

valoraban internamente si concedían o no el crédito al solicitante. El primer sistema se creó en 1958 y en 1970 se aplicó a tarjetas de crédito bancarias. Aunque la fórmula exacta para su cálculo es desconocida, los pesos aproximados de los distintos componentes del FICO son los siguientes, según Moral Rincón (2014):

- A. Historial de pagos anteriores: 35 por 100. Pagos retrasados de facturas o créditos (hipotecas, tarjetas de crédito, préstamos para la adquisición de automóviles, etc.) causan una reducción del FICO. El principal indicador de solvencia es si la persona ha cumplido siempre a tiempo con sus obligaciones financieras.
- B. Utilización de crédito: 30 por 100. La ratio de la deuda actual sobre el límite total de crédito disponible. Existe menos probabilidad de impago si la persona ha consumido un menor porcentaje de su línea de crédito disponible.
- C. Extensión del historial crediticio: 15 por 100. A medida que el historial crediticio aumenta, si no se producen impagos, el FICO aumenta. Cuanto más tiempo el individuo haya utilizado productos de financiamiento se le otorgará un mayor puntaje.
- D. Tipos de crédito utilizados: 10 por 100. Un historial crediticio con múltiples tipos de créditos (hipotecario, al consumo, etcétera) supone un FICO elevado. Si la persona ha gestionado responsablemente varias clases de préstamos, se considerará que la probabilidad de incumplimiento es menor.
- E. Búsqueda reciente de crédito o cantidad de crédito obtenido recientemente: 10 por 100. Múltiples solicitudes de nuevo crédito pueden perjudicar el FICO. Acumular muchas solicitudes de deuda puede bajar la puntuación del *credit scoring*.

Los componentes del FICO muestran por qué resulta complicado conseguir un crédito o una tarjeta, al menos en Estados Unidos, con un historial crediticio breve o sin historial. Asimismo en países con poco desarrollo financiero, donde

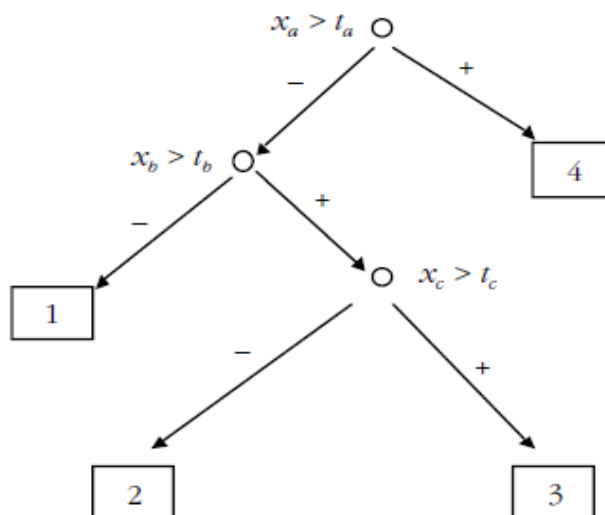
interno de aprendizaje divide los datos en muestra de training y testing. La comprensión de lo que ocurre dentro del modelo requiere de conocimientos muy especializados.

5.3 Árboles de regresión y clasificación (CART)

Friedman (1977) es de los primeros autores que emplea la técnica de los árboles de decisión al *credit scoring*. Posteriormente, Makowski (1985) y Carter y Carlett (1987) aplicaron modelos de árboles de decisión para clasificar a los clientes que solicitaban un crédito. Los árboles de decisión son una técnica no paramétrica de clasificación binaria que permite clasificar a los solicitantes de un crédito en categorías previamente establecidas. La utilización de árboles de decisión de *credit scoring* utiliza la selección de variables para la aceptación o negación de una solicitud de crédito, que se genera mediante la ejecución de un proceso interno que es automático. La técnica de árboles de decisiones permite la división óptima de la muestra, de tal forma que la variable respuesta indique diferentes perfiles de riesgo. Entre las ventajas del *credit scoring* basado en árboles de decisión se encuentran que representa una relación visual entre las variables, su flexibilidad, su firme estructura a transformaciones de las variables independientes, su interpretabilidad, al ser de fácil comprensión de los resultados obtenidos, y además, no se requiere establecer supuestos previos a la construcción del modelo acerca de la distribución de la población objeto de estudio. Entre los inconvenientes que tiene este modelo están el sobreaprendizaje y que no presenta explícitamente las probabilidades que un cliente no cumpla con el pago del crédito. El modelo CART (Classification and Regression Trees), utilizado por Friedman (1977), es una técnica cuya función trata en dividir sucesivamente la muestra original en submuestras procurando que éstas sean cada vez más homogéneas, utilizando para ello reglas univariantes que permitan determinar cuál es la variable independiente que mejor discrimina la muestra. Su utilización en problemas de clasificación es relativamente reciente si lo comparamos con otros modelos; cabe citar los trabajos de Breiman (1984), Frydman (1985) y Marais (1984).

CART presenta una estructura en forma de árbol compuesto por una sucesión de nodos y ramas que constituyen respectivamente los grupos y divisiones que se van realizando de la muestra original. Cuando ya no sea posible realizar ninguna otra división que mejore la homogeneidad de los subgrupos se llega a un nodo terminal, que será la clase asignada por el modelo. El error total del árbol se calculará sumando los correspondientes a cada uno de estos nodos terminales. El principal problema con el que se enfrenta este modelo es la complejidad de su estructura, que es común para todos los modelos no paramétricos, y puede llevar a un desembocamiento en el sobreaprendizaje de las observaciones que componen la muestra. Por este motivo, en la construcción de los modelos CART no solamente se persigue crear conjuntos homogéneos, sino que también se pretende obtener aquella estructura que presente una complejidad óptima. El modelo, como vemos en la Figura 1, se estructura en forma de árbol compuesto de una sucesión de nodos y ramas que constituyen, respectivamente, los grupos y divisiones que se van realizando de la muestra original. Cada uno de los nodos terminales representa aquel grupo cuyo coste esperado de error sea menor, es decir, aquellos que presenten menor riesgo. El proceso finaliza cuando sea imposible realizar una nueva división que mejore la homogeneidad existente. El riesgo total del árbol se calcula sumando los correspondientes a cada uno de los nodos terminales.

Figura 1: Árboles de Clasificación



6. Conclusiones

El motivo por el cual he realizado este trabajo es simplemente con fines educativos e informativos. Para ello he tenido que aportar una gran cantidad de referencias bibliográficas que considero que serán de gran ayuda para el lector. Con el objetivo de que el lector pueda obtener una clara visión sobre el *credit scoring*, partiendo desde una base de conocimiento nula sobre este tema. He recopilado toda la información que he considerado relevante para la comprensión de dicho tema, que ayudará a comprender la importancia y el impacto que tiene el *credit scoring* en la economía y en la sociedad actual. Soy consciente de que no he nombrado todas las técnicas paramétricas y no paramétricas, tan solo he citado aquellas que considero que son las más comunes, menos complejas y más sencillas de comprender. A pesar de eso considero que este trabajo aporta todo lo necesario para entender cómo funciona y para qué sirve el *credit scoring*. Por lo tanto, se puede concluir que el trabajo cumple con los objetivos planteados:

1. Se ha definido qué es el *credit scoring*. Se ha conocido su origen gracias al trabajo de Durand (1941) por aplicar las técnicas estadísticas de Fischer (1936) a la economía. Durante el resto del siglo XX se ha visto cómo diferentes autores y científicos han ido desarrollando diferentes técnicas que han permitido mejorar el sector financiero. En España no fue hasta 1983 que se implementaron dichas técnicas, que permitieron sustituir a las tradicionales basadas en el análisis del patrimonio, siendo éstas subjetivas, lentas y costosas.
2. Se ha comentado las diferentes funciones, aplicaciones y técnicas del *credit scoring* y el por qué ha adquirido tanta importancia para el sector bancario. Así como sus ventajas y desventajas.
3. Se ha contado la historia del Comité de Supervisión Bancaria de Basilea. En el cual se detalla el origen y la importancia de sus acuerdos, así como los pilares de Basilea II, al tener un papel muy importante a la hora de supervisar y regular los bancos.

4. Se ha visto lo que es el método FICO, cómo funciona y para qué sirve. También se ha hablado de algunas técnicas del *credit scoring*, clasificándolas en modelos paramétricos y no paramétricos.
5. Tras haber nombrado las distintas técnicas estadísticas y matemáticas que se emplean en el *credit scoring*, se ha resumido sus ventajas y desventajas en las siguientes tablas 6.1 y 6.2 respectivamente.

Tabla 6.1: Ventajas y desventajas (modelos paramétricos)

Modelos paramétricos	Ventajas	Desventajas
Análisis discriminante	Buen rendimiento en muestras grandes Probabilidades a posteriori son fáciles de obtener	Rigidez para el cumplimiento de las hipótesis Difícil calcular la probabilidad de impago de un cliente
Probabilidad lineal	Resultados fáciles de interpretar Alto poder predictivo	Problemas importantes al estimar la variable <i>dummy</i>
Logit	Flexible y fácil de calcular No requiere establecer una hipótesis inicial respecto al comportamiento de las variables	Dificultad de interpretación de los parámetros
Probit	Muestran la probabilidad de impago de un cliente No requiere establecer una hipótesis inicial respecto al comportamiento de las variables	Dificultad en la interpretación de los parámetros Proceso de estimación tiende a ser complicado

Tabla 6.2: Ventajas y desventajas (modelos no paramétricos)

Modelos no paramétricos	Ventajas	Desventajas
Programación lineal	Flexible Admite gran cantidad de variables No necesita establecer supuestos previos a su formulación	Difícil comprensión Predicciones no exactas
Redes neuronales	Flexible Alto nivel de predicción cuando las muestras son pequeñas Aprendizaje no supervisado	Difícil comprensión No presenta explícitamente las probabilidades de que un cliente cumpla o no con el pago Requiere conocimientos muy especializados
Árboles de decisión	Flexible Representa una relación visual entre las variables Firme estructura a transformaciones de las variables independientes Fácil de comprender No requiere establecer supuestos previos a la construcción del modelo	Sobreaprendizaje No presenta explícitamente las probabilidades de que un cliente cumpla o no con el pago

7. Bibliografía

- Altman, E. (1968). Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4), 589-610.
- Altman, E. (2002). Revisiting Credit Scoring Models in a Basel II Environment,” in Credit Rating: Methodologies, Rationale, and Default Risk. *London Risk Books*.
- Artís Ortuño, M., & Guillén Estany, M. (2013, septiembre). *Variables dependientes cualitativas*.
- Banco de España. (n.d.). *El Comité de Supervisión Bancaria de Basilea (BCBS)*.
- Banda Ortiz, H., & Garza Morales, R. (2014). Aplicación teórica del método Holt-Winters al problema de credit scoring de las instituciones de microfinanzas. *Revista de Investigación y Análisis*, 15(2), 5-21.
- Bank for International Settlements. (n.d.). *History of the Basel Committee*.
- Bank for International Settlements. (2001, enero). *New Basel Capital Accord*. Bank for International Settlements.
- Bonfim, D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking and Finance*, 33(2), 281-299.
- Bonilla, M., Olmeda, I., & Puertas, R. (1999). *Clasificación Crediticia Mediante Modelos de Agregación*.
- Bonilla, M., Olmeda, I., & Puertas, R. (2003). Modelos paramétricos y no paramétricos en problemas de credit scoring. *Revista española de financiación y contabilidad*, 32(118), 833-869.

- Boyes, W. J., Hoffman, D. L., & Low, S. A. (1989, enero). A econometric analysis of the bank credit scoring problem. *Journal of Econometrics*, 40(1), 3-14.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. Taylor & Francis.
- Campbell, T., & Dietrich, J. (1983). The determinants of default on insured conventional residential mortgage loans. *The Journal of Finance*, 38(5), 1569-1581.
- Carter, C., & Carlett, J. (1987). Assessing credit card applications using machine learning. *IEEE Expert. Fall*, 71-79.
- Cheung, S. (1996). Provincial credit ratings in Canada: An ordered probit analysis. In *Working Paper Bank of Canada*.
- Dabós, M., & Universidad de Belgrano. (n.d.). *Credit Scoring*.
- Davis, R. H., Edelman, D. B., & Gamberman, A. J. (1992, enero 1). Machine-learning algorithms for credit-card applications. *IMA Journal of Management Mathematics*, 4(1), 43-51.
- de Miguel Domínguez, J. C., Miranda Torrado, F., Pallas Gonzalez, J., Peraza Fandiño, C., & Universidad de Santiago de Compostela. (n.d.). *La medición del riesgo de crédito y el nuevo Acuerdo de Capital del Comité de Basilea*.
- Domínguez, J., Torrado, F., González, J., & Fandiño, C. (2003). *La medición del riesgo de crédito y el nuevo acuerdo de capital del Comité de Basilea*.
- Durand, D. (1941, enero). Risk Elements in Consumer Installment Financing. *National Bureau of Economic Research*.

- Edwin Torrico, S., & Universidad Privada Boliviana. (2014, noviembre 20). *Macro Credit Scoring as a Proposal for Quantifying Credit Risk*.
- Efron, B. (1975). The efficiency of logistic regression compared to normal discriminant analysis. *Journal of the American Statistical Association*, 70(352), 892-898.
- Espin García, O., & Rodríguez Caballero, C. V. (2013). *Metodología para un scoring de clientes sin referencias crediticias*.
- Falbo, T. (1991). El impacto de los abuelos en los resultados de los niños en China. *Revisión de matrimonio y familia*, 16(3-4), 369-376.
- Fisher, R. A. (1936). *The use of multiple measurements in taxonomic problems* (Vol. 7). Ann. Eugenics.
- Freed, N., & Glover, F. (1981a). A linear programming approach to the discriminant problem. *Decision Sci.*, 12, 68-74.
- Freed, N., & Glover, F. (1981b). Simple but powerful goal programming formulations for the discriminant problem. *European J. Oper. Res.*, 7, 44-60.
- Friedman, J. H. (1977). A recursive partitioning decision rule for nonparametric classification. *IEEE Transaction on Computer*, 26(4), 404-509.
- Frydman, H., Altman, E., & Kao, D. (1985). Introducing recursive partitioning for financial classification: The case of financial distress. *The Journal of Finance*, 40(1), 269-291.
- Glover, F., Keene, S., & Duea, B. (1988). A new class of models for the discriminant problem. *Decision Sciences*, 19(2), 269-280.

Gonçalves, R., & Braga, M. (2008). Determinantes de risco de liquidez em cooperativas de crédito: uma abordagem a partir do modelo logit multinomial. *Revista de Administração Contemporânea*, 12(4), 1019-1041.

Gutierrez Girault, M. A., & Banco Central de la República Argentina. (2007, octubre). *Modelos de credit scoring: qué, cómo, cuándo y para qué*.

Kim, J. (2005, mayo). *A Credit Risk Model for Agricultural Loan Portfolios Under The New Basel Capital Accord*. Texas.

Kolesar, P., & Showers, J. L. (1985, febrero). A robust credit screening model using categorical data. *Management Science*, 31(2), 122-133.

Lam, K., Choo, E., & Moy, J. (1996). Minimizing deviations from the group mean: A new linear programming approach for the two-group classification problem. *European Journal of Operational Research*, 358-367.

Lawrence, E., & Arshadi, N. (1995). A multinomial logit analysis of problem loan resolution choices in banking. *Journal of Money, Credit and Banking*, 27(1), 202-216.

Lee, T.-S., Chiu, C.-C., Lu, C.-J., & Chen, I.-F. (2002). Credit scoring using the hybrid neural discriminant technique. *Expert Systems with Applications*, 23(3), 245-254.

Lewis, E. (1992). *An Introduction to Credit Scoring* (2nd ed.). Athena Press.

Makowski, P. (1985). Credit scoring branches out: decision tree-recent technology. *Credit World*, 30-37.

Marais, M. L., Patell, J., & Wolfson, M. (1984). The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications. *Journal of Accounting Research*, 22(1), 87-114.

Moody's KMV Company. (2017, febrero 18). Tratamiento de Cuestiones Planteadas por el Método Basado en Calificaciones Internas del Nuevo Convenio de Capital de Basilea. *Plataforma de Calificación Interna de Moody 's KMV y Métodos Basados en Calificaciones Internas (MÉTODO IRB) de Basilea II*.

Moral Rincón, M. J. (Ed.). (2014). *Nuevos Negocios Bancarios*.

Myers, J. H., & Forgy, E. W. (1963, septiembre). The Development of Numerical Credit Evaluation Systems. *Journal of the American Statistical Association*, 58(303), 799-806.

Orgler, Y. (1970). A credit scoring model for commercial loans. *Journal of Money, Credit & Banking*, 435-445.

Pérez, F., & Fernández, H. (2007). Las Redes Neuronales y la Evaluación del Riesgo de Crédito. *Revista Ingenierías Universidad de Medellín*, 6(10), 77-91.

Plotnikci, B. (2005). Modelo de comportamiento y predicción de incumplimiento crediticio: el caso de empresas pyme en argentina. *Temas de Management*, 15-19.

Porté Beck, D. (2020, mayo 26). *Altman Z-Score: la fórmula para detectar empresas en riesgo de quiebra*. Bolsaexpertos.

Press, J., & Wilson, S. (1978, diciembre). Choosing between logistic regression and discriminant analysis. *Journal of the American Statistical Association*, 73(364), 699-705.

Puertas Medina, R., & Martí Selva, M. L. (2013). Análisis del Credit Scoring (R. Ratner Rochman, Ed.). *Revista de Administración de Empresas*, 53(3), 303-315.

Ripley, B. D. (1994). Neural networks and related methods for classification. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(3), 409-437.

Rodríguez Guevara, D. E., Becerra Arévalo, J. A., & Cardona Valencia, D. (2017). Modelos y metodologías de credit score para personas naturales: una revisión literaria. *Revista CEA*, 3(5), 13-28.

Saavedra García, M. L., & Saavedra García, M. J. (2018, octubre). Modelos para medir el riesgo de crédito de la banca.

Showers, J. L., & Chakrin, L. M. (1981, diciembre). Reducing uncollectable revenue from residential telephone customers. *Interfaces*, 11(6), 21-34.

Steenackers, A., & Goovaerts, M. (1989, marzo). A credit scoring model for personal loans. *Insurance: Mathematics and Economics*, 8(1), 31-34.

Tam, K. Y., & Kiang, M. Y. (1992, julio 1). Aplicaciones gerenciales de redes neuronales: el caso de las predicciones de quiebras bancarias. *Ciencias de la gestión.*, 38(7), 926-947.

Thomas, L. (2000). A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16(2), 149-172.

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit Scoring and Its Applications*. SIAM.

Tsaih, R., Liu, J., & Lien, Y.-L. (2004, octubre). Credit scoring system for small loans. *Decision Support System*, 38(1), 91-99.

Vargas Zuñiga, F. (2001). *Introducción al Pilar 1 de Basilea II*. Banco de España.

Westreicher, G. (2018, septiembre 07). *Comité de Basilea*. Economipedia.com.

Westreicher, G. (2018, septiembre 10). *Calificación Crediticia (Credit Scoring)*. Economipedia.

Wiginton, J. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, 15(3), 757-770.

