

Aplicación de Machine Learning en la Gestión de Riesgo de Crédito Financiero: Una revisión sistemática

Ericka Liseth Mendoza Saltos

17 de octubre de 2023

Resumen

Este artículo examina la gestión de riesgos bancarios, centrándose en riesgos como el crediticio, de mercado, operativo y de liquidez, con especial énfasis en el riesgo crediticio. El estudio se enfoca en la implementación de algoritmos de machine learning para gestionar el riesgo crediticio y evalúa su rendimiento a través de una revisión sistemática de la literatura. Se seleccionaron 12 artículos para el análisis, revelando que el riesgo crediticio es de suma importancia en el sector financiero. Aunque algunos algoritmos de machine learning se han implementado, enfrentan desafíos, especialmente en cuanto a su falta de explicabilidad y transparencia, requisitos esenciales para los organismos regulatorios. Como solución, se sugieren modelos híbridos y el uso de métodos como SHAP para mejorar la interpretación de estos modelos complejos.

1. Introducción

La importancia del machine learning (ML) en diversos campos, especialmente en la gestión de riesgos financieros. Las empresas están invirtiendo significativamente en algoritmos de ML para predecir datos futuros, especialmente en el sector financiero, donde se enfocan en la gestión de riesgos crediticios. A medida que las organizaciones reconocen la utilidad de estos algoritmos, se están moviendo de modelos estadísticos tradicionales a soluciones más sofisticadas de ML para cuantificar y mitigar el riesgo de manera más efectiva. La investigación se centra en revisar la literatura existente para identificar los algoritmos de ML utilizados en la gestión de riesgos crediticios, evaluando sus ventajas y desventajas. El documento está estructurado en secciones que explican los conceptos, la metodología de investigación, los resultados y las conclusiones.

2. Conceptos y Técnicas de Machine Learning

2.1 Inteligencia Artificial y Machine Learning

Machine learning (ML) es una parte de la inteligencia artificial (IA) que permite a las máquinas imitar la inteligencia humana y mejorar con la recopilación de datos, sin requerir programación explícita. Se categoriza en aprendizaje supervisado, no supervisado, semisupervisado y por reforzamiento según cómo aprenden los algoritmos. La explosión en la investigación de ML desde los años noventa se atribuye a la colaboración entre diferentes comunidades de investigación y a la aplicación de técnicas de aprendizaje automático en una variedad de problemas, desde el reconocimiento de voz hasta el análisis de datos y el control de robots.

2.2 Aprendizaje supervisado

Este conjunto de algoritmos utiliza variables etiquetadas para proceder con un entrenamiento. A través de esto, el modelo puede ser capaz de realizar predicciones de ejemplos sin etiquetar. Principalmente, se asocia con problemas de clasificación, regresión y “ranking problem”. Algunos tipos de algoritmos que utilizan el aprendizaje supervisado son regresión lineal, regresión logística, k-nearest neighbor (KNN), support vector machine (SVM), random forest, árboles de decisión y Naive Bayes.

2.3 Aprendizaje no Supervisado

Este conjunto de algoritmos tiene como principal tarea detectar patrones a partir de datos no etiquetados. En el entrenamiento no existen datos etiquetados disponibles y el principal enfoque que se les otorga es resolver problemas de “clustering”, detección de valores atípicos, reducción de dimensionalidad y detección de anomalías.

2.4 Aprendizaje Profundo (Deep Learning)

Este tipo de aprendizaje se encuentra dentro del machine learning que tiene un campo de estudio muy amplio, debido a la complejidad de sus modelos y que muchas veces es representado como “caja negra” que extrae representaciones de diferentes capas de características.

3. Metodologia

Para la realización del presente trabajo, se adoptó una revisión sistemática de la literatura (SRL, por sus siglas en inglés). Se optó por realizar una metodología similar a la propuesta por Nina et al. (2021), la cual consta de una planificación para identificar el interés, la población y el contexto actual, el planteamiento de preguntas relevantes para la revisión, una búsqueda de artículos por palabras clave y una construcción de cadenas de búsqueda, y, finalmente, una filtración y selección de artículos. Los pasos detallados son los siguientes:

Planificación de estudio

En primer lugar, se requirió realizar una planificación del estudio utilizando la técnica PICO:

- Población: Gestión de Riesgos de Crédito
- Interes: Aplicación de Machine Learning en el Riesgo de Crédito
- Contexto: Sector bancario, Fintech o financieras, plataformas de préstamos

Como interés de la presente revisión, se lograron identificar 4 preguntas:

- ¿Qué relevancia presenta el riesgo de crédito dentro de la gestión de riesgos en el sector financiero?
- ¿Cómo el Machine Learning mejora la predicción del riesgo de crédito financiero frente a métodos tradicionales?
- ¿Qué retos presenta el Machine Learning frente a los métodos tradicionales?
- ¿Cómo identificar el algoritmo de Machine Learning más eficaz para la Gestión de Riesgo de Crédito Financiero?

Búsqueda de artículo

En segundo lugar, se utilizó la base de datos de bibliografía indexada Scopus, la cual es frecuente para estudios de revisión sistemática de la literatura. Dado que cuenta con una herramienta que permite realizar búsquedas avanzadas, se optó por identificar palabras claves que permitiesen realizar una cadena de búsqueda y así realizar un filtrado de artículos.

2.3 Selección de artículos

En tercer lugar, se realizaron criterios de inclusión y exclusión

4. Resultados

4.1 ¿Qué relevancia presenta el riesgo de crédito dentro de la gestión de riesgos en el sector financiero?

El riesgo crediticio, fundamental para los bancos, requiere una gestión efectiva para evitar pérdidas. La gestión de riesgos financieros (FRM) se divide en varias categorías, siendo el riesgo crediticio el más crítico y enfocado en la mayoría de los estudios. Evaluar la capacidad del prestatario y calcular pérdidas potenciales es crucial; la calificación crediticia desempeña un papel central. De acuerdo con los Acuerdos de Basilea III, las técnicas para gestionar el riesgo crediticio pueden ser estandarizadas o basadas en calificaciones internas. La falta de una buena práctica en la gestión de riesgos crediticios puede poner en peligro a cualquier entidad, ya que representa una oportunidad clave para mejorar el desempeño y mantener una ventaja competitiva.

4.2 ¿Cómo el Machine Learning mejora la predicción del riesgo de crédito frente a métodos tradicionales? Los algoritmos clasificadores de

machine learning, como las support vector machines y las redes neuronales, han demostrado ser significativamente más precisos en la calificación crediticia que los métodos estadísticos tradicionales como el análisis discriminante lineal o la regresión logística. Estos métodos de aprendizaje automático, tanto supervisados como no supervisados, han reemplazado las evaluaciones subjetivas y han mejorado la precisión en la predicción del riesgo crediticio. La combinación de algoritmos de machine learning con métodos tradicionales, como la regresión logística y los árboles de decisión, ha mostrado resultados prometedores, destacando la importancia de integrar enfoques diversos para una gestión efectiva de riesgos crediticios.

4.3 ¿Qué retos presenta el Machine Learning frente a los métodos tradicionales?

El texto aborda los desafíos en la adopción de modelos de machine learning para la gestión de riesgos, especialmente en el ámbito crediticio. Los algoritmos clasificadores de machine learning, aunque altamente precisos, a menudo son considerados “cajas negras” debido a su complejidad y dificultad para explicar

cómo obtienen resultados. Las entidades regulatorias exigen explicabilidad, interpretabilidad y transparencia en los modelos, requisitos difíciles de cumplir para ciertos algoritmos, lo que ha generado resistencia en su implementación. Además, se destaca el dilema entre la precisión y la interpretación: los modelos complejos ofrecen alta precisión predictiva pero limitada interpretación, mientras que los modelos estadísticos "simples" son más interpretables pero menos precisos. Para abordar este problema, se están utilizando herramientas como SHAPley Additive exPlanations para mejorar la explicabilidad de los modelos. Además, se mencionan desafíos relacionados con sesgos en la selección de datos y la necesidad de personal capacitado, ética en la recopilación de datos y protección de la privacidad.

[1]

4.4¿Cómo identificar el algoritmo de Machine Learning más eficaz para la Gestión de Riesgo de Crédito Financiero?

Son diversos algoritmos de machine learning e identificar uno de estos como el más adecuado es una tarea complicada; sin embargo, se realizan diversas acciones y métricas para poder determinar si el modelo es adecuado. Para esta función, se toma en cuenta la matriz de confusión y, a partir de esta, poder determinar la sensibilidad y especificidad del modelo. Asimismo, se analiza la curva ROC, o AUC en inglés, para calcular el área bajo su curva

5. Conclusiones

Los algoritmos de machine learning se llevan utilizando un periodo de tiempo considerable, por lo cual es notoria la evolución que han tenido, asimismo ha habido un aumento importante en la gestión de riesgos de cualquier entidad que se dedique al rubro financiero, pero ha existido un mayor énfasis en la gestión de riesgo de crédito, debido a que representa el mayor riesgo, ya que evita pérdidas y ganancias en las organizaciones.

Asimismo, se ha podido determinar que el machine learning representa ventajas frente a los métodos tradicionales, ya que obtienen mejor éxito para clasificar a los potenciales clientes. [1]

Referencias

- [1] Jordi Navarro y Andrés González Pere Vandellós. Machinelearning. 2017.