aprendizaje supervisado y no supervisado para analizar grandes conjuntos de datos, identificar patrones y hacer predicciones o tomar decisiones sin ser programados explícitamente para cada escenario. En esta investigación se

utilizará para entender cómo el machine learning puede utilizarse como una herramienta que permite a las instituciones financieras incorporar y procesar eficientemente datos no tradicionales (open data) para evaluar la solvencia y el potencial de crecimiento de empresas.

De esta manera se podrá hacer un análisis de riesgo crediticio con el uso de modelos matemáticos y procesamiento de datos masivos que permita abrir la posibilidad de evaluar empresas que son excluidas del sistema financiero por los modelos de scoring tradicionales. Según George Akerlof, Michael Spence y Joseph Stiglitz, la falta de información completa entre prestamistas y prestatarios puede llevar a ineficiencias en el mercado crediticio. En su famoso artículo "The Market for Lemons" demuestran cómo la asimetría de información puede llevar a selección adversa en los mercados (Akerlof, 1970). Teniendo en cuenta lo anterior, se respalda la definición propuesta dando a entender cómo el machine learning es una herramienta para alinear datos masivos de manera simétrica y generar evaluaciones más precisas no accesibles por el análisis humano.

La variable se puede medir de diversas maneras, pero hay cuatro métricas fundamentales dentro del contexto que se desea aplicar. La principal a considerar es la medición en la aprobación de créditos, esta hace referencia a medir el incremento de créditos aprobados por el algoritmo frente a la tasa de morosidad. La otra métrica sería la precisión en el score crediticio; según el Banco de la República, un modelo de scoring alternativo efectivo debe alcanzar un área bajo la curva del ROC (AUC-ROC) de al menos 0.80 y una precisión balanceada superior al 75% (Banco de la República, 2023). Estas dos métricas serán fundamentales para evaluar la capacidad del modelo de identificar PYMES de alto y bajo riesgo crediticio y asimismo identificar su impacto en dicho mercado financiero.

De no ser posible medir estas dos métricas de la variable de forma precisa, el modelo de machine learning podría ser una amenaza para la salud financiera de la entidad financiera que la implemente, presentando mayor riesgo de pérdidas por default de pagos y asimismo un impacto negativo en las PYMES por favorecer altos niveles de endeudamiento sin tener una capacidad de pago, representando riesgos regulatorios y financieros para aquellos que esta investigación busca apoyar mediante una solución crediticia. Por último se debe poder medir esta variable con uno de los conceptos más

importantes dentro del marco regulatorio, y es el cumplimiento con la Circular Externa 029 de 2014 de la Superintendencia Financiera de Colombia, que exige las entidades financieras puedan explicar las decisiones de sus modelos de riesgo crediticio (Superintendencia Financiera de Colombia, 029/2014). Está cumplimiento se puede medir con SHAP (shapley additive explanations) y LIME (local interpretable model-agnostic explanations) para dar confiabilidad del cumplimiento de la métrica regulatoria del modelo de machine learning como uso de scoring alternativo.

2.3 Estado del arte:

En los últimos años, la combinación de open data y machine learning ha ganado protagonismo en el ámbito del análisis crediticio, particularmente en el desarrollo de sistemas de scoring alternativo. Estas tecnologías podrían permitir a las instituciones financieras evaluar el riesgo crediticio de las PYMES de manera más inclusiva y precisa, superando las limitaciones de los modelos tradicionales basados en datos financieros formales. La integración de open data y machine learning facilita una evaluación más detallada y precisa de los solicitantes de crédito, proporcionando a los bancos herramientas más sofisticadas para medir el riesgo crediticio y mejorar la inclusión financiera.

Open data hace referencia a conjuntos de datos accesibles públicamente que pueden ser utilizados para diferentes propósitos, incluidos los análisis financieros. En el contexto de credit scoring, open data incluye información sobre transacciones comerciales, comportamientos en redes sociales, datos macroeconómicos, y registros públicos, que pueden proporcionar información adicional sobre las PYMES que carecen de un historial crediticio formal. Kumar et al. (2021) destacan que la utilización de open data permite a los bancos acceder a fuentes de datos complementarias que pueden enriquecer los modelos de evaluación crediticia y proporcionar una imagen más completa del riesgo asociado a un solicitante de crédito.

La investigación de Bello O. A (2024) sugiere que los datos abiertos, como registros comerciales y datos macroeconómicos, pueden utilizarse para crear perfiles de riesgo más detallados y precisos, que van más allá de los simples historiales financieros. Al integrar estos datos, los bancos pueden

identificar patrones de comportamiento crediticio en las PYMES, lo que les permite otorgar créditos a empresas que de otro modo no habrían calificado bajo modelos tradicionales. Además, Bounoua et al. (2020) argumentan que el uso de open data reduce el sesgo en la evaluación crediticia, ya que proporciona una perspectiva más completa de los prestatarios, lo que promueve la equidad y la transparencia en el acceso al crédito.

Por otro lado, un desafío destacado por Kshetri (2020) es que la calidad y disponibilidad de los datos abiertos pueden variar. Sin embargo, con una adecuada infraestructura y regulación, open data puede ser una herramienta eficaz para mejorar el scoring crediticio, al permitir a las instituciones financieras analizar una gama más amplia de factores en la evaluación del riesgo crediticio de las PYMES.

Machine learning ha revolucionado la forma en que las instituciones financieras abordan la evaluación crediticia. Esta tecnología permite analizar grandes volúmenes de datos, identificar patrones y generar predicciones precisas sobre el comportamiento crediticio. En el contexto de credit scoring, machine learning facilita el procesamiento de datos no estructurados y semi-estructurados, como los obtenidos de open data, para predecir de manera más precisa la probabilidad de incumplimiento de un prestatario.

Un estudio de Zhai et al. (2022) demostró que el uso de machine learning en modelos de scoring crediticio ha permitido mejorar la precisión de las predicciones de riesgo, lo que ha llevado a una mayor eficiencia en la asignación de crédito. Además, estos modelos son capaces de procesar grandes cantidades de open data, lo que permite identificar patrones de comportamiento crediticio que no son evidentes con los métodos tradicionales. Ali et al. (2022) señalan que los algoritmos de machine learning pueden aprender de los datos históricos y aplicar estos aprendizajes a nuevos solicitantes de crédito, lo que mejora la toma de decisiones y minimiza los errores en la evaluación crediticia.

Por su parte, Bagnall et al. (2019) resaltan que machine learning no solo aumenta la precisión del scoring crediticio, sino que también permite la personalización de las ofertas de crédito. A través del

análisis predictivo de machine learning, los bancos pueden adaptar sus productos financieros a las necesidades específicas de las PYMES, optimizando tanto la experiencia del cliente como la rentabilidad del banco. Esto es posible gracias a la capacidad de los algoritmos de aprendizaje automático para detectar patrones complejos en los datos abiertos, proporcionando una visión más rica y profunda del perfil de crédito del solicitante.

La combinación de open data y machine learning ha demostrado ser una solución innovadora para mejorar la inclusión financiera de las PYMES. Al integrar ambas tecnologías, los bancos pueden desarrollar modelos de scoring alternativos que no dependen únicamente de los historiales financieros tradicionales. Según Bitch et al. (2021), el uso de open data en conjunto con machine learning permite que las instituciones financieras creen modelos de riesgo más inclusivos, facilitando el acceso al crédito para las PYMES que han sido excluidas tradicionalmente del sistema financiero.

Zhai et al. (2022) encontraron que la combinación de open data con algoritmos de machine learning no solo mejora la precisión de las predicciones de riesgo crediticio, sino que también reduce la dependencia de datos financieros formales. Esto es especialmente relevante para las PYMES que no cuentan con un historial crediticio sólido. El uso de open data permite a los modelos de Machine learning extraer información clave de fuentes no convencionales, lo que mejora la evaluación del riesgo y permite a los bancos tomar decisiones de crédito más informadas.

Por último, Kitchin (2020) subraya que la adopción conjunta de open data y machine learning transforma el panorama del credit scoring al permitir que los bancos adapten sus modelos de riesgo a las características y necesidades de las PYMES. Esta combinación tecnológica no solo mejora la precisión y equidad en la asignación del crédito, sino que también fomenta la transparencia y confianza en los sistemas de evaluación crediticia, promoviendo un entorno financiero más inclusivo y eficiente.

2.4 Hipótesis:

“La dependencia de modelos tradicionales basados en historial crediticio tiene una relación negativa con el acceso al crédito formal para PYMES del sector comercio en Colombia.”

Esta hipótesis parte del reconocimiento de que los modelos tradicionales de evaluación crediticia en Colombia están profundamente anclados en el análisis de variables como el historial financiero previo, los estados financieros auditados y la disponibilidad de garantías reales. Si bien estos enfoques han sido históricamente efectivos para grandes empresas consolidadas, resultan excluyentes para un amplio segmento de PYMES, particularmente aquellas que operan en entornos informales, no cuentan con suficiente historial crediticio o se encuentran en etapa temprana de desarrollo.

Diversos estudios, como el informe del (Banco de la República, 2023) y el (BID 2023), evidencian que este tipo de modelos perpetúan la exclusión financiera al limitar el acceso a crédito formal a empresas que no cumplen con estos requisitos estructurales, a pesar de tener flujos operativos viables. Las PYMES del sector comercio son especialmente vulnerables a este sesgo, ya que muchas de ellas operan con prácticas contables rudimentarias, no están registradas formalmente, o han recurrido históricamente a fuentes informales de financiamiento.

La relación negativa entre el uso exclusivo de estos modelos y el acceso al crédito se manifiesta en las bajas tasas de aprobación de solicitudes, condiciones poco favorables, y la necesidad de recurrir a financiamiento informal con mayores costos y riesgos. Esto no solo afecta la capacidad de crecimiento y formalización de las PYMES, sino que también limita el potencial de desarrollo económico y generación de empleo en el país.

Esta hipótesis permite analizar empíricamente si la presencia de barreras estructurales en los sistemas tradicionales de evaluación está directamente asociada con un menor nivel de inclusión financiera en este segmento empresarial, estableciendo una base comparativa para contrastar la efectividad de modelos más inclusivos, como el scoring alternativo basado en open data.

3. Marco teórico:

3.1 Conceptualización de la inclusión financiera: de la teoría a la práctica

La inclusión financiera ha cambiado su definición tradicional que se conoce como una simple posesión de cuentas bancarias a un nuevo entendimiento en un modelo integral que abarca acceso, uso, calidad y sostenibilidad de servicios financieros. Según el el banco mundial en el 2022, este concepto combina tres dimensiones clave: Acceso (disponibilidad de servicios), uso (adopción y entendimiento de los productos) y calidad (oportunidad y protección al consumidor). En el contexto de las PYMES en colombia, implica no solo disponer de créditos formales, sino también su extensión para acceder a seguros agrícolas, plataformas de pagos electrónicos y herramientas de gestión financiera (OECD, 2021).

En América Latina, la inclusión financiera ha seguido una trayectoria definida por crisis económicas y avances tecnológicos. Durante la década de 1990, las reformas neoliberales impulsaron la bancarización masiva, pero con enfoques excluyentes en su definición en la actualidad hacia sectores rurales y PYMES (CEPAL, 2018). Después de la crisis del 2008 se fomentaron modelos más inclusivos. Brasil implementó el programa Bolsa familiar para vincular transferencias sociales con cuentas bancarias, mientras que México desarrolló una ley de instituciones de tecnología financiera en el 2018 conocida como la ITF para regular plataformas digitales (BID,2020). Colombia por su parte logró combinar políticas de garantías estatales (Fondo Nacional de Garantías) con innovaciones como el sistema de pagos rápidos Bre-B, inspirado en Pix de Brasil y UPI en India (MinHacienda, 2023).

La inclusión financiera se considera un acelerador para los ODS (Objetivos de Desarrollo Sostenible). El ODS 1: Fin de la pobreza, se vincula con el acceso a microcréditos para emprendimientos rurales, mientras que el ODS 5: Igualdad de género, exige eliminar barreras para mujeres empresarias que se ven afectadas por la alta exclusión financiera quienes representan el 34% de las PYMES Colombianas pero solo reciben el 22% del crédito formal (ONU Mujeres, 2023). Asimismo, el ODS 9: Industria, innovación e infraestructura, resalta la necesidad de crear nuevas infraestructuras digitales que

fomenten la inclusión en zonas rurales, donde el 65% de las PYMES carecen de conectividad estable (MinTIC, 2023).

3.2 Barreras estructurales: un análisis multidimensional

El modelo sobre racionamiento crediticio de Stiglitz y Weiss (1981) sigue vigente en Colombia. En promedio las entidades financieras exigen 12 documentos para aprobar un crédito a una PYME que al comparar con Chile en el cual son 5 se ve una gran barrera al crédito, lo que desincentiva a empresas informales o sin historial crediticio según la Superintendencia Financiera en el 2023. Este fenómeno afecta desproporcionadamente a PYMEs en áreas rurales, en las cuales el 78% no cuentan con un estado financiero auditado y el 63% depende de préstamos informales con tasas de interés de hasta 30% mensual según el DANE en el 2023. Adicionalmente la falta de datos estandarizados afecta negativamente la evaluación de riesgos, mientras el 89% de las PYMES urbanas reportan ventas de manera digital solo el 24% de las rurales lo hacen (ANDI, 2023).

Según el DANE en el 2023 la informalidad es un círculo vicioso, 43% de las PYMES colombianas no cuentan con registro mercantil el cual les impide acceder a productos como el factoring que solo el 8% lo utiliza y el leasing que solo el 4% lo utiliza. Adicionalmente el 92% de sus activos no son aceptados como garantía por la banca tradicional según Asobancaria en el 2022. Esta exclusión obliga a estas PYMES a recurrir a “gota a gota”, en las cuales el 45% de los empresarios de pequeñas empresas son víctimas de extorsión crediticia (Fiscalía General, 2023). La adopción de tecnología también se ve afectada por la informalidad ya que el 71% de estas empresas no utilizan un software contable, y el 58% almacenan datos de manera física las cuales son vulnerables a pérdidas (MinTIC, 2023).

Las diferencias geográficas también tienen un gran impacto. En la Guajira el 88% de las PYMES no tienen acceso a cuentas corrientes, mientras que en Bogotá solo el 11% no tiene (Bancóldex, 2023).

3.3 Innovación tecnológica: de la disrupción a la inclusión

Los pagos en Colombia se han transformado por la revolución digital. El sistema de Transfiya procesó 2.1 billones de pesos en 2023, un aumento del 145% comparando al 2020, está facilitando transacciones instantáneas entre PYMES y proveedores (ACH Colombia, 2024). Adicionalmente plataformas como DaviPlata y Nequi han bancarizado a 8 millones de colombianos permitiendo que las PYMES reciban pagos sin comisiones y acceder a créditos preaprobados basándose en su flujo de efectivo (Superintendencia Financiera, 2023). El sistema lanzado en el 2023 Bre-B permitió transferencias 24/7 usando solo el número de celular el cual fue inspirado en Pix (Brasil) y UPI (India), este ha logrado procesar 43 millones de transacciones mensuales, especialmente en zonas rurales (MinHacienda, 2024).

La *Ley Fintech (Ley 2128 de 2021)* sentó las bases para la banca abierta en Colombia, permitiendo a las PYMES compartir datos financieros entre instituciones bajo su consentimiento. Esto ha permitido el desarrollo de aplicaciones como *SumaMóvil*, que agrega cuentas de múltiples bancos y sugiere estrategias de ahorro basadas en patrones de gasto (Superintendencia Financiera, 2024) los cuales pueden utilizarse como patrones de información para sistemas de scoring alternativo. En el Reino Unido, la banca abierta aumentó la competencia, reduciendo tasas de interés para PYMES en un 18% (FCA, 2022), un efecto esperado en Colombia hacia 2026.

3.4 Conceptualización del scoring alternativo y su relevancia

Las limitaciones de los modelos crediticios permitieron que surja como respuesta el scoring alternativo, ya que estas dependen de historiales crediticios formales, garantías físicas y estados financieros auditados. Este enfoque utiliza datos no convencionales como transacciones comerciales, pagos de servicios públicos, actividad en redes sociales, patrones de consumo digital y métricas de reputación empresarial para evaluar la solvencia de PYMES, especialmente aquellas en la informalidad o sin acceso a bancarización (Banco Mundial, 2022). Investigaciones recientes destacan que el uso de datos alternativos permite a instituciones financieras reducir asimetrías de información y ampliar la cobertura crediticia en un 15-20% en países emergentes (Provenir, 2024).

3.5 Evidencia empírica y validación de modelos

Estudios como el proyecto *Puntajes de crédito alternativos para PYMES* (2014-2024), respaldado por Innovations for Poverty Action, demostraron que entre los modelos basados en características demográficas y lo que están basados en transacciones comerciales predicen el riesgo crediticio (IPA, 2024). En este estudio, se aleatorizó el acceso a créditos utilizando ambos enfoques en países en desarrollo, observándose una reducción del 12% en la morosidad en el grupo evaluado con datos transaccionales (IPA, 2024). Casos como *Mercado Crédito* en Argentina validan esta teoría, que al analizar historiales de compras en MercadoLibre, otorgaron 45 millones de créditos a emprendedores con mercado pago (Mercado Libre, 2023). Además, soluciones de scoring alternativo han mejorado la precisión en la evaluación de riesgo en un 30% (FICO, 2021).

3.6 Desafíos en contextos de alta informalidad

La Informalidad laboral en América latina es de un 48% en un promedio de 11 países (OIT, 2023), las PYMES enfrentan barreras estructurales. Un estudio de la CEPAL (2023) reveló que solo el 12% del crédito total de América latina se destina a pequeñas empresas mientras que en países de la OCDE cuentan con un 25%. La falta de documentación financiera estandarizada y la dependencia del efectivo (85% de transacciones en Centroamérica) limitan la aplicabilidad de modelos tradicionales (BID, 2022). por otra parte iniciativas como TiendaPago en México y Perú utilizan datos de compras de inventario para otorgar financiamiento a comercios informales, demostrando que los algoritmos de *machine learning* pueden identificar patrones de solvencia incluso en ausencia de historiales formales (TiendaPago, 2023) adi.

3.7 Tecnologías habilitadoras: IA, blockchain y banca abierta

La inteligencia artificial está revolucionando el scoring alternativo, fintec como Addi en Colombia ya emplean IA para analizar huellas digitales, entre estas los datos que utilizan son el historial de pagos en Rappi, Uber o plataformas de comercio electrónico (PayU), la reputación comercial a través de reseñas en redes sociales y el cumplimiento de obligaciones tributarias, así como también utilizan

datos ambientales para evaluar el impacto de fenómenos climáticos en las cosechas de PYMES agrícolas, con esto redujeron la morosidad al 4.5% (Addi, 2023). En India, también se evidenció el sistema Aadhaar (tarjeta de identificación nacional de la India) que permite bancarizar el a 450 millones de personas mediante identificación biométrica, un modelo replicable para PYMEs rurales (Reserve Bank of India, 2022). Además, la tendencia en banca abierta (Open Banking), facilita el intercambio seguro de datos entre instituciones financieras, permitiendo scoring basado en flujos de efectivo multisistema (Superintendencia Financiera de Colombia, 2023).

3.8 Fintechs en Colombia

En este contexto, las fintechs son actores clave en la transformación del acceso al crédito para las PYMES en Colombia. Ya que estas empresas de base tecnológica han logrado desarrollar soluciones financieras de manera rápida, accesible y con una neutralidad que la permite centrarse en el usuario, y tienen la habilidad de romper con las barreras que impone el scoring crediticio tradicional. El uso de datos alternativos alimentados de plataformas digitales junto a machine learning, permite que las fintechs ofrezcan procesos de originación crediticia más rápidos, con menor documentación y mejor adaptabilidad al perfil real de las empresas. La capacidad de las fintech para innovar en productos financieros, como créditos inteligentes, líneas rotativas digitales o factoring automatizado permite ampliar la cobertura a segmentos excluidos históricamente.

3.9 Lecciones globales: casos de éxito y adaptabilidad

Volviendo a lo comentado anteriormente, el sistema de identificación biométrica Aadhaar en India, con más de 1,300 millones de usuarios registrados ha sido clave para una transformación financiera inclusiva sin precedentes en la historia del país. Además, este sistema ha facilitado el desarrollo de créditos comunitarios, en los que los préstamos grupales se otorgan con base en la reputación colectiva, sustituyendo las garantías físicas tradicionales (Reserve Bank of India, 2023). Este sistema ofrece lecciones valiosas para Colombia, donde la integración de datos del Registro Civil, la DIAN y la Cámara de Comercio en una plataforma digital unificada podría permitir la verificación de

identidad de manera ágil y segura, abriendo la puerta a créditos sin papeleo ya una mayor inclusión financiera para las PYMES.

También existe el caso de M-Pesa, una billetera móvil que utiliza el 80% de la población en Kenia que ha revolucionado el sistema financiero del país al procesar alrededor del 60% del PIB nacional a través de transacciones digitales (Stadler, C. 2024). Su éxito se centra en mecanismos innovadores tales como créditos preaprobados vía SMS que otorgan préstamos de hasta USD 500 en un tiempo aproximado de apenas 90 segundos, el sistema se alimenta de historial de recargas del usuario como referencia (N26 Magazine. 2024). Este sistema es un modelo que se puede replicar en Colombia, donde operadores móviles como Claro y Tigo podrán mediante alianzas con entidades financieras poder implementar soluciones similares.

3.10 Políticas públicas: logros, limitaciones y recomendaciones

Entre 2020 y 2023, el fondo nacional de garantías (FNG) respaldó créditos por un valor alrededor de 48 billones de pesos, beneficiando a aproximadamente a más de 500 mil PYMES. Por otro lado, su estructura presenta limitaciones que han restringido su impacto. En términos geográficos, más del 60% de los recursos se concentraron en Bogotá, Valle del Cauca y Antioquia dejando otros departamentos más vulnerables con poca visibilidad económica (FNG, 2024). Asimismo, el decreto 2654 de 2014 marcó un hito en la regulación del sector financiero digitalizado en Colombia, se logró identificar las fuentes en tres principales categorías: pagos, crédito e inversión, y establecer un requisito de capital mínimo de USD 200.000 exigidos para operación con relación a préstamos. Esto logró un crecimiento de aproximadamente del 300% de fintech registradas, que pasaron de 120 en 2021 a 480 en 2024, con la ayuda de una reducción del 40% en costos de transacción para PYMES (Colombia Fintech, s.f.).

4. Metodología de la investigación:

4.1 Tipo y diseño de la investigación

En el presente estudio se utilizó un enfoque mixto que combina investigación exploratoria y descriptiva. Es exploratoria porque se investigó un modelo relativamente nuevo y poco estudiado en el contexto del sector crediticio en Colombia; descriptiva porque se buscó caracterizar las principales barreras de acceso al crédito para PYMES, y su impacto en la inclusión financiera y porque se logró establecer la relación entre el uso de estas tecnologías y la mejora en el acceso al crédito para las PYMES.

4.2 Población y muestra

Para este estudio se seleccionó el sector de PYMES de comercio en Colombia, considerando que en el país hay un estimado de 1,7 millones de empresas formales activas y 5,7 millones de informales (Ospina, 2024). De las empresas formales, el 94% son PYMES, las cuales constituyeron la población objetivo de la investigación. Sin embargo, dado el enfoque específico en el sector comercio, se delimitó la población a aproximadamente 200.000 PYMES de comercio formalmente registradas en áreas urbanas, con especial atención a ciudades capitales y principales centros económicos como Bogotá, Medellín, Cali y Barranquilla. Esta delimitación se justifica por la mayor disponibilidad de datos y la relevancia de estas regiones en la actividad económica del país.

La muestra se calculó utilizando un nivel de confianza del 95% y un margen de error del 10%, lo que arrojó un tamaño de muestra de 97 empresas. Este cálculo se hizo con base a la población de 200.000 PYMES de comercio y se realizó utilizando la fórmula de muestreo aleatorio simple ($n = (N \times Z^2 \times p \times q) / [(e^2 \times (N-1)) + (Z^2 \times p \times q)]$), donde se utilizó $Z=1.96$ para el nivel de confianza, $p=q=0.5$ para máxima variabilidad, resultando en $n=97$.

4.3 Instrumentos de recolección

Para la recolección de datos, se diseñó una encuesta de 22 preguntas (ver Anexo 1), dirigida a empresarios de PYMES del sector comercial en Colombia, esta fue diseñada en base a los objetivos específicos de la investigación y tomando en cuenta una revisión bibliográfica de estudios similares sobre inclusión financiera y scoring alternativo. La cual se dividió en 4 partes iguales:

1. **Datos generales:** Caracterización de la empresa (ubicación, tamaño, actividad económica, ingresos).
2. **Acceso al crédito formal:** Identificación de barreras, experiencias de rechazo y uso de alternativas financieras.
3. **Percepción sobre modelos tradicionales:** Percepción de las PYMES sobre los requisitos documentales y garantías.
4. **Apertura hacia un nuevo modelo:** Disposición para compartir datos y confianza a un nuevo sistema

La encuesta se administró de manera digital mediante google forms, enviada directamente a las PYMES asegurando accesibilidad y facilitando el análisis automatizado de respuestas.

4.4 Procedimientos

El proceso de recolección de datos se llevó a cabo en tres etapas:

Selección y contacto inicial: De una base de datos extraída de EMIS se extrajo información de 10,000 PYMES formales, luego estas fueron filtradas para tener solo correos válidos y actualizados, una vez realizado se enviaron a 3,000 invitaciones para llenar la encuesta acompañadas de una carta de presentación institucional que explicaba los objetivos, confidencialidad y avales académicos del estudio.

Seguimiento y ampliación de la muestra: Para mitigar la baja tasa de respuesta se implementó una segunda ronda de contacto telefónico a empresas seleccionadas aleatoriamente, adicional a esto se realizaron 4 rondas de recordatorios semanales vía correo electrónico de marzo a abril de 2025.

Validación y análisis de datos: Las 98 respuestas válidas obtenidas se filtraron para filtrar inconsistencias, y luego los datos se pasaron a un excel en el cual se analizaron a mayor detalle

5. Análisis de los resultados:

La encuesta realizada nos permitió identificar barreras estructurales y percepciones actuales sobre los modelos de evaluación crediticia en las PYMES del sector de comercio en Colombia. A través de una encuesta estructurada de 22 preguntas aplicada a múltiples PYMES, se recopilieron datos de 98 respuestas relevantes que validan la hipótesis de investigación y aportan evidencia empírica al cuerpo teórico revisado previamente. La encuesta fue diseñada para entender tanto las dificultades experimentadas por las PYMES al acceder a créditos formales como su disposición frente a modelos de scoring alternativo basados en open data y machine learning.

Uno de los hallazgos más significativos fue que el 73% ver anexo 2 de los encuestados manifestó haber enfrentado al menos una barrera al intentar acceder a créditos formales. Las barreras más mencionadas fueron: "altos costos financieros (tasas de interés elevadas)" (16%), "historial crediticio insuficiente" (21%) y "exceso de trámites burocráticos" (21%) ver anexo 3. Esta información evidencia una desconexión entre las estructuras actuales del sistema financiero y las necesidades reales de las PYMES. Adicionalmente, el 81% ver anexo 4 de los empresarios se mostró en desacuerdo o neutral frente a la afirmación de que estos modelos eran adecuados para evaluar el riesgo de sus empresas. Esta percepción respalda empíricamente la hipótesis central del estudio: "La dependencia de modelos tradicionales basados en historial crediticio tiene una relación negativa significativa con el acceso al crédito formal para PYMES del sector comercio en Colombia".

En cuanto a la apertura hacia nuevas metodologías, el 58% ver anexo 5 de los encuestados indicó que estaría interesado o muy interesado en que su empresa fuera evaluada mediante modelos de scoring alternativo que utilicen datos no tradicionales. El 33% de los empresarios contestó que está muy interesado a que sus empresas fueran evaluada a través de datos no tradicionales (como ventas digitales, pagos a proveedores, o flujos de caja) y el 39% contestó que también estarían dispuestos a compartirlo pero dependiendo del manejo de sus datos ver anexo 6. No obstante, una parte significativa de los participantes expresó que su disposición estaría condicionada a la forma en que los datos fueran utilizados, destacando la importancia de la transparencia y la protección de la información.

Estos resultados contribuyen de manera sustancial a la literatura revisada en los primeros capítulos del trabajo. Autores como Narayan et al. 2020 sostienen que los datos alternativos mejoran la inclusión financiera al proveer información que suple las deficiencias de los modelos tradicionales. Los hallazgos de la encuesta corroboran estas afirmaciones al evidenciar que una parte considerable de las PYMES excluidas por el sistema financiero formal estaría dispuesta a ser evaluada mediante metodologías más inclusivas y tecnológicamente avanzadas. Asimismo, se valida el planteamiento de Bounoua et al. 2020, quien argumenta que el uso de open data reduce el sesgo y promueve una asignación de crédito más equitativa.

Desde un enfoque teórico, los resultados se alinean con la teoría de la asimetría de información expuesta por Akerlof 1970, al mostrar que muchas PYMES con operaciones viables no logran acceder a financiamiento debido a la falta de historial crediticio formal. La implementación de modelos de machine learning alimentados por open data permitiría mitigar estas asimetrías, proporcionando a los bancos una visión más precisa y completa del riesgo asociado a cada solicitante. También se confirma el valor del enfoque de aprendizaje estadístico de Vapnik 1999, según el cual los algoritmos pueden detectar patrones relevantes en grandes volúmenes de datos, permitiendo una evaluación del riesgo más objetiva.

Al evidenciarse que existe interés y disposición para adoptar metodologías de scoring alternativo, se abre una ventana de oportunidad para el rediseño de los mecanismos de evaluación crediticia en Colombia, con miras a fomentar una economía más equitativa y dinámica.

Sin embargo, la aplicabilidad de estos resultados debe contemplarse con cautela. La muestra está compuesta principalmente por empresas urbanas, lo cual podría no reflejar completamente la realidad de las PYMES rurales o informales. Asimismo, aunque existe interés por parte de los empresarios, también se identificó una preocupación por la privacidad de los datos y la necesidad de que los modelos sean explicables. Un 84% de los encuestados manifestó que considera importante o muy importante que los modelos de scoring expliquen cómo se toman las decisiones crediticias, en línea con lo exigido por la Circular Externa 029 de 2014 de la Superintendencia Financiera de Colombia.

En conclusión, los resultados obtenidos fortalecen la hipótesis y los objetivos del estudio, al demostrar que los modelos tradicionales de evaluación de riesgo crediticio excluyen a un amplio grupo de PYMES, mientras que existe una disposición creciente por parte de estas empresas a adoptar modelos alternativos basados en datos no convencionales. Estos hallazgos, además de respaldar el marco teórico planteado, ofrecen un punto de partida para futuras investigaciones, así como para el diseño de políticas públicas e instrumentos financieros orientados a mejorar la inclusión financiera en Colombia.

5.1 Limitaciones de la investigación

En el desarrollo de la investigación, la implementación de la encuesta presentó limitaciones que afectaron la representatividad de los resultados. Se compartió la encuesta a tres mil PYMES registradas en Colombia. Se evidenció una fuerte cautela por parte de los empresarios al ingresar al enlace enviado por correo electrónico, fenómeno atribuible a los altos niveles de fraude electrónico en Colombia, que han generado un alto nivel de desconfianza en la población. Este escepticismo se manifiesta en comportamientos tales como verificación exhaustiva de la legitimidad del estudio antes de responder, esto redujo significativamente la tasa de participación en la investigación. De las tres mil empresas contactadas, únicamente se obtuvieron 98 respuestas válidas, esto reflejó una limitación crítica asociada a la ciberseguridad y la percepción de riesgo entre los destinatarios.

Adicionalmente, se identificaron obstáculos operativos durante la fase de contacto. La base de datos que se utilizó se extrajo de plataformas como EMIS y otros registros institucionales contando con información de diez mil PYMES en Colombia. Estas bases de datos presentaron grandes inconsistencias, aproximadamente dos mil correos electrónicos resultaron inexistentes o desactualizados. Para mitigar este problema, se implementó la estrategia con contactos directos vía telefónica a empresarios previamente identificados. No obstante, esta aproximación también presentó desafíos, como la disponibilidad limitada de los encuestados. Estas limitaciones resaltan la complejidad de recopilar datos sensibles en contextos de alta desconfianza cibernética y cuestionan la necesidad de fortalecer los protocolos de validación de bases de datos para futuras investigaciones.

6. Discusión

La inclusión financiera de las PYMEs puede ser uno de los pilares de mayor importancia para la consolidación de una economía equitativa y sostenibles en países en desarrollo como Colombia. Esta investigación parte del hecho de que dichas empresas no solo consiguen la mayoría del tejido empresarial colombiano, si no que también son responsables de una gran porción del PIB nacional y del empleo formal tanto como informal. En este contexto mejorar su acceso al crédito puede considerarse como una política prioritaria que impacta directamente el desarrollo social económico del país. La investigación reafirma esta idea al evidenciar como los actuales mecanismos de evaluación crediticia basados en modelos tradicionales excluyen de manera sistemática a un amplio espectro de estas empresas.

Uno de los aportes principales del estudio fue la validación de la hipótesis según la cual la dependencia exclusiva de información financiera histórica como estados financieros auditados e historial crediticio está correlacionada negativamente con el acceso efectivo al crédito por parte de las PYMES. Esta relación fue reafirmada por los datos empíricos obtenidos en las encuestas donde un 81% de los empresarios manifestó desconfianza o desacuerdo con la efectividad de los modelos tradicionales. A esto se le puede sumar que el 73% de los encuestados han enfrentado obstáculos en el momento de adquirir un crédito formal, lo que refuerza el argumento de que si existe una brecha estructural entre la lógica operativa del sector financiero y su realidad contable.

También se logró contrastar con la investigación dos caminos potenciales para la implementación del scoring alternativo en Colombia: la adopción directa por parte de los bancos tradicionales y la utilización de agencias independientes especializadas. A través del análisis y la revisión de literatura internacional, se concluye que, al menos en el corto y mediano plazo del contexto Colombiano, la segunda opción representa una vía mas realista y efectiva. Las agencias externas tienen mayor flexibilidad para operar con datos no estructurados, cuentan con neutralidad metodológica y pueden adaptarse de manera más rápida a cambios tecnológicos sin estar atadas a las rígidas estructuras internas de los bancos. Tales casos ya se han evidenciado de manera parcial con Mercado Crédito en Argentina y Addi en Colombia que refuerzan esta tesis, al lograr demostrar que los modelos externos pueden ser más inclusivos en la evaluación de riesgos sin perder su precisión.

Sin embargo, el uso de agencias independientes también presenta limitaciones. Estas entidades, al ser de reciente creación, enfrentan desafíos importantes en términos de sostenibilidad financiera, confianza de los usuarios y cumplimiento normativo. Como se identificó en el análisis regulatorio, la viabilidad de estos modelos está sujeta a marcos jurídicos que aún se encuentran en construcción en Colombia. Aunque leyes de la Superintendencia Financiera han logrado abrir espacios normativos para la innovación de estas empresas, todavía existen muchos vacíos regulatorios en torno al uso de datos sensibles, la trazabilidad algorítmica y la gobernanza de modelos predictivos. La exigencia de que todo modelo de scoring sea explicable, auditado y alineado con los principios de protección al consumidor representan un estándar que muchas fintech aún no logran cumplir de manera exitosa.

Con respecto a lo anterior, se suma una barrera cultural relevante en el país: la desconfianza hacia nuevas tecnologías por parte de los empresarios. Si bien el 58% de los encuestados mostró interés en adoptar metodologías no tradicionales, esta disposición está fuertemente condicionada a factores tales como la transparencia en el uso de datos, la claridad en la toma de decisiones y mitigaciones y aceptación frente a sesgos y errores del modelos. De hecho fue de gran relevancia evidenciar que el 84% de los participantes destacó como muy importante que los modelos expliquen cómo se asigna los puntajes de crédito, el cual con esto se evidencia la preocupación existente de la incomprensión de cómo funciona el machine learning.

Estas limitaciones dan lugar a nuevas inquietudes que exceden el marco original de esta investigación, pero que deben ser consideradas en futuros estudios y en la implementación de estos modelos son que una de las principales preguntas es si estas agencias de scoring alternativo en caso de que se consoliden en Colombia, es sí estas terminan siendo absorbidas por la banca tradicional, corren el riesgo de perder su neutralidad metodológica y de convertirse en simples extensiones de los modelos que originalmente buscan corregir. otra inquietud mira desde el rol del Estado, siendo: ¿están las instituciones gubernamentales dispuestas a promover activamente estos modelos alternativos, o mantendrán su postura histórica de favorecer a los bancos tradicionales mediante políticas de concentración financiera y bajo incentivo a la competencia?

Otra reflexión crítica que surge sobre los riesgos de exclusión dentro de la inclusión es: ¿acaso los modelos de scoring alternativo podrían terminar replicando los sesgos estructurales si estos son entrenados con datos contaminados o no representativos? al igual que lo advierten Narayan et al. (2020), sin una supervisión y gobernanza algorítmica adecuada e independiente, estos modelos podrían volverse opacos o hasta discriminatorios. Esto es especialmente relevante ya que Colombia es un país con altos niveles de informalidad y desigualdad regional, en el cual el diseño de estos modelos debe estar guiado junto a principios de equidad y protección de sectores vulnerables tratando de excluir sesgos en contra de estos.

Desde un punto de vista optimista, los hallazgos de esta tesis sugieren que el scoring alternativo basado en machine learning y open data no solo es algo viable si no positivo siempre que este se diseñe y se implemente bajo un enfoque regulado, ético, progresivo y explicable. En el corto plazo las alianzas entre bancos y agencias independientes pueden representar un modelo de transición que combine lo mejor de ambos mundos sin que uno afecte a otro negativamente: la experiencia institucional y de gestión de riesgo del sector bancario, y la flexibilidad tecnológica y capacidad de innovación de las fintecs. En el mediano y largo plazo el éxito de estos modelos dependerá de gobiernos para garantizar normas favorables, fomentar una cultura de confianza digital y competencia entre empresarios, entidades financieras y consumidores.

Para terminar, la investigación realizada aporta evidencia teórica y empírica que respalda la necesidad de reformar los modelos de evaluación de riesgo crediticios en Colombia. Se ha evidenciado que los sistemas tradicionales centrados en el pasado financiero documentado ya no son suficientes para evaluar el potencial de empresas modernas, especialmente en sectores donde la informalidad y la digitalización coexisten. Al adoptar estos modelos más inclusivos basados en datos actualizados, análisis predictivo y comportamientos reales el sistema financiero colombiano puede dar un salto importante en eficiencia, cobertura y equidad.

7. Conclusión

El scoring alternativo especialmente el que se basa en open data y herramientas algorítmicas como machine learning se presenta como una solución viable y transformadora. La investigación evidenció que los modelos basados en open data permite agrandar las fuentes de información disponibles para evaluar la solvencia de una empresa, incluyendo datos transaccionales, reputaciones y operativos fuera de lo común antes usado en bancos tradicionales. Al integrarse con técnicas de machine learning, es posible construir modelos más precisos y dinámicos para PYMES colombianas.

Asimismo, los hallazgos de la tesis muestran una disposición positiva de parte de los empresarios a ser evaluados mediante esta nueva tecnología, siempre que se demuestre transparencia el cual se logra de manera más efectiva con agencias independientes que trabajen de la mano de los bancos pero que sigan siendo terceridad como se respaldo en la revisión de literatura. Sin embargo estos sistemas en dichas agencias no están exentas de desafíos: vacíos regulatorios, barreras culturales y el riesgos de replicar sesgos estructurales ya fomentados en bancos tradicionales que no fomentan la inclusión y equidad.

Esta tesis aporta una visión crítica y proactiva sobre la inclusión financiera en economías emergentes. No solo propone una alternativa tecnológica si no un replanteamiento del concepto de riesgo, en la cual hay una lógica excluyente y documental hacia un enfoque más integral basado en

comportamientos reales y datos actualizados. Asimismo se logra emprender un camino a un sistema financiero más eficiente, inclusivo y alineado a los retos que han surgido en el siglo XXI.

Para cerrar esta investigación deja sembradas varias inquietudes para el futuro como: ¿Qué rol debe jugar el estado en la promoción e implementación de estos modelos?, ¿Cómo garantizar una supervisión y gobernanza algorítmica independiente y ética?, ¿Cuál es el impacto que tendrán estos modelos en zonas rurales, sectores informales y poblaciones tradicionalmente marginadas? y ¿sí estos modelos terminan siendo absorbidos por la banca tradicional, pueden correr el riesgo de convertirse en simples extensiones de los modelos que originalmente buscan corregir?

8. Recomendaciones a futuras investigaciones

La presente investigación evidencia la necesidad de profundizar en diversas áreas. A partir de los hallazgos y limitaciones encontradas a lo largo de la investigación, se proponen las siguientes recomendaciones para guiar trabajos futuros dentro de esta área de estudio. En primer lugar, es prioritario fortalecer las metodologías de recolección de datos para superar la desconfianza cibernética detectada en los empresarios. Esto podría lograrse mediante la incorporación de mecanismo de verificación de seguridad en las comunicaciones electrónicas, esto se logra a través de enlaces a plataformas para validar que el enlace no tenga virus (por ejemplo, NordVPN). Asimismo, adjuntar un resumen ejecutivo en formato PDF que detalla los objetivos, metodología y avales institucionales del estudio permitiendo validar la legitimidad de la investigación aumentando la confianza con los empresarios.

Los resultados obtenidos en la investigación plantean varias interrogantes que requieren exploración académica con profundidad, como el rol del estado en la promoción de marcos regulatorios que fomenten la adopción de scoring alternativo sin comprometer la estabilidad del sector financiero. Para esto se incluye analizar instrumentos de sandboxes financieros o incentivos fiscales. Adicionalmente, es urgente investigar mecanismos de gobernanza algoritmos que garanticen la transparencia en los modelos alternativos, evitando de esta forma sesgos en poblaciones vulnerables. Otro eje temático es evaluar la viabilidad técnica y social de estos modelos en contextos marginados, como zonas rurales o sectores informales. Paralelamente se recomienda estudiar las dinámicas de la absorción institucional, esto hace referencia si la integración de scoring alternativo en la banca tradicional puede reproducir

exclusiones sistemáticas o, por el contrario, democratizar el acceso al crédito mediante productos inclusivos.

Desde la perspectiva ética y regulatoria, futuras investigaciones deben abordar desafíos vinculados con la privacidad y la equidad. Esto sugiere explorar diseños técnicos y legales que logren equilibrar la innovación con el respeto y cuidado de la privacidad informática. Además, se debe analizar como los modelos alternativos podrían perpetuar desigualdades ya existentes en el sistema si se entrenan con datos históricos sesgados. Con base a lo anterior, se sugiere, desarrollar estudios longitudinales que midan el impacto socioeconómico del scoring alternativo en la formalización, productividad y sostenibilidad de las PYMES a largo plazo, de esta manera se podrá llevar a cabo un análisis comparativo entre países de la región (como México, Argentina o Brasil) que permita identificar buenas prácticas adaptables al contexto Colombiano, considerando diferencias en infraestructura digital y niveles de informalidad.

Finalmente, es importante destacar las colaboraciones interdisciplinarias que integren enfoques de sociología tecnológica, economía conductual y derecho digital tienen que ser fomentadas, ya que esta integración permitirá comprender las barreras culturales, normativas y técnicas que influyen en la adopción de los modelos alternativos de evaluación crediticia. En términos generales estas recomendaciones no solo buscan solucionar las limitaciones indicadas en el estudio sino que también buscan impulsar un debate académico, político y cultural orientado a construir un sistema financiero alineado con los desafíos del siglo XXI.

Referencias

Addi. (2023). Informe de impacto: Tecnología y crédito inclusivo. Recuperado el 13 de febrero de 2025, de <https://co.addi.com/blog/proposito-y-rsc/inclusion-financiera-que-es>

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500. Recuperado el 22 de noviembre de 2024, de <https://www.jstor.org/stable/1879431>

ANDI. (2023). Informes económicos y sectoriales. Recuperado el 7 de marzo de 2025, de <https://www.andi.com.co/Uploads/Balance%202023%20Perspectivas%202024.pdf>

Alam, N., Gupta, L., & Zameni, A. (2025). FinTech: a literature review of emerging financial technologies and applications. *Financial Innovation*, 11(1). Recuperado el 12 de enero de 2025, de <https://link.springer.com/article/10.1186/s40854-024-00668-6>

Asian Development Bank. (2019). Alternative credit scoring: The potential for financial inclusion. Recuperado el 3 de diciembre de 2024, de <https://www.adb.org/sites/default/files/linked-documents/52218-002-ld-02.pdf>

Asobancaria. (2022). Ensayos sobre inclusión financiera en Colombia II. Recuperado el 18 de octubre de 2024, de

https://asobancaria.com/wp-content/uploads/Ensayos_sobre_Inclusion_Financiera_en_Colombia-II.pdf

Asobancaria. (2022). Estudios sobre inclusión financiera en Colombia. Recuperado el 1 de mayo de 2025, de <https://www.asobancaria.com/estudios-economicos/inclusion-financiera/>

Asobancaria. (2024). Lecciones globales, acciones locales: regulación de los datos abiertos en Colombia. Recuperado el 12 de enero de 2025, de <https://www.asobancaria.com/wp-content/uploads/2024/07/1432-BE.pdf>

Bagnall, A. J., Bianchi, S., & L. J. (2019). Forecasting using ensemble methods: An application to financial markets. *Journal of Forecasting*, 38(7), 629–639. Recuperado el 29 de marzo de 2025, de <https://doi.org/10.1002/for.2557>

Baker, M., & Wurgler, J. (2020). Behavioral finance: Big data and AI. *Journal of Finance*, 75(3), 1235–1264. Recuperado el 12 de febrero de 2025, de https://www.researchgate.net/publication/350339608_Behavioral_Finance_Research_in_2020_Cui_Bono_et_Quo_Vadis

Banca de las Oportunidades. (2021). Reporte de Inclusión Financiera 2021. Recuperado el 25 de octubre de 2024, de <https://www.bancadelasoportunidades.gov.co/sites/default/files/2022-07/Reporte%20de%20inclusi%C3%B3n%20financiera%202021.pdf>

Banca de las Oportunidades. (2024). Reporte de Inclusión Financiera 2023. Recuperado el 4 de abril de 2025, de

https://www.bancadelasoportunidades.gov.co/sites/default/files/2024-06/Reporte%20de%20Inclusi%C3%B3n%20Financiera%202023_2.pdf

Banco de la República. (2023). Informe de política monetaria, agosto 2023. Recuperado el 9 de noviembre de 2024, de

<https://repositorio.banrep.gov.co/server/api/core/bitstreams/87fc8f1a-48e4-42c2-8c76-de4bda494037/content>

Banco de la República. (2023). Reporte de Estabilidad Financiera - Primer Semestre 2023.

Recuperado el 16 de enero de 2025, de <https://www.banrep.gov.co/es/reporte-estabilidad-financiera>

Banco Interamericano de Desarrollo. (2020). Inclusión financiera en América Latina y el Caribe:

Coyuntura actual y desafíos para los próximos años. Recuperado el 28 de noviembre de 2024, de

<https://publications.iadb.org/es/inclusion-financiera-en-america-latina-y-el-caribe-coyuntura-actual-y-desafios-para-los-proximos>

Banco Interamericano de Desarrollo. (2023). BID aprueba crédito de US\$200 millones para facilitar el financiamiento sostenible en Colombia. Recuperado el 20 de marzo de 2025, de

<https://www.iadb.org/es/noticias/bid-aprueba-credito-de-us200-millones-para-facilitar-el-financiamiento-sostenible-e>

Banco Mundial. (2022). Inclusión financiera: Panorama general. Recuperado el 3 de febrero de 2025, de <https://www.bancomundial.org/es/topic/financialinclusion/overview>

Banco Mundial. (2017). Alternative data transforming SME finance. Recuperado el 10 de diciembre de 2024, de <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/701331497329509915/alternative-data-transforming-sme-finance>

BBVA Research. (2024, febrero). MiPymes en Colombia: Retos y oportunidades. Recuperado el 6 de enero de 2025, de https://www.bbvarsearch.com/wp-content/uploads/2024/02/202401_MiPymes_Colombia-1.pdf

Bello, O. A. (2024). The role of data analytics in enhancing financial inclusion in emerging economies. *International Journal of Developing and Emerging Economies*, 11(3), 90–112. Recuperado el 27 de febrero de 2025, de https://www.researchgate.net/publication/383264963_The_Role_of_Data_Analytics_in_Enhancing_Financial_Inclusion_in_Emerging_Economies

Bhardwaj, B. (2020). The impact of artificial intelligence on the financial sector: Analysis of the effect of AI on the financial sector and how AI is changing the traditional banking and financial services [Tesis de maestría, National College of Ireland]. NORMA eResearch @NCI. Recuperado el 11 de diciembre de 2024, de <https://norma.ncirl.ie/4433/1/bharatbhardwaj.pdf>

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data: The story so far. *International Journal on Semantic Web and Information Systems*, 5(3), 1–22. Recuperado el 30 de octubre de 2024, de https://www.researchgate.net/publication/225070216_Linked_Data_The_Story_so_Far

Brown, I. L. J. (2012). Basel II compliant credit risk modelling [Tesis doctoral, University of Southampton]. Repositorio institucional. Recuperado el 19 de noviembre de 2024, de <https://eprints.soton.ac.uk/341517/>

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. Recuperado el 8 de febrero de 2025, de <https://ieeexplore.ieee.org/document/9069875>

CNBV. (2023). Sandbox regulatorio en México. Recuperado el 21 de septiembre de 2024, de <https://www.cnbv.gob.mx>

Colombia Fintech. (s.f.). Regulación Fintech en Colombia: Decreto 2654 de 2014 y mecanismos de control necesarios. Recuperado el 19 de enero de 2025, de <https://colombiafintech.co/lineaDeTiempo/articulo/regulacion-fintech-colombia>

Duarte, A., Frost, J., Gambacorta, L., Koo Wilkens, P., & Song Shin, H. (2022). Central banks, the monetary system and public payment infrastructures: Lessons from Brazil's Pix (52nd ed.) [Ebook]. BIS. Recuperado el 14 de octubre de 2024, de <https://www.bis.org/publ/bisbull52.pdf>

eMarketservices. (2024). Kenia. M-Pesa, el sistema de dinero móvil más extendido del mundo.

Recuperado el 26 de marzo de 2025, de

<https://www.emarketservices.es/es/menu-principal/actualizate/detalle-noticias.Kenia-M-Pesa-dinero-movil.news202409>

Finde Gateway. (2023). Scoring alternativos para la inclusión financiera. Recuperado el 21 de abril de

2025, de <https://www.findevgateway.org/es/noticias/scoring-alternativos-para-la-inclusion-financiera>

Colombia Fintech . (2023). Estadísticas del ecosistema fintech en Colombia. Recuperado el 19 de

enero de 2025, de <https://colombiafintech.co/lineaDeTiempo/articulo/fintech-snapshot-2024->

FNG. (2024). Histórico de indicadores de gestión. Recuperado el 8 de marzo de 2025, de

<https://www.fng.gov.co/nosotros/rendicion-de-cuentas/historico-de-indicadores>

GSMA. (2023). Impacto de M-Pesa en Kenya. Recuperado el 7 de marzo de 2025, de

<https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/programme/mobile-money/mobile-money-and-evolving-fintech-in-conversation-with-m-pesa/>

IEF & Altria Corpo. (2025). Quinto Barómetro Empresarial de Fintech y Financiación Alternativa.

FundsSociety. Recuperado el 18 de abril de 2025, de

<https://www.fundssociety.com/es/noticias/alternativos/solo-el-43-de-las-empresas-esperan-buscar-financiacion-alternativa-en-2025-segun-el-quinto-barometro-empresarial-de-fintech-y-financiacion-alternativa/>

Innovations for Poverty Action. (2014). Puntajes de crédito alternativos para pymes. Recuperado el 21 de abril de 2025, de <https://es.poverty-action.org/study/alternative-credit-scores-smes>

Kitchin, R., & Lauriault, T. (2018). Digital data and data infrastructures. In J. Ash, R. Kitchin, & A. Leszczynski (Eds.), Digital geographies. Sage. Recuperado el 2 de noviembre de 2024, de <https://sk.sagepub.com/book/edvol/digital-geographies/toc>

Mercado Libre. (2023). Informe de impacto de Mercado Crédito. Recuperado el 24 de abril de 2025, de <https://www.mercadolibre.com.ar>

Ministerio de Comercio, Industria y Turismo. (2024, 23 de agosto). Colombia tiene cerca de 1,56 millones de empresas formales activas, la mayor cifra de la última década. Recuperado el 2 de mayo de 2025, de <https://www.mincit.gov.co/prensa/noticias/industria/colombia-mayor-cifra-empresas-formales-activas>

MinTIC. (2023). Estadísticas de inclusión digital. Recuperado el 28 de abril de 2025, de <https://www.mintic.gov.co/portal/inicio/Servicios/1737:Estadisticas>

N26 Magazine. (2024). M-Pesa: Cómo Kenia ha revolucionado los pagos móviles. Recuperado el 17 de abril de 2025, de <https://mag-es.n26.com/m-pesa-c%C3%B3mo-kenia-ha-revolucionado-los-pagos-m%C3%B3viles-d5eff43cb9ec>

OECD. (2021). Financial inclusion. Recuperado el 17 de marzo de 2025, de

<https://www.oecd.org/finance/financial-inclusion/>

OECD. (2025). Financing SMEs and Scoreboard: 2025 Highlights. SME Finance Forum. Recuperado el 21 de abril de 2025, de

<https://smefinanceforum.org/post/launch-of-the-oecd-financing-smes-and-scoreboard-2025-highlights>

OIT. (2023). Panorama Laboral 2023: América Latina y el Caribe. Recuperado el 9 de abril de 2025, de

https://www.ilo.org/sites/default/files/wcmsp5/groups/public/%40americas/%40ro-lima/documents/publication/wcms_906617.pdf

ONU Mujeres. (2023, mayo). Mujeres y PYMES: Brechas en el acceso al crédito. Recuperado el 21 de febrero de 2025, de

<https://www.unwomen.org/sites/default/files/2023-11/generation-equality-accountability-report-2023-es.pdf>

Reserve Bank of India. (2023). Annual Report Publications. Recuperado el 7 de abril de 2025, de

<https://www.rbi.org.in/Scripts/AnnualReportPublications.aspx>

Stadler, C. (2024, 11 de junio). M-Pesa: Why The World's First Large Mobile Payment Platform Keeps On Winning. Forbes. Recuperado el 7 de abril de 2025, de

<https://www.forbes.com/sites/christianstadler/2024/06/11/m-pesa-why-the-worlds-first-large-mobile-payment-platform-keeps-on-winning/>

Superintendencia Financiera de Colombia. (2023). Reporte de inclusión financiera 2023. Recuperado el 28 de octubre de 2025, de <https://www.superfinanciera.gov.co>

TiendaPago. (2023). Financiamiento para comercios informales. Recuperado el 21 de abril de 2025, de <https://www.tiendapago.com>

Anexo:

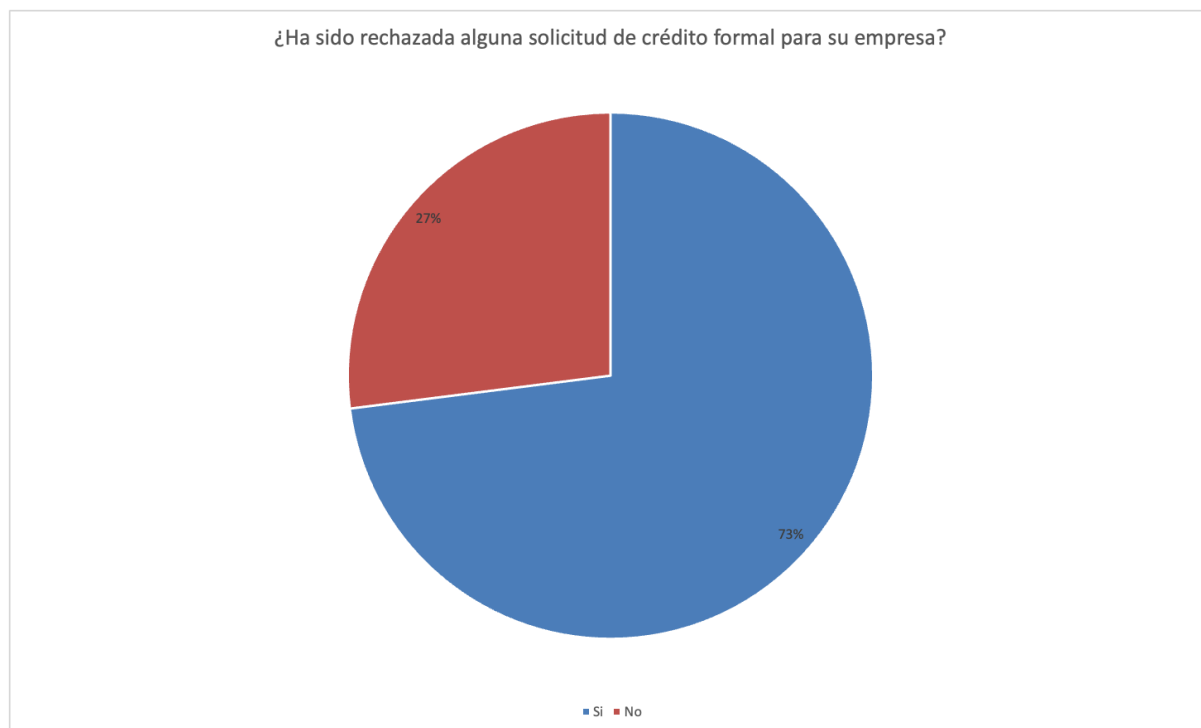
Investigación (Encuesta) Anexo 1

- 1. ¿Dónde se ubica tu empresa?
 - Abierta
- 2. ¿Cuántos empleados tiene tu empresa?
 - Menos de 10
 - Entre 10 y 50
 - Entre 51 y 100
 - Entre 101 y 150
 - Entre 151 y 200
- 3. ¿Cuál es su actividad económica (CIIU)?
- Abierta
- 4. ¿Cuántos son sus ingresos al mes (Millones)?
 - Menos de 20
 - Entre 20 y 30
 - Entre 30 y 40
 - Más de 40
- 5. ¿Ha enfrentado alguna de estas barreras al intentar acceder a créditos formales?
Seleccionar varias opciones.
 - Altos costos financieros (tasas de interés elevadas)
 - Exceso de trámites burocráticos
 - Falta de garantías o colaterales
 - Desconocimiento de productos financieros
 - Historial crediticio insuficiente
 - Otras (especifique)
- 6. ¿Ha sido rechazada alguna solicitud de crédito formal para su empresa?
 - Si
 - No
- 7. ¿Cuál cree que fue la razón principal?
 - Falta de historial crediticio
 - Falta de garantías o colaterales
 - Ingresos insuficientes
 - Otras (especifique)
- 8. ¿Qué alternativas de financiamiento distintas al sistema financiero formal ha utilizado?
(Seleccione todas las que apliquen)
 - Proveedores (créditos comerciales)
 - Recursos propios (ahorros o reinversión de utilidades)
 - Financiamiento a través de familiares o amigos
 - Plataformas de crowdfunding
 - Factoring o leasing
 - Otras (especifique)
- 9. La falta de historial crediticio adecuado dificulta el acceso a financiación formal de su empresa
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 10. ¿Los requisitos excesivos de documentación y garantías dificultan el acceso a financiación formal de su empresa?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral

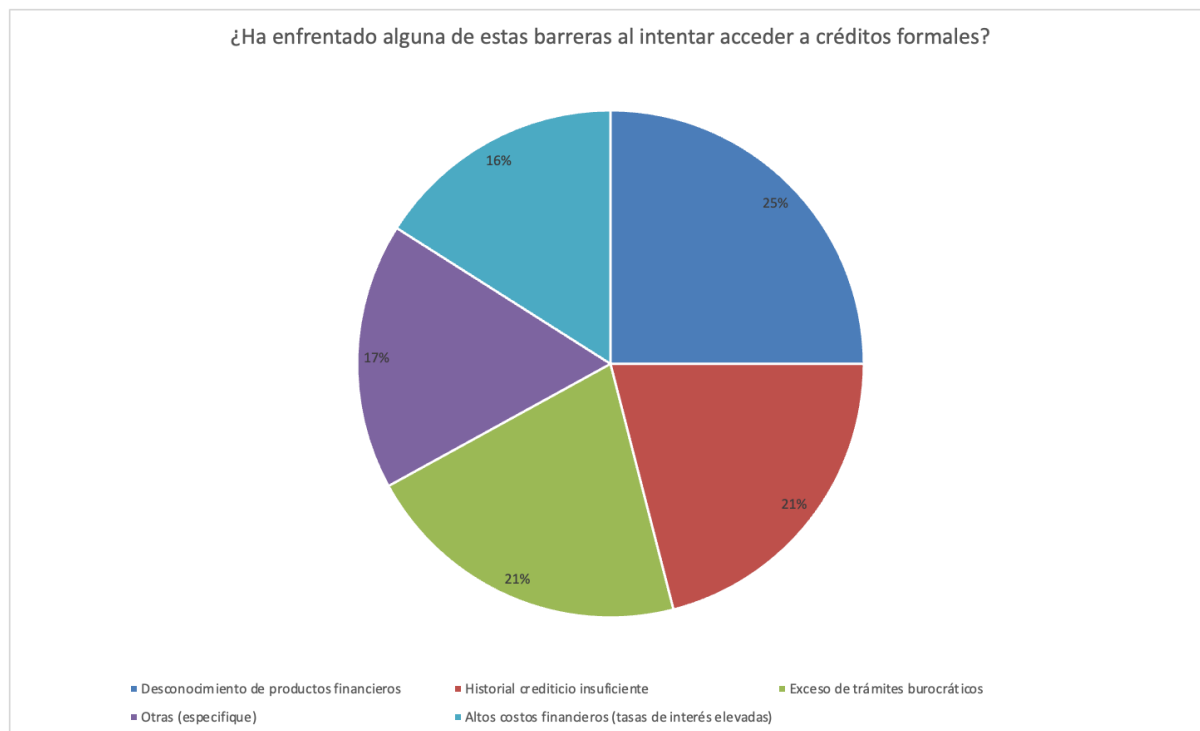
- De acuerdo
- Totalmente de acuerdo
- 11. ¿Los altos costos asociados con el acceso al crédito dificultan el acceso a financiación formal de su empresa?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 12. La falta de conocimiento sobre las alternativas de crédito disponibles dificultan el acceso a financiación formal de su empresa
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 13. Los altos tiempos de aprobación dificultan el acceso a financiación formal de su empresa
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 14. ¿Los montos y plazos de crédito aprobados por las entidades financieras son suficientes para cubrir las necesidades de su empresa?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 15. ¿Conoce y ha utilizado productos financieros digitales (como banca electrónica, factoring electrónico, o plataformas de crowdfunding) para acceder a créditos?
 - Sí, los conozco y los he utilizado
 - Sí, los conozco pero no los he utilizado
 - No los conozco
- 16. ¿Estaría dispuesto a compartir datos alternativos (como flujos de caja, ventas digitales, o historial de pagos a proveedores, entre otros) para mejorar su perfil crediticio y acceder a créditos?
 - Sí, totalmente dispuesto
 - Depende del tipo de datos solicitados
 - No, no estoy dispuesto
- 17. ¿Ha solicitado créditos de redescuento (como los ofrecidos por Bancoldex, Finagro, entre otros) para su empresa?
 - Sí, los he utilizado
 - No, pero los conozco
 - No los conozco
- 18. ¿Los modelos tradicionales de scoring crediticio (basados en historial crediticio y garantías) son adecuados para evaluar el riesgo de su empresa?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo

- 19. ¿Estaría interesado en que su empresa fuera evaluada mediante modelos de scoring alternativo que utilicen datos no tradicionales (como ventas digitales, pagos a proveedores, o flujos de caja)?
 - Sí, estaría muy interesado
 - Depende de cómo se utilicen los datos
 - No, no estoy interesado
- 20. ¿Qué datos no financieros está dispuesto a compartir (como transacciones en plataformas digitales, redes sociales, o patrones de consumo) para mejorar su perfil crediticio y acceder a créditos?
- Abierta
- 21. ¿Confiaría en un sistema de evaluación crediticia basado en inteligencia artificial (IA) que utilice datos alternativos para otorgar créditos?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo
- 22. ¿Es importante que los modelos de scoring crediticio sean transparentes y expliquen cómo se toman las decisiones crediticias?
 - Totalmente en desacuerdo
 - En desacuerdo
 - Neutral
 - De acuerdo
 - Totalmente de acuerdo

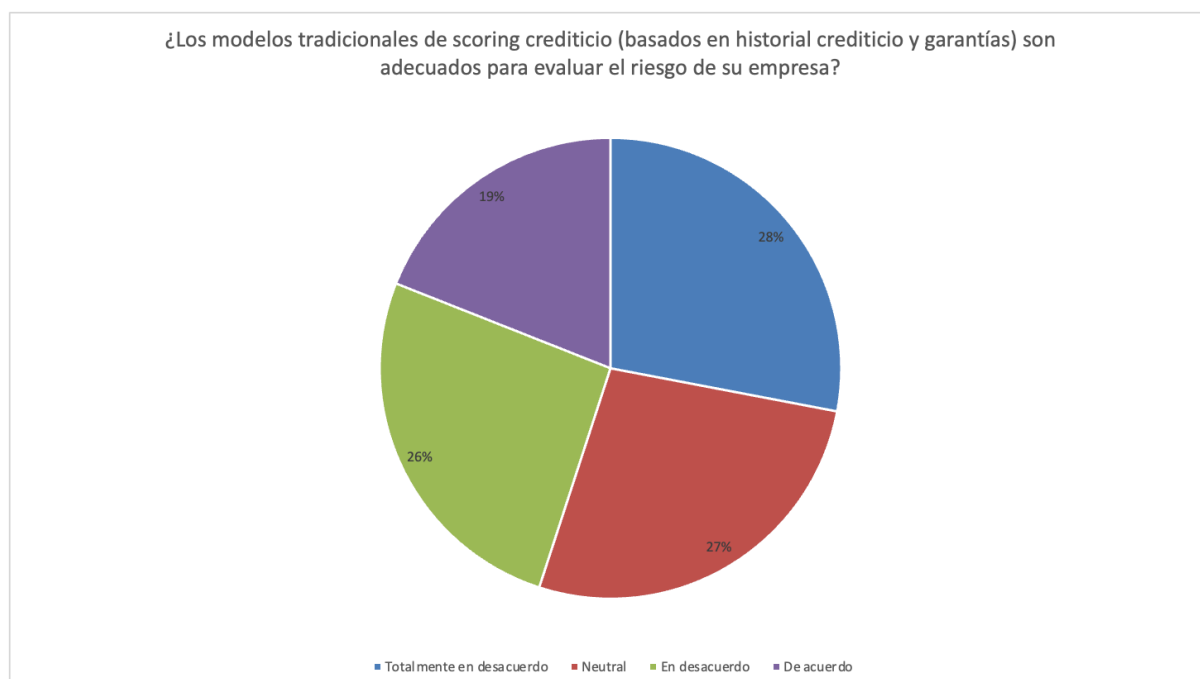
Anexo 2

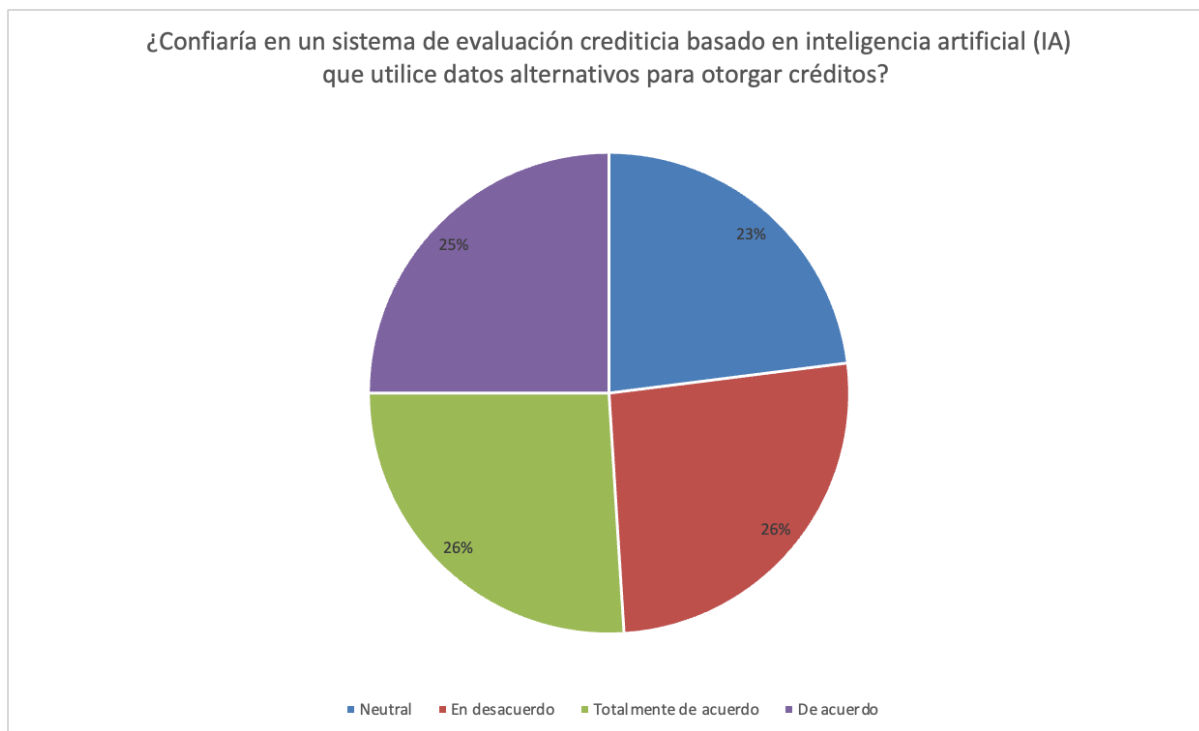


Anexo 3



Anexo 4



Anexo 5**Anexo 6**