

Auditoría de modelos estadísticos en entornos financieros: caso práctico en créditos para automóviles

Natalia Janeiro Fong

Máster en Técnicas Estadísticas

Curso 2024-2025

En el presente documento se resume el Trabajo de Fin de Máster titulado «Auditoría de modelos estadísticos en entornos financieros: caso práctico en créditos para automóviles». No se autoriza la publicación de dicho trabajo en la web del Máster por motivos de confidencialidad derivados de la colaboración con ABANCA Corporación Bancaria S. A., entidad en cuyo departamento de Auditoría de Modelos se desarrollaron las prácticas curriculares vinculadas a la Modalidad B de este máster. La memoria emplea datos reales de operaciones de financiación de vehículos de clientes de la entidad y describe en detalle el modelo interno de scoring crediticio en producción, cuya divulgación pública comprometería la confidencialidad de los procesos de gestión del riesgo, la protección de los datos personales de los titulares de las operaciones y los intereses legítimos de ABANCA como entidad colaboradora.

Introducción y motivación del trabajo

El riesgo de crédito, definido como la probabilidad de que un prestatario incumpla sus obligaciones contractuales de pago, constituye uno de los pilares fundamentales de la gestión bancaria. Su correcta cuantificación determina directamente el capital regulatorio que las entidades deben mantener conforme a los Acuerdos de Basilea II y Basilea III, articulados en torno a tres parámetros clave: la Probabilidad de Default (PD), la Pérdida Dado el Incumplimiento (LGD) y la Exposición al Incumplimiento (EAD). En este marco, la Autoridad Bancaria Europea (EBA) y el Banco Central Europeo (BCE) exigen que los modelos utilizados bajo el enfoque IRB (Internal Ratings-Based) sean robustos, interpretables, estables y sometidos a auditoría independiente periódica.

Los modelos de scoring crediticio son la herramienta central de este proceso. Asignan a cada operación una puntuación que sintetiza el nivel de riesgo del solicitante, permitiendo discriminar entre prestatarios solventes y aquellos con mayor probabilidad de incumplimiento. Su validación y auditoría constituyen, por tanto, una función crítica dentro de las áreas de control interno de las entidades financieras. El presente trabajo se enmarca precisamente en esa función, al desarrollarse en el departamento de Auditoría de Modelos de Riesgo Crediticio de ABANCA Corporación Bancaria S. A., en el contexto de las prácticas curriculares de la Modalidad B del Máster en Técnicas Estadísticas.

El objeto concreto de estudio es un modelo de scoring crediticio en producción para la financiación de automóviles y motocicletas, cuyo objetivo es estimar la probabilidad de impago a doce meses en estas operaciones. La arquitectura del modelo en producción está basada en una regresión logística con variables transformadas mediante la técnica Weight of Evidence (WoE), estándar en el sector bancario regulado por su interpretabilidad y trazabilidad ante el supervisor. El trabajo parte de la replicación de dicho modelo, evalúa

alternativas metodológicas y propone una especificación mejorada dentro de la misma arquitectura mediante una selección de variables distinta.

Objetivos del estudio realizado

El objetivo principal del trabajo es auditar y replicar el modelo en producción, y determinar si es posible mejorar su capacidad predictiva, ya sea mediante una selección de variables diferente dentro de la misma arquitectura GLM+WoE, ya sea adoptando una familia metodológica alternativa, manteniendo en todo caso los requisitos de interpretabilidad y supervisabilidad del marco IRB. Para ello se evalúan diez familias de modelos estadísticos distintas, todas bajo las mismas condiciones de test, con el fin de garantizar la comparabilidad de los resultados.

Un objetivo secundario es la identificación de un núcleo de predictores robusto y consistente entre metodologías distintas. La convergencia entre familias metodológicas muy diferentes en la selección de variables aporta evidencia empírica sobre qué factores son verdaderamente relevantes para estimar el riesgo de impago en el segmento de automoción, con independencia del enfoque de modelización.

Metodología de la solución propuesta

La muestra de trabajo comprende operaciones de financiación de vehículos formalizadas entre 2022 y 2023, con información del perfil socioeconómico del solicitante, características de la operación y variables externas de comportamiento crediticio. La variable objetivo es la ocurrencia de mora a doce meses. El principal reto metodológico es el fuerte desequilibrio entre clases: la tasa de impago es marcadamente minoritaria, lo que invalida el uso de métricas de rendimiento convencionales como el accuracy. La evaluación se realiza en su totalidad mediante el Área Bajo la Curva ROC (AUC), el coeficiente Kappa de Cohen y la Balanced Accuracy, métricas robustas ante el desequilibrio de clases. El desbalance se trata explícitamente en todos los métodos evaluados.

El proceso de preparación de los datos incluye depuración, recategorización de variables y transformación WoE, que convierte cada variable en un valor numérico continuo proporcional a la separación entre la distribución de buenos y malos pagadores en cada categoría. Esta transformación introduce linealidad entre los predictores y el log-odds de impago, facilita la comparación entre variables de naturaleza distinta y es directamente interpretable, lo que la hace especialmente adecuada para su uso en modelos IRB.

Se evalúan las siguientes familias metodológicas: regresión logística con WoE, regularización mediante Ridge, LASSO y Elastic Net, reducción de dimensión mediante PCR y PLS-DA, modelos aditivos generalizados (GAM), árboles de decisión CART, métodos ensemble (Balanced Random Forest y Real AdaBoost), máquinas de soporte vectorial con kernel RBF y redes neuronales con una capa oculta. Dentro de la familia GLM+WoE, la selección de variables se realiza mediante tres estrategias en paralelo (stepwise, LASSO y Elastic Net) que convergen en el mismo subconjunto reducido de predictores relacionados con el perfil del solicitante y las características de la operación, distinto del subconjunto empleado por el modelo en producción.

Resultados

El modelo en producción de la entidad obtiene un AUC de 0,805 sobre el conjunto de test. El GLM con codificación WoE resultante de la nueva selección de variables alcanza un AUC de 0,8975, Kappa de Cohen de 0,3569 y Balanced Accuracy de 0,7185, lo que representa una mejora sustancial sobre el modelo de referencia manteniendo exactamente la misma arquitectura. Los modelos regularizados LASSO y Elastic Net alcanzan métricas prácticamente equivalentes y seleccionan el mismo subconjunto de variables, confirmando la estabilidad de la especificación propuesta.

Entre los métodos no lineales, el Balanced Random Forest y el Real AdaBoost son los más competitivos (AUC = 0,848), pero no superan al GLM con la nueva selección de variables. El GAM completo alcanza AUC = 0,779 con sensibilidad del 62,5%; el árbol CART obtiene la mayor sensibilidad de la comparación (66,7%) con AUC = 0,770; la SVM y la red neuronal registran AUC = 0,738 y AUC = 0,693, respectivamente. Los métodos de reducción de dimensión resultan los menos competitivos, con AUC inferiores a 0,61. Ninguna familia metodológica alternativa supera la especificación GLM+WoE con variables reseleccionadas.

Un hallazgo adicional de valor práctico es la identificación de un núcleo predictivo estable: un subconjunto de variables reconocidas convergentemente por tres familias metodológicas muy distintas (modelos lineales, regularización y métodos basados en árboles). Esta convergencia aporta evidencia empírica sobre cuáles son los factores fundamentales del riesgo de impago en el segmento estudiado, con independencia del enfoque de modelización.

Conclusiones

El principal hallazgo del trabajo es que una reselección de variables dentro de la arquitectura GLM+WoE, manteniendo la misma metodología que el modelo en producción, permite mejorar sustancialmente la capacidad discriminante: el AUC pasa de 0,805 a 0,8975. Esta mejora se obtiene sin incrementar la complejidad del modelo ni comprometer su interpretabilidad o supervisabilidad, lo que la hace directamente aplicable en un entorno regulado bajo el marco IRB.

Ninguna de las familias metodológicas alternativas evaluadas supera al GLM con la nueva selección de variables. La transformación WoE aporta una estructura de información que los métodos de mayor flexibilidad no logran superar con el volumen de datos disponible, y el incremento de complejidad metodológica no se traduce en mejora predictiva. La auditoría concluye que la arquitectura presente del modelo es la más adecuada para el entorno regulado en el que opera la entidad, y que la especificación propuesta de interpretabilidad, estabilidad y supervisabilidad, constituye una mejora fundamentada empíricamente sobre el modelo actualmente en producción.