SADAF
DOCENTES DE ADMINISTRACIÓN FINANCIERA

46 Jornadas Nacionales de Administración Financiera
Septiembre 17 y 18, 2026

Inteligencia artificial aplicada al análisis financiero

Desarrollo de un *dashboard* interactivo para la predicción de fracasos empresariales en el mercado de capitales argentino

Ricardo A. Yagüe
Universidad Nacional de Cuyo

SUMARIO

1. Introducción y justificación del estudio
2. Marco teórico: Modelos de predicción de fracasos
3. La herramienta: diseño y arquitectura
4. Aplicación a casos reales: metodología
5. Resultados
6. Discusión: límites y contribuciones
7. Conclusiones

Para comentarios:
ricardo.yague@fce.uncu.edu.ar

Resumen

El presente trabajo describe el proceso de diseño, desarrollo y validación de una herramienta interactiva de análisis financiero basada en modelos académicos de predicción de fracasos empresariales, construida con asistencia de inteligencia artificial generativa (Claude, Anthropic). La herramienta integra siete modelos de predicción −Beaver (1966), Altman Z' para mercados emergentes (2006), Dutto-Gallacher (2009), Alonso *et al*. (SADAF), Milanesi (2015/2019), Pascale (1988) y Sandín-Porporato (2007)−, varios de ellos presentados en ediciones anteriores de las Jornadas de Administración Financiera.

Los modelos se aplican a una muestra depurada de treinta y seis empresas del mercado de capitales argentino en el período 2024-2025: seis que entraron en default o concurso preventivo y treinta que cumplieron regularmente sus obligaciones. La ampliación del grupo de control desde cinco hasta treinta empresas —respecto de la versión preliminar de este trabajo— modifica sustancialmente las conclusiones sobre el poder discriminante de los modelos individuales. Se incorpora además un grupo separado de seis entidades financieras que opera como control de inaplicabilidad.

El trabajo aporta cuatro contribuciones: (i) un protocolo documentado y auditable de carga e implementación de modelos académicos con asistencia de IA; (ii) evidencia de que la mayoría de los modelos clásicos exhibe tasas de falsos positivos superiores al 45 % sobre grandes empresas cotizantes argentinas, mientras que el índice de consenso ponderado separa ambos grupos sin superposición; (iii) el diagnóstico de una anomalía de escala en la implementación del modelo Sandín-Porporato que anula su capacidad discriminante; y (iv) la demostración de que existen poblaciones (las entidades financieras) en las que estos modelos no producen predicciones degradadas sino predicciones sin contenido informativo.

1. Introducción y justificación del estudio

La predicción de dificultades financieras empresariales constituye uno de los problemas más relevantes de la teoría financiera aplicada. Desde los trabajos seminales de Beaver (1966) y Altman (1968), una vasta literatura ha documentado modelos capaces de anticipar el fracaso con base en información contable pública. Sin embargo, la aplicabilidad de estos modelos a contextos específicos (en particular mercados emergentes con alta volatilidad macroeconómica, cambios abruptos de régimen y restricciones cambiarias) plantea interrogantes que el presente trabajo busca abordar. A ello se suma una limitación adicional: muchos de estos modelos resultan inaplicables a entornos distintos de aquellos para los cuales fueron calibrados, de modo que aplicar los modelos de Beaver o Altman a economías latinoamericanas puede arrojar resultados que no se condicen con la realidad local.

El bienio 2024-2025 ofreció al mercado de capitales argentino un conjunto inusualmente rico de eventos de fracaso: Celulosa Argentina, Generación Mediterránea, Los Grobo Agrope-

cuaria, Red Surcos, Bioceres S.A. y SanCor Cooperativas acumularon deudas en default superiores a los USD 2.000 millones, en un entorno donde simultáneamente el país normalizaba su macroeconomía y empresas como Pampa Energía, Ternium Argentina e YPF accedían fluidamente al mercado de obligaciones negociables. Esta dicotomía convierte al período en un caso de estudio excepcional para evaluar el poder predictivo de los modelos clásicos.

Al mismo tiempo, el surgimiento de modelos de lenguaje de gran escala como herramientas de asistencia al desarrollo profesional abre una nueva dimensión metodológica: en qué medida puede la IA generativa acelerar y enriquecer la construcción de herramientas analíticas complejas, y bajo qué condiciones el resultado de ese proceso resulta auditable por un tercero. Este trabajo documenta esa experiencia de forma transparente, describiendo el proceso iterativo de co-creación entre el autor y el modelo Claude (Anthropic), y proponiendo un protocolo explícito de trazabilidad.

Esta versión del trabajo incorpora una ampliación sustancial del grupo de control respecto de la versión preliminar. Aquella se apoyaba en cinco empresas solventes; la presente utiliza treinta. La ampliación no fue neutral: varias conclusiones de la versión anterior sobre el desempeño relativo de los modelos individuales no sobreviven al aumento de la muestra, y ese hecho constituye en sí mismo uno de los resultados del trabajo. Un grupo de control reducido tiende a sobreestimar el poder discriminante de modelos que, aplicados a un conjunto amplio de empresas, clasifican como fracaso a la mayoría de las solventes.

Los objetivos del estudio son:

- Construir una herramienta interactiva que implemente los siete modelos de predicción seleccionados sobre datos reales del mercado argentino, con una arquitectura que permita auditar la correspondencia entre cada especificación publicada y su implementación.
- Evaluar el poder discriminante de cada modelo sobre una muestra que incluya un grupo de control suficientemente amplio como para estimar tasas de falsos positivos.
- Identificar las limitaciones de validez externa de modelos calibrados con submuestras específicas cuando se aplican a grandes empresas cotizantes de sectores de capital intensivo.
- Delimitar las poblaciones en las que estos modelos no resultan aplicables, distinguiendo la pérdida de precisión de la ausencia de contenido informativo.
- Documentar la metodología de desarrollo asistido por IA y extraer lecciones para su uso en el ámbito académico y profesional.

2. Marco teórico: Modelos de predicción de fracasos

La literatura sobre predicción de insolvencia empresarial reconoce tres grandes familias metodológicas: análisis univariado de ratios contables (Beaver, 1966); modelos multivariados de discriminante lineal y regresión logística (Altman, 1968; Ohlson, 1980); y modelos basados en teoría de opciones y precios de mercado (Merton, 1974; Milanesi, 2015). Los siete modelos

implementados en la herramienta cubren las tres familias y aportan calibraciones realizadas total o parcialmente sobre empresas argentinas o latinoamericanas.

2.1 Modelo de Beaver (1966)

William Beaver publicó en 1966 el primer modelo estadístico formal para la predicción de quiebras empresariales. Su enfoque es univariado: analiza cada ratio financiero de forma independiente y compara sus medias entre una muestra de empresas sanas y otra de empresas fracasadas.

Beaver analizó los estados contables de 79 empresas en bancarrota pertenecientes a 38 industrias y un grupo de control de igual tamaño, evaluando 30 indicadores financieros. Su hallazgo más relevante fue que el ratio flujo de fondos sobre pasivo total es el predictor con mayor poder discriminante, seguido por ganancia neta sobre activo total y deuda sobre activo total. Las cuatro proposiciones básicas del modelo, *ceteris paribus*, son: a mayores reservas de efectivo, menor probabilidad de fracaso; a mayor flujo de efectivo operativo, menor probabilidad de fracaso; a mayor apalancamiento, mayor probabilidad de fracaso; y a mayores gastos operativos relativos, mayor probabilidad de fracaso.

En la implementación de la herramienta, el modelo de Beaver se operacionaliza evaluando cinco señales de riesgo, con el veredicto siguiente: cuatro o más señales activas indican fracaso; dos o tres señales, zona gris; menos de dos, empresa sana. Corresponde señalar que esta regla de agregación no forma parte del modelo original, que es estrictamente univariado y no define un criterio de decisión conjunto.

> Gnecco *et al*. (2020) aplicaron el modelo de Beaver a una muestra de 88 empresas de Mar del Plata y reportaron un 78 % de coincidencia global, con una tasa de acierto del 85 % para empresas sanas y del 45 % para fracasos, evidenciando la asimetría característica del modelo.

2.2 Modelo Z' de Altman para mercados emergentes (2006)

Edward Altman (1968) introdujo el análisis discriminante múltiple al campo de la predicción de quiebras, permitiendo considerar varias variables en forma simultánea. Su modelo original fue calibrado con empresas manufactureras cotizantes de Estados Unidos. La versión implementada es la adaptación para países emergentes (Altman & Hotchkiss, 2006), que reemplaza el valor de mercado del patrimonio por el valor contable, elimina la variable de rotación y ajusta los coeficientes:

$$Z' = 6{,}56 \cdot X_1 + 3{,}26 \cdot X_2 + 6{,}72 \cdot X_3 + 1{,}05 \cdot X_4 + 3{,}25$$

Donde X_1 es capital de trabajo sobre activo total; X_2, ganancias retenidas sobre activo total; X_3, EBIT sobre activo total; y X_4, patrimonio neto sobre pasivo total. La constante 3,25 estandariza la escala de modo que un puntaje de cero equivalga a un bono con calificación D. Las zonas de clasificación implementadas son: Z' menor a 4,15, alta probabilidad de quiebra; entre 4,15 y 5,85, zona gris; superior a 5,85, baja probabilidad de quiebra.

Nota sobre el punto de corte inferior: la convención más difundida en la literatura ubica el límite de la zona de distress del modelo EMS en 4,35, resultante de sumar la constante de 3,25 al corte de 1,10 del modelo Z'' subyacente. Existen versiones que lo sitúan en 3,75. El valor de 4,15 adoptado en la herramienta es intermedio y se conserva por consistencia con la versión preliminar del trabajo; su sustitución por 4,35 no altera ningún veredicto de la muestra.

Gnecco *et al*. (2020) reportaron para el modelo de Altman una coincidencia global del 66 % en la muestra de Mar del Plata, con 86 % de acierto en empresas sanas y solo 25 % en fracasos, evidenciando sesgos hacia falsos negativos.

2.3 Modelo Dutto-Gallacher (2009)

Martín Dutto y Marcos Gallacher (2009) desarrollaron un modelo de regresión logística calibrado específicamente con empresas constructoras de las provincias de Santa Fe y de la región de Rosario en el período 2007-2008. La variable dependiente es dicotómica y las variables independientes resultantes tras la selección son:

$$Z = 1{,}5796 - 1{,}5311 \cdot \log(\text{Activo}/\text{Ventas}) + 0{,}6568 \cdot \log(\text{Disp.}/\text{Deuda CP}) - 2{,}4554 \cdot (\text{Deuda}/\text{Activo})$$

La probabilidad de no fracaso se obtiene mediante la función logística $P(\text{sano}) = 1/(1+e^{-z})$. El umbral original del paper es $P(\text{sano}) \geq 0{,}50$. Como se discute en la sección 5.3, ese umbral no resulta adecuado para empresas de capital intensivo, dado que el coeficiente dominante penaliza sistemáticamente a las empresas con alta intensidad de activos fijos respecto de sus ventas anuales.

El modelo logró una precisión del 72% al 75% sobre la muestra original de constructoras. Dutto (2008) había demostrado previamente, mediante análisis univariado, que las variables más discriminantes son ventas sobre pasivo, disponibilidades sobre deuda de corto plazo, deuda sobre activo, flujo de fondos sobre deuda y activo corriente sobre ventas.

El presente ensayo identifica una limitación de validez externa crítica: el modelo fue calibrado con constructoras pyme cuyo ratio Activo sobre Ventas oscila en el rango de 1,0 a 1,5. Aplicado a grandes empresas de capital intensivo −energía, acero, cemento, gas− con ratios de 2 a 7, el modelo colapsa sistemáticamente hacia probabilidades de solvencia inferiores al 12 %, con independencia de la situación financiera real.

2.4 Modelo Alonso *et al.* sobre pymes bonaerenses

Aldo Alonso, Santiago Barraza, Gustavo Guglielmetti y Ana Legato (Universidad Nacional del Centro de la Provincia de Buenos Aires, SADAF XXVI) desarrollaron un modelo logístico de predicción de default para pymes del centro de la Provincia de Buenos Aires. La metodología incluye depuración de variables mediante índice de Gini, eliminación de colinealidades y selección forward de covariables. Las tres variables seleccionadas son:

$$Z = -2{,}310 + 2{,}951 \cdot (\text{Créd. Ventas/Deuda}) + 18{,}347 \cdot (\text{Disponibilidades/Activo}) + 0{,}693 \cdot (\text{PN/Deuda})$$

El modelo alcanzó un poder predictivo del 71,4 % sobre la muestra de desarrollo. Los errores de Tipo I (asignar bajo riesgo a un incumplidor) fueron del 26,1 %, y los de Tipo II (asignar alto riesgo a un cumplidor) del 30,8 %. El coeficiente de mayor peso, 18,347 sobre disponibilidades sobre activo, refleja la importancia crítica de la liquidez inmediata para pymes con acceso limitado al crédito.

La implementación en la herramienta no contempla zona gris para este modelo: el veredicto es binario según la probabilidad estimada supere o no el 50 %. Esta característica, que replica la formulación original, tiene consecuencias sobre la comparabilidad con los demás modelos que se discuten en la sección 5.2.

2.5 Modelo de opciones barrera de Milanesi (2015/2019)

Gastón Milanesi (Universidad Nacional del Sur) propone una familia de modelos basados en la teoría de opciones exóticas del tipo knock-out para valuar el patrimonio de la empresa y estimar probabilidades de insolvencia.

El punto de partida es la observación de Black y Scholes (1973) y Merton (1974) de que el patrimonio de una empresa apalancada puede asimilarse a una opción de compra europea sobre sus activos. La debilidad de este enfoque es que el valor de la opción crece con la volatilidad del subyacente, lo que conduciría a la conclusión inaceptable de que mayor riesgo implica mayor valor patrimonial. Las opciones barrera del tipo down-and-out call corrigen esa debilidad: el contrato se extingue si el valor de los activos cae por debajo de un nivel definido como barrera de insolvencia, produciendo la relación inversa entre volatilidad y valor del patrimonio que resulta coherente con la lógica empresarial.

La probabilidad total de insolvencia incorpora tanto la probabilidad de default al vencimiento como la probabilidad de que los activos atraviesen la barrera en cualquier instante previo, lo que resulta más informativo que los modelos contables para evaluar empresas con deuda en moneda extranjera y alta volatilidad de activos.

> Milanesi *et al*. (2019) aplicaron el modelo a 53 empresas cotizantes del mercado argentino a diciembre de 2018 y encontraron una correlación entre volatilidad del activo y probabilidad de default de 0,28 para el modelo de opciones barrera frente a 0,18 para el modelo naíf basado en Black-Scholes.

2.6 Modelo de Pascale (1988)

Ricardo Pascale desarrolló en 1988 un modelo de análisis discriminante múltiple calibrado con empresas de Uruguay, con validaciones posteriores para el contexto latinoamericano. Su relevancia reside en ser uno de los pocos modelos clásicos calibrado en la región, con variables y coeficientes ajustados a las características de las economías emergentes del Cono Sur.

$$Z = -2{,}9511 + 0{,}2756 \cdot (\text{UN/Ventas}) + 2{,}6271 \cdot (\text{PN/Pasivo}) + 0{,}1364 \cdot (\text{AC/PC}) + 0{,}0469 \cdot (\text{Ventas/Activo})$$

Las cuatro variables capturan la rentabilidad sobre ventas, la solvencia estructural, la liquidez corriente y la eficiencia en el uso de activos. Las zonas de clasificación son: Z superior a 0,50, empresa sana; entre -0,50 y 0,50, zona gris; Z igual o inferior a -0,50, fracaso.

El modelo tiene la ventaja de haber sido calibrado en un entorno macroeconómico similar al argentino. Su debilidad, señalada en la literatura, es el tamaño reducido de la muestra de calibración original. Como se documenta en la sección 5.2, la estructura de la ecuación (una constante negativa de magnitud considerable que solo el término de solvencia puede compensar) impone un requisito implícito de patrimonio neto superior al pasivo total para alcanzar la zona de solvencia, condición que la mayoría de las grandes empresas cotizantes argentinas no satisface.

2.7 Modelo de Sandín y Porporato (2007)

Montserrat Sandín (Universidad Carlos III) y Marcelo Porporato (York University) publicaron en 2007 un modelo logístico calibrado con empresas argentinas del período 1995-2001, incluyendo la crisis de la Convertibilidad, lo que lo convierte en el modelo de mayor especificidad geográfico-temporal para el análisis del riesgo corporativo en Argentina.

$$Z = 6{,}507 - 6{,}041 \cdot (\text{UN/Activo}) - 2{,}304 \cdot (\text{EBIT/Activo}) + 0{,}025 \cdot (\text{Pasivo/PN}) - 0{,}659 \cdot (\text{Ventas/Activo})$$

La probabilidad de fracaso se obtiene como $P = \frac{1}{1+e^{-Z}}$, con umbrales de 0,50 para fracaso y 0,35 para el límite inferior de la zona gris. Los signos de los coeficientes son intuitivos: mayor rentabilidad reduce el riesgo, mayor apalancamiento lo aumenta, y el signo negativo sobre rotación es coherente con el hallazgo de Dutto-Gallacher respecto de las empresas de capital intensivo.

La aplicación de este modelo a la muestra ampliada revela una anomalía que se analiza en detalle en la sección 5.4 y que compromete su utilidad en la implementación actual.

2.8 Consenso ponderado: integración de los siete modelos

La herramienta implementa un sistema de consenso ponderado que agrega los veredictos individuales en un índice único de riesgo. Los pesos asignados reflejan la especificidad geográfica del modelo, la complejidad metodológica de la técnica empleada y el desempeño documentado en la literatura previa.

El índice se calcula como el promedio ponderado de los veredictos individuales, asignando valor 1,0 al fracaso, 0,5 a la zona gris y 0,0 a empresa sana, y normalizando por la suma de los pesos efectivamente aplicados. Un índice superior a 0,60 indica riesgo elevado; entre 0,35 y 0,60, zona de incertidumbre; inferior a 0,35, situación de solvencia según el agregado.

Corresponde declarar explícitamente que estos pesos constituyen un juicio experto del autor y no el resultado de una optimización estadística sobre la muestra. Una asignación derivada de un procedimiento de optimización sobre estos mismos datos incurriría en circularidad, y el desempeño del índice que se reporta en la sección 5.6 dejaría de ser informativo.

Tabla 1: Ponderación de los modelos considerados

Modelo	*Referencia*	*Peso*	*Justificación*
Beaver	1966	10%	Univariado — análisis complementario
Altman Z'	2006	20%	Mayor difusión y validación internacional
Dutto-Gallacher	2009	15%	Calibrado en Argentina, con ajuste por sector
Alonso et al.	SADAF	15%	Calibrado en pymes bonaerenses
Milanesi	2015/19	15%	Incorpora volatilidad y deuda dolarizada
Pascale	1988	10%	Calibrado en LatAm, contexto regional
Sandín-Porporato	2007	15%	Calibrado en Argentina 1995-2001

3. La herramienta: diseño y arquitectura

3.1 Descripción funcional del *dashboard*

El *dashboard* es una aplicación web contenida en un único archivo HTML de aproximadamente 1.300 líneas, que integra dos módulos funcionales accesibles mediante una barra de navegación.

Módulo 1 - ONcalc: calculadora de obligaciones negociables. Permite seleccionar un instrumento de la base de datos de ONs vigentes en el mercado argentino, ingresar precio de compra y parámetros de valuación, y calcular en tiempo real la tasa interna de retorno, la duración de Macaulay, la duración modificada, la convexidad, el precio sucio y los intereses devengados. Incluye la proyección completa del flujo de fondos descontado en formato tabular y gráfico.

Módulo 2 - Fracasos corporativos Argentina 2024-2025. Implementa los siete modelos académicos sobre una base de cuarenta y siete emisores, con selector interactivo por empresa, panel de indicadores clave, tabla de ratios contra referencias históricas, cronología del caso, tabla comparativa cruzada de todos los modelos contra todas las empresas, control de umbral para el modelo Dutto-Gallacher y panel de consenso ponderado. De esos cuarenta y siete emisores, treinta y seis integran la muestra analizada en la sección 5 y el resto se excluye conforme a los criterios de la sección 4.1.

La interfaz fue diseñada con criterio editorial −tipografías Playfair Display, IBM Plex Sans e IBM Plex Mono, paleta inspirada en papel de periódico− para facilitar la lectura densa de información financiera cuantitativa.

Tabla 2: Características del dashboard

Característica	*Especificación técnica*
Tecnología	HTML5 / CSS3 / JavaScript ES5, sin dependencias externas de terceros
Portabilidad	Archivo único .html; corre en cualquier navegador sin servidor de aplicaciones
Base de ONs	Se carga dinámicamente desde un repositorio público en GitHub; requiere conexión a internet para actualizarse
Emisores incluidos	47 en la base; 36 en la muestra analizada más 6 entidades financieras como control
Modelos implementados	Beaver, Altman Z', Dutto-Gallacher, Alonso et al., Milanesi, Pascale, Sandín-Porporato
Cálculo en tiempo real	TIR, duración, convexidad, probabilidades de default y consenso ponderado
Control de umbral	Selector interactivo de 40% a 80% para el modelo Dutto-Gallacher

3.2 Protocolo de carga e implementación de los modelos

La expresión *cargar un modelo* en una herramienta asistida por inteligencia artificial designa tres operaciones distintas que conviene separar analíticamente, porque cada una tiene un riesgo de error propio y exige una verificación diferente. La primera es la extracción de la especificación desde el paper original: identificar la forma funcional, las variables explicativas con su definición contable exacta, los coeficientes estimados y los puntos de corte. La segunda es la normalización: reducir siete definiciones heterogéneas de variables (cada autor nombra y construye sus ratios de manera distinta) a un conjunto común de magnitudes contables observables. La tercera es la transcripción a código ejecutable y su verificación numérica.

Esta sección documenta las tres operaciones de forma que un tercero pueda auditar o replicar la implementación. El criterio metodológico adoptado fue que la especificación operativa de cada modelo debe ser recuperable desde el artefacto de software y no desde el registro de la conversación con el asistente de IA. Las razones de este criterio se exponen en el apartado 3.2.5.

1) Corpus de entrada y extracción de la especificación

El proceso se inició con la carga simultánea de los papers originales al contexto del modelo de lenguaje, acompañada de una instrucción única de propósito general: analizar los modelos contenidos en los textos y construir una herramienta que permita evaluarlos sobre estados contables públicos. No se suministró una especificación técnica previa, ni un esquema de datos, ni una lista de variables. La decisión fue deliberada: el objetivo era observar en qué medida el asistente podía inferir la estructura común de un conjunto de modelos heterogéneos a partir del material académico crudo.

La respuesta identificó correctamente, en una única iteración, la pertenencia metodológica de cada modelo (univariado, discriminante múltiple, logístico o basado en teoría de opciones),

su conjunto de variables y su número de coeficientes. Este resultado no debe sobreinterpretarse: identificar la estructura de un modelo publicado es una tarea de reconocimiento sobre literatura ampliamente difundida, no un acto de inferencia original. El valor del paso está en la reducción del tiempo de transcripción, no en la generación de conocimiento.

Los modelos de Pascale (1988) y Sandín-Porporato (2007) se incorporaron en una etapa posterior del desarrollo, cuando la herramienta ya estaba operativa con los cinco modelos iniciales. Esa incorporación tardía produjo un episodio metodológicamente relevante que se documenta en el punto 5c.

2) El vector contable normalizado

El elemento central de la arquitectura −y la decisión de diseño que hace posible la comparación entre modelos− es que las siete implementaciones no leen estados contables. Leen un mismo vector normalizado de dieciocho campos, construido una sola vez por empresa. Trece campos corresponden a magnitudes contables observables y cinco son parámetros de valuación exigidos exclusivamente por el modelo de opciones barrera.

Tabla 3: Campos para el análisis

Los cinco campos sombreados son parámetros de valuación y no magnitudes contables observables: se estiman, y su elección constituye un supuesto del analista. El resto se lee directamente de los estados contables.

Campo	Magnitud contable	Campo	Magnitud contable
at	Activo total	vt	Ventas totales del ejercicio
ac	Activo corriente	eb	EBIT (resultado operativo)
pc	Pasivo corriente	un	Utilidad neta del ejercicio
pt	Pasivo total	ff	Flujo de fondos operativo
pn	Patrimonio neto	inv	Bienes de cambio
dp	Disponibilidades	gr	Ganancias retenidas
cv	Créditos por ventas	—	—
sg	Volatilidad del activo (σ)	ro	Rendimiento esperado del activo (μ)
rf	Tasa libre de riesgo	T	Horizonte de la deuda, en años
B	Barrera como fracción del pasivo		

La consecuencia de esta normalización es que las diferencias entre los veredictos de los siete modelos, para una misma empresa, son atribuibles enteramente a la especificación de cada modelo y no a diferencias en los datos de entrada. Se trata de un requisito indispensable para que la comparación de la sección 5 sea válida.

La normalización también hace explícito un problema que la lectura aislada de cada paper tiende a ocultar: autores distintos usan el mismo nombre para ratios que no son el mismo ratio.

El endeudamiento de Dutto-Gallacher es Pasivo total sobre Activo total; el de Sandín-Porporato es Pasivo total sobre Patrimonio neto; el de Pascale entra invertido, como Patrimonio neto sobre Pasivo total. Las tres construcciones responden a la misma intuición económica y tienen comportamientos numéricos radicalmente distintos, en particular ante patrimonio neto negativo.

3) Transcripción a código y contrato de salida

Cada modelo se implementó como una función independiente que recibe el vector contable y devuelve una estructura de formato común: identificación del modelo, año de referencia, valores intermedios de sus variables, el estadístico propio del modelo (puntaje discriminante, probabilidad logística o probabilidad de default) y un veredicto categórico en tres estados: fracaso, zona gris o empresa sana.

La uniformidad del formato de salida habilita tres funcionalidades que de otro modo requerirían tratamiento caso por caso: la tabla comparativa cruzada, el índice de consenso ponderado y la degradación controlada ante datos insuficientes. Respecto de esto último, la función que ejecuta los siete modelos en secuencia descarta los que no pueden computarse y el consenso renormaliza los pesos sobre los modelos efectivamente disponibles.

El modelo de Milanesi incorpora además una degradación interna: la rutina calcula en primer término la probabilidad de default de Merton mediante la distancia al default, y solo después intenta la valuación completa con barrera de absorción. Si el cálculo de la barrera no converge (situación posible ante combinaciones extremas de volatilidad y horizonte), el resultado retenido es el de Merton. La herramienta nunca devuelve un valor indefinido; devuelve el modelo más simple de la misma familia.

La función de distribución normal acumulada, requerida por Merton y por la formulación de barrera, se implementó mediante la aproximación polinómica de Abramowitz y Stegun, con error absoluto acotado en el orden de 10^{-7}, suficiente para la precisión de los datos contables de entrada.

4) Casos límite y guardas numéricas

Los modelos publicados se especifican sobre poblaciones de empresas en funcionamiento, donde los denominadores de sus ratios son positivos. Una herramienta que se aplica precisamente a empresas en crisis debe resolver situaciones que los papers originales no contemplan: patrimonio neto negativo, ventas nulas, pasivo corriente inexistente.

La implementación adopta dos guardas. La primera es la adición de un épsilon de 0,001 a los denominadores de los ratios en los modelos de Dutto-Gallacher, Alonso, Pascale y Sandín-Porporato, que evita la división por cero. La segunda acota el argumento del logaritmo a un mínimo positivo en el modelo de Dutto-Gallacher, cuya forma funcional aplica logaritmos a dos de sus tres variables.

Estas guardas resuelven el problema numérico pero no el problema económico, y corresponde declararlo: con patrimonio neto negativo, el ratio Pasivo sobre Patrimonio neto que utiliza Sandín-Porporato adopta valor negativo, y el modelo lo interpreta como apalancamiento bajo. El coeficiente de esa variable es positivo, de modo que un patrimonio negativo reduce el puntaje y empuja el veredicto hacia empresa sana. En el caso de Celulosa Argentina (patrimonio neto de -$23.774 millones) el ratio resultante es de -14,7 y el veredicto final se

mantiene en fracaso únicamente porque los términos de rentabilidad, cuyos coeficientes son de magnitud muy superior, dominan la ecuación. El resultado es correcto por compensación, no por construcción. Se trata de una limitación estructural del modelo original al aplicarse a empresas con patrimonio consumido, no de un defecto de la transcripción.

5) Intervenciones deliberadas sobre la especificación original

Dos elementos de la implementación se apartan de las especificaciones publicadas. Ambos son decisiones del autor y se declaran aquí de manera explícita, porque su omisión comprometería la replicabilidad del trabajo.

a) Corte adaptativo en el modelo Dutto-Gallacher. El punto de corte del modelo original es una probabilidad estimada de 0,50, valor estándar en modelos logísticos. Su aplicación directa a la muestra clasifica como fracaso a la totalidad de las empresas de capital intensivo, incluidas las solventes, por las razones que se analizan en la sección 5.3.

La implementación incorpora en consecuencia un corte condicional: cuando la relación activo sobre ventas supera 2,0, los umbrales se desplazan a 0,08 para empresa sana y 0,03 para zona gris; por debajo de ese valor rigen los umbrales originales de 0,50 y 0,40. Este corte no proviene del paper. Es una recalibración empírica derivada de la propia muestra y, como tal, está sujeta a la crítica de sobreajuste: los umbrales fueron elegidos observando la separación efectiva entre los dos grupos. Su validación exige una muestra independiente. Se lo conserva porque, sin él, el modelo no aporta información discriminante alguna sobre grandes empresas cotizantes, y porque el corte hace visible el problema en lugar de disimularlo.

b) Ponderaciones del consenso. Los pesos de la sección 2.8 constituyen un juicio experto del autor y no el resultado de una optimización estadística sobre la muestra, según se declara allí.

c) Trazabilidad: el artefacto como fuente de verdad. Durante el desarrollo se produjo un episodio que resultó ser el hallazgo metodológico más relevante del proceso. Al redactar la sección 2 de este trabajo (en la misma sesión en que se habían incorporado los modelos de Pascale y Sandín-Porporato) el asistente no pudo recuperar los coeficientes de esos dos modelos: la extensión de la conversación había desplazado de su ventana de contexto el material que él mismo había producido horas antes. La especificación se recuperó leyendo el código fuente del dashboard.

La conclusión metodológica es directa y aplica a cualquier trabajo asistido por modelos de lenguaje. El diálogo con el asistente no es un registro confiable: es volátil, no está versionado y no resulta auditable por un tercero. El artefacto de software sí lo es. En consecuencia, el criterio adoptado fue que ningún coeficiente, umbral o definición de variable se considera establecido hasta estar escrito en el código, y que toda verificación posterior se realiza contra el código y contra el paper original, nunca contra lo afirmado en la conversación.

Esta observación responde por anticipado a la objeción central que puede formularse a un trabajo de esta naturaleza: la posibilidad de que los coeficientes hayan sido generados

por confabulación del modelo de lenguaje en lugar de extraídos de las fuentes. La respuesta no es una garantía sobre el comportamiento del asistente, que no puede ofrecerse, sino un procedimiento: la especificación operativa está íntegramente contenida en un artefacto inspeccionable y es contrastable, línea por línea, contra las publicaciones originales. La tabla 4 sistematiza esa correspondencia. La sección 5.4 documenta un caso concreto en que este procedimiento permitió detectar una discrepancia que la lectura del texto por sí sola no habría revelado.

Tabla 4: Trazabilidad

CT = capital de trabajo; GR = ganancias retenidas; FF = flujo de fondos operativo; CV = créditos por ventas; Disp = disponibilidades. La columna de desvíos identifica toda diferencia entre la especificación publicada y la implementada.

Modelo	*Forma funcional implementada*	*Umbrales*	*Desvío respecto del original*
Beaver (1966)	Recuento de 5 señales binarias: FF/PT<0,30; UN/AT<0; PT/AT>0,50; CT/AT<0,10; AC/PC<1,50	≥4 señales → fracaso; 2-3 → gris; <2 → sana	Operacionalización propia: el modelo original es univariado y no define regla de agregación
Altman Z' EMS (2006)	Z = 3,25 + 6,56·CT/AT + 3,26·GR/AT + 6,72·EBIT/AT + 1,05·PN/PT	Z<4,15 → fracaso; 4,15-5,85 → gris; >5,85 → sana	Coeficientes sin modificación. Corte inferior en 4,15; la convención más difundida lo ubica en 4,35 (ver nota en 2.2)
Dutto-Gallacher (2009)	Z = 1,5796 − 1,5311·ln(AT/VT) + 0,6568·ln(Disp/PC) − 2,4554·PT/AT; P = 1/(1+e⁻ᶻ)	Adaptativo según AT/VT (ver 3.2.5.a)	Coeficientes sin modificación; puntos de corte recalibrados sobre la muestra
Alonso *et al*.	Z = −2,310 + 2,951·CV/PT + 18,347·Disp/AT + 0,693·PN/PT; P = 1/(1+e⁻ᶻ)	P≥0,50 → sana; P<0,50 → fracaso	Sin zona gris: la implementación es binaria, a diferencia de los demás modelos
Milanesi (2015/2019)	Opción barrera de absorción sobre el valor del activo, con retroceso a distancia al default de Merton	PD>0,55 → fracaso; 0,30-0,55 → gris	Tasa de dividendos fijada en 0,03 y no parametrizada; barrera expresada como fracción del pasivo
Pascale (1988)	Z = −2,9511 + 0,2756·UN/VT + 2,6271·PN/PT + 0,1364·AC/PC + 0,0469·VT/AT	Z>0,50 → sana; −0,50 a 0,50 → gris; Z≤−0,50 → fracaso	Sin modificación
Sandín-Porporato (2007)	Z = 6,507 − 6,041·UN/AT − 2,304·EBIT/AT + 0,025·PT/PN − 0,659·VT/AT; P = 1/(1+e⁻ᶻ)	P>0,50 → fracaso; 0,35-0,50 → gris; P≤0,35 → sana	Variables ingresadas como fracciones decimales; ver anomalía de escala en 5.4

3.3 Rol de la inteligencia artificial en el desarrollo

El desarrollo de la herramienta se realizó íntegramente con asistencia del modelo Claude (Anthropic), a través de una conversación iterativa de aproximadamente treinta turnos. El proceso constituye en sí mismo un objeto de análisis metodológico sobre las capacidades y limitaciones de la IA generativa como asistente de desarrollo profesional, y puede describirse en cinco fases.

Fase 1: Diseño conceptual y marco teórico

Consistió en la carga y procesamiento de los papers académicos, conforme al protocolo descripto en 3.2.1. El asistente analizó los textos, extrajo las fórmulas, identificó los coeficientes de calibración y los valores de referencia de los grupos de empresas sanas y fracasadas. El análisis habría demandado varias horas de trabajo manual; con asistencia de IA tomó aproximadamente veinte minutos.

Fase 2: Implementación matemática

La implementación de los modelos requirió traducir fórmulas académicas escritas en notación matemática a código ejecutable, incluyendo la función de distribución normal acumulada, el algoritmo de Newton-Raphson para el cálculo iterativo de la TIR, las fórmulas de duración y convexidad, y las expresiones de la teoría de opciones barrera. El código generado resultó sintácticamente correcto en el primer intento en todos los casos. Corresponde precisar que la corrección sintáctica y la corrección de la transcripción son cosas distintas: la anomalía documentada en la sección 5.4 permaneció inadvertida durante toda esta fase y solo se hizo visible al ampliar el grupo de control.

Fase 3: Recolección y carga de datos

Los datos financieros fueron obtenidos de comunicados a la CNV, reportes de calificadoras, prensa especializada y estados contables publicados. El asistente colaboró en la extracción de datos relevantes de los textos de los comunicados, estimó valores no disponibles de forma consistente con los ratios sectoriales y documentó el nivel de confianza de cada estimación.

Fase 4: Depuración y refinamiento iterativo

Fue la fase más reveladora del proceso. Se detectaron y corrigieron iterativamente errores de sintaxis por apóstrofes dentro de cadenas de texto, cierres de arreglos duplicados por ediciones sucesivas, incompatibilidades de interfaces de programación en el entorno de ejecución y problemas de alcance con variables definidas después de las funciones que las referencian. El asistente demostró capacidad diagnóstica efectiva, pero también tendencia a la acumulación de errores en ediciones sucesivas sobre archivos extensos, lo que sugiere que para proyectos de mediana complejidad un entorno con acceso directo al sistema de archivos resulta más eficiente que el chat conversacional.

Fase 5: Análisis crítico de resultados

Los dos hallazgos analíticos del trabajo —la limitación de validez externa del modelo Dutto-Gallacher y la anomalía de escala del modelo Sandín-Porporato— emergieron en esta fase. En ambos casos la secuencia fue la misma: el autor identificó un resultado incompatible con las expectativas teóricas, y el asistente ejecutó en minutos la descomposición matemática que permitió localizar la causa. La distinción es relevante y se retoma en la sección 6.3: el reconocimiento de un resultado como anómalo requiere conocimiento del dominio; el diagnóstico de por qué se produce es la parte del trabajo que la IA acelera de manera sustantiva.

3.4 Módulo ONcalc: Valuación de obligaciones negociables

El módulo ONcalc implementa la metodología estándar de valuación de renta fija conforme a los criterios del IAMC y BYMA para el mercado local. El algoritmo central es la resolución iterativa (Newton-Raphson, hasta 300 iteraciones con tolerancia de 10^{-9}) de la ecuación de precio implícita:

$$P = \frac{\sum FC_i}{(1+TIR)^{t_i/365}}$$

donde FC_i es el flujo neto en la fecha t_i, expresada en días desde la fecha de compra.

La herramienta soporta bonos bullet y amortizantes lineales, en pesos o en dólares, con periodicidad de cupón configurable. Los resultados incluyen precio limpio, precio sucio, intereses devengados, duración de Macaulay, duración modificada y convexidad, calculados según las relaciones estándar de la teoría de renta fija.

La base de instrumentos se carga dinámicamente desde un repositorio público, de modo que la actualización del universo de ONs no requiere modificar el archivo de la herramienta. Esta decisión de diseño tiene como contrapartida que el módulo requiere conexión a internet para refrescar su base, a diferencia del módulo de predicción de fracasos, que es enteramente autónomo.

4. Aplicación a casos reales: metodología

4.1 Construcción y depuración de la muestra

La base del *dashboard* reúne cuarenta y siete emisores del mercado argentino con obligaciones negociables vigentes o en situación de incumplimiento durante el período analizado. No todos ellos resultan aptos para integrar la muestra sobre la que se evalúa el poder discriminante de los modelos. Esta sección documenta los criterios de inclusión y las exclusiones practicadas.

Criterio de inclusión en el grupo de fracasos: haber declarado concurso preventivo o quiebra, o haber incumplido pagos de obligaciones negociables, durante el período enero 2024 - julio 2025, con estados contables públicos disponibles al momento del análisis.

Criterio de inclusión en el grupo de control: ser emisor con obligaciones negociables vigentes al 31/12/2024, sin antecedentes de incumplimiento en el período, con estados contables auditados disponibles.

Exclusiones practicadas

Emisores del sector público (cuatro casos). La Provincia de Buenos Aires, la Provincia de Mendoza, la Provincia de Salta y la Municipalidad de Córdoba fueron excluidas en su totalidad. Los siete modelos operan sobre magnitudes de estados contables empresariales —patrimonio neto, ventas, EBIT, activo corriente— que no tienen equivalente en la contabilidad presupuestaria de un ente público. No se trata de una limitación de validez externa sino de una imposibilidad de aplicación: las variables de entrada no existen.

Entidades financieras (siete casos, seis entidades económicas). Los seis bancos y el holding financiero de la base se retiran del cómputo del poder discriminante y se reservan como grupo de control de inaplicabilidad, cuyo tratamiento se desarrolla en la sección 5.5. La justificación es que su estructura patrimonial responde a la lógica de la intermediación financiera y no a la de una empresa productiva, de modo que las variables de los modelos no conservan el significado económico que estos les atribuyen. Banco Supervielle S.A. y Grupo Supervielle S.A. figuran en la base como registros separados pese a corresponder a la misma entidad económica; se computa uno solo.

Practicadas ambas exclusiones, la muestra analizada queda conformada por treinta y seis casos: seis fracasos y treinta empresas del grupo de control.

Tabla 5: Casos de fracaso

El grupo de control abarca energía y servicios regulados, transporte y distribución de gas, electricidad, telecomunicaciones, retail, agroindustria, metalurgia, real estate, petróleo y tecnología. Su composición completa se detalla en la tabla 7 (sección 5.2).

Empresa	*Sector*	*Clasificación*	*Deuda USD*
Celulosa Argentina	Papel / Industrial	Fracaso — Concurso	128 M
Generación Mediterránea	Energía	Fracaso — Default ONs	1.100 M
Los Grobo Agropecuaria	Agro	Fracaso — Concurso	207 M
Red Surcos S.A.	Agroquímicos	Fracaso — Concurso	73 M
Bioceres S.A.	Agro / Biotech	Fracaso — Quiebra	39 M
SanCor Cooperativas	Lácteo	Fracaso — Concurso	250 M
30 emisores adicionales	*Diversos*	*Grupo de control*	*n/a*

La ampliación del grupo de control desde cinco hasta treinta empresas merece una observación metodológica. En la versión preliminar de este trabajo, con cinco empresas solventes, varios modelos aparentaban un desempeño razonable. Con treinta, la mayoría exhibe tasas de falsos positivos que invalidan su uso individual. El grupo de control reducido no era simplemente menos preciso: producía una conclusión cualitativamente distinta y equivocada. La asimetría es característica de este tipo de evaluación, porque los casos de fracaso son escasos por naturaleza mientras que el grupo de control puede ampliarse a voluntad, y la tentación de equilibrar ambos grupos conduce a sobreestimar sistemáticamente la especificidad de los modelos.

4.2 Fuentes de datos y criterio de clasificación

Los datos financieros fueron obtenidos de las siguientes fuentes primarias: estados contables publicados en la Comisión Nacional de Valores, auditados o con revisión limitada; comunicados de hechos relevantes publicados por las propias emisoras; informes de calificadoras de

riesgo (Moody's Local Argentina, FIX SCR, Standard & Poor's); y cobertura de fuentes periodísticas especializadas.

Para cada caso se utilizó el último ejercicio cerrado con información pública disponible antes del evento de fracaso, o el último ejercicio cerrado al 31/12/2024 en el caso de las empresas del grupo de control. Los parámetros de valuación requeridos por el modelo de Milanesi (volatilidad del activo, rendimiento esperado, horizonte y nivel de barrera) fueron estimados por sector y documentados como tales en la herramienta.

Corresponde señalar que una porción de los valores contables de la base son estimaciones construidas a partir de información fragmentaria, particularmente en los casos de empresas que dejaron de publicar estados contables completos tras el evento de fracaso. Esta limitación se analiza en la sección 6.1.

5. Resultados

5.1. Casos de fracaso: análisis comparado

La tabla 6 presenta el veredicto de cada modelo para los seis casos de fracaso. Ningún caso elude el diagnóstico negativo en la totalidad de los modelos y la unanimidad es prácticamente completa.

Tabla 6: Casos de fracaso según cada modelo

F = fracaso · G = zona gris · S = empresa sana. D-G = Dutto-Gallacher con corte adaptativo. S-P = Sandín-Porporato. El consenso es el índice ponderado de la sección 2.8.

Empresa	*Beaver*	*Altman*	*D-G*	*Alonso*	*Milanesi*	*Pascale*	*S-P*	*Consenso*
Celulosa Argentina	F	F	F	F	F	F	F	**100%**
Generación Mediterránea	G	G	F	F	F	F	F	**85%**
Los Grobo Agropecuaria	F	F	F	F	F	F	F	**100%**
Red Surcos S.A.	F	F	F	F	F	F	F	**100%**
Bioceres S.A.	F	F	F	F	F	F	F	**100%**
SanCor Cooperativas	F	F	F	F	F	F	F	**100%**

El caso de Generación Mediterránea merece atención especial. Es la única empresa del grupo donde los modelos de Beaver y Altman no clasifican directamente como fracaso, sino como zona gris, lo que resulta consistente con la realidad del caso: la empresa mantenía EBITDA positivo del orden de USD 117 millones, operaciones normales y acceso al mercado de capitales hasta setenta días antes del default. El problema era exclusivamente de estructura de financiamiento, con una relación Deuda sobre EBITDA de 11,4 veces en moneda dura. Se trata de una dimensión que los modelos contables tradicionales capturan con menor sensibilidad que el modelo de Milanesi, que incorpora explícitamente la volatilidad del activo.

La lectura de esta tabla exige una advertencia que la versión preliminar de este trabajo no formulaba. La detección de seis fracasos sobre seis no constituye por sí sola evidencia de poder predictivo. Un modelo que clasificara como fracaso a toda empresa que evalúa obtendría idéntico resultado en esta tabla. La capacidad discriminante solo puede establecerse contrastando este desempeño con el del grupo de control, y es allí donde el comportamiento de los modelos diverge de manera decisiva.

5.2 Grupo de control ampliado: poder discriminante

La tabla 7 presenta los veredictos de los siete modelos sobre las treinta empresas del grupo de control (que están ordenadas de menor a mayor índice de consenso).

La tabla 8 sintetiza el desempeño de cada modelo. La sensibilidad mide la proporción de fracasos correctamente identificados; la especificidad, la proporción de empresas solventes correctamente clasificadas. En ambos casos los veredictos de zona gris se computan con valor medio. La exactitud balanceada es el promedio de ambas medidas, y su valor de referencia es 50 %: ese es el resultado que obtendría una clasificación aleatoria o una regla constante.

El resultado principal de esta tabla es que la sensibilidad no informa prácticamente nada. Cinco de los siete modelos detectan la totalidad de los fracasos, pero tres de ellos lo hacen porque clasifican como fracaso a la mayoría de las empresas que evalúan. La jerarquía real surge de la especificidad.

Milanesi y Altman son los únicos dos modelos con desempeño utilizable. El modelo de opciones barrera no produce un solo falso positivo sobre treinta empresas solventes y detecta los seis fracasos, alcanzando una exactitud balanceada del 97,5 %. El modelo de Altman tampoco produce falsos positivos: sus errores son de zona gris, es decir, de indefinición y no de clasificación errónea. Ambos comparten una propiedad estructural que explica el resultado: sus formas funcionales no permiten que una única variable domine el puntaje. En Altman, la constante de 3,25 y el reparto de coeficientes impiden que el apalancamiento arrastre el resultado; en Milanesi, la probabilidad de default depende de la distancia entre el valor del activo y la barrera en unidades de volatilidad, magnitud que no se distorsiona con la intensidad de capital.

Pascale y Beaver fallan por razones distintas pero con consecuencias equivalentes. En el modelo de Pascale, la ecuación parte de una constante de -2,9511 y el único término con capacidad real de compensarla es Patrimonio neto sobre Pasivo total, con coeficiente 2,6271. Manteniendo los demás términos en valores típicos, alcanzar la zona de solvencia exige una relación Patrimonio sobre Pasivo superior a 1,22. Solo cuatro de las treinta empresas del grupo de control satisfacen esa condición. El modelo no está evaluando el riesgo de las empresas argentinas: está midiendo si su patrimonio supera a su deuda, y penalizando como fracaso a toda estructura de financiamiento con apalancamiento moderado. En el modelo de Beaver, el problema es la regla de agregación: cuatro de las cinco señales se activan simultáneamente en empresas de servicios regulados con capital de trabajo negativo y alto endeudamiento estructural, perfil que caracteriza a buena parte del grupo de control.

El modelo de Alonso mantiene una especificidad del 53 %, superior a la reportada en la versión preliminar de este trabajo, donde clasificaba como fracaso a la totalidad del grupo de control. La diferencia se explica por la composición de la muestra: las cinco empresas de aquel

Tabla 7: Resultado de cada modelo para las empresas del grupo de control

F = fracaso · G = zona gris · S = empresa sana. Todas las empresas de esta tabla cumplieron regularmente sus obligaciones durante el período analizado; en consecuencia, todo veredicto de fracaso constituye un falso positivo.

Empresa	*Beaver*	*Altman*	*D-G*	*Alonso*	*Milanesi*	*Pascale*	*S-P*	*Consenso*
Pampa Energía S.A.	S	S	S	S	S	S	F	15%
Ternium Argentina	S	S	S	S	S	S	F	15%
Aluar Aluminio Argentino	S	S	G	S	S	S	F	22%
Loma Negra S.A.	S	S	F	S	S	S	F	30%
Ledesma S.A.	G	S	G	S	S	G	F	32%
Molinos Río de la Plata	F	S	S	S	S	F	F	35%
Carboclor S.A.	F	S	G	S	S	F	F	43%
San Cornelio S.A.	F	S	G	S	S	F	F	43%
Disco S.A.	F	G	S	S	S	F	F	45%
Asoc. de Cooperativas Arg.	F	G	S	S	S	F	F	45%
IRSA Propiedades Comerciales	G	S	G	F	S	G	F	48%
Capsa Petróleo S.A.	F	S	F	S	S	F	F	50%
Electricidad de los Andes	F	S	F	S	S	F	F	50%
La Pampa Energía S.A.	F	S	F	S	S	F	F	50%
Cencosud S.A.	F	G	G	S	S	F	F	53%
Molinos Agro S.A.	F	G	G	S	S	F	F	53%
Mastellone Hnos. S.A.	F	G	G	S	S	F	F	53%
Autopistas del Sol S.A.	F	S	F	F	S	F	F	65%
Transener S.A.	F	S	F	F	S	F	F	65%
Metrogas S.A.	F	S	F	F	S	F	F	65%
Transp. de Gas del Norte	F	S	F	F	S	F	F	65%
Transp. de Gas del Sur	F	S	F	F	S	F	F	65%
EDENOR S.A.	F	S	F	F	S	F	F	65%
EDESUR S.A.	F	S	F	F	S	F	F	65%
AMX Argentina (Claro)	F	S	F	F	S	F	F	65%
YPF S.A.	F	G	S	F	G	F	F	68%
Cresud S.A.C.I.F. y A.	F	G	G	F	S	F	F	68%
Mercado Libre	G	S	F	F	G	F	F	68%
Telecom Argentina	F	G	F	F	S	F	F	75%
AySA (Agua y Saneamientos)	F	G	G	F	G	F	F	75%

Tabla 8: Desempeño de cada modelo

Ordenamiento por exactitud balanceada. El valor de 50% corresponde a la ausencia total de poder discriminante.

Modelo	*Sensibilidad*	*Especificidad*	*Exactitud balanceada*	*Falsos positivos*
Milanesi	100%	95%	97,5%	0 / 30
Altman Z'	92%	85%	88,3%	0 / 30
Alonso *et al*.	100%	53%	76,7%	14 / 30
Dutto-Gallacher	100%	37%	68,3%	14 / 30
Pascale	100%	17%	58,3%	24 / 30
Beaver	92%	18%	55,0%	23 / 30
Sandín-Porporato	100%	0%	50,0%	30 / 30

grupo eran todas de gran tamaño y capital intensivo, precisamente el perfil que penaliza el coeficiente sobre disponibilidades de 18,347. Al incorporar empresas de retail, agroindustria y comercialización, con mayor liquidez inmediata relativa, el modelo recupera parte de su capacidad. Sigue siendo un modelo calibrado para pymes y su aplicación a grandes empresas requiere recalibración, pero la afirmación de que las clasifica sistemáticamente como fracaso no se sostiene sobre la muestra ampliada.

El caso de Sandín-Porporato es de naturaleza distinta y se trata por separado en la sección 5.4.

5.3 Análisis crítico del modelo Dutto-Gallacher

El término dominante de la ecuación logística es -1,5311 multiplicado por el logaritmo de activo sobre ventas. Para una constructora pyme, población sobre la que el modelo fue calibrado, ese ratio se ubica típicamente entre 1,0 y 1,5, de modo que el logaritmo es próximo a cero y el impacto sobre el puntaje resulta neutro. Para una empresa de capital intensivo la situación es radicalmente distinta.

Sobre la muestra reducida original, el análisis revelaba una separación limpia con corte empírico en torno al 5 %: los fracasos se ubicaban por debajo del 3 % y las empresas solventes por encima del 8 %. La solución adoptada en la herramienta fue la detección automática del perfil de capital intensivo y la aplicación del corte ajustado, según se documenta en la sección 5.2.

La muestra ampliada obliga a matizar esa conclusión. Con el corte adaptativo en funcionamiento, el modelo conserva una especificidad del 37 %: catorce de las treinta empresas solventes siguen clasificadas como fracaso. La recalibración por umbrales mejora el desempeño pero no lo lleva a un nivel utilizable. La razón es que el corte adaptativo corrige el desplazamiento sistemático del puntaje, pero no restituye capacidad de separación: dentro del grupo de empresas de capital intensivo, la distribución de probabilidades estimadas de fracasos y solventes se superpone. El problema no es dónde se traza la línea, sino que las poblaciones no están separadas en esa dimensión.

Tabla 9: Análisis del modelo Dutto-Gallacher

Empresa	*Activo/Ventas*	*log(A/V)*	*Aporte al puntaje*	*P(sano)*
Constructora sana (paper)	1,08	0,08	−0,12	50,2 %
Loma Negra	1,95	0,67	−1,03	20,3 %
Telecom Argentina	1,93	0,66	−1,01	8,8 %
YPF S.A.	2,38	0,87	−1,33	9,5 %
Pampa Energía	4,26	1,45	−2,22	12,1 %
Generación Mediterránea	6,85	1,92	−2,95	2,5 %

Esta observación refuerza en lugar de invalidar el hallazgo original. La limitación de validez externa del modelo es real y su magnitud es mayor de lo que la muestra reducida permitía apreciar. Una recalibración adecuada exigiría reestimar los coeficientes sobre una muestra de grandes empresas cotizantes argentinas, no simplemente ajustar el punto de corte.

5.4 Anomalía de escala en el modelo Sandín-Porporato

El modelo de Sandín-Porporato clasifica como fracaso a la totalidad de los treinta y seis casos de la muestra: los seis fracasos y las treinta empresas solventes. Su especificidad es del 0% y su exactitud balanceada, del 50 %, equivalente a la de una regla constante. Un modelo publicado, calibrado sobre empresas argentinas y con coeficientes de signo teóricamente correcto, no debería comportarse de este modo. El protocolo de trazabilidad descripto en el apartado 3.2 permitió localizar la causa.

La descomposición del puntaje revela que la constante de la ecuación, de valor 6,507, es de un orden de magnitud superior al de la suma de los cuatro términos restantes cuando las variables se expresan como fracciones decimales. Para Pampa Energía, los cuatro términos suman -0,816, dejando un puntaje de 5,69 y una probabilidad estimada de fracaso del 99,7 %. Para Ternium Argentina, la suma es de -0,491 y la probabilidad resultante, del 99,8 %.

Tabla 10: Análisis del modelo Sandin-Porporato

Aportes de cada término al puntaje, sobre una constante de +6,507. La probabilidad de fracaso se satura en valores superiores al 99 % para toda la muestra, con independencia de la situación real de la empresa.

Empresa	*UN/AT*	*EBIT/AT*	*PT/PN*	*VT/AT*	*P*
Pampa Energía	−0,430	−0,248	+0,017	−0,155	99,7 %
Ternium Argentina	−0,160	−0,109	+0,004	−0,226	99,8 %
Aluar Aluminio	−0,637	−0,355	+0,018	−0,527	99,3 %
Celulosa Argentina	+2,471	+0,568	−0,368	−0,523	100 %

Para que el puntaje descienda por debajo de cero (condición necesaria para que la probabilidad estimada de fracaso caiga del 50 %) se requeriría, actuando únicamente sobre el término de rentabilidad, una utilidad neta equivalente al 97 % del activo total. Ninguna empresa en funcionamiento satisface esa condición. La implementación actual del modelo no puede emitir un veredicto de solvencia bajo ninguna configuración de datos económicamente posible.

La hipótesis más plausible es que los coeficientes originales fueron estimados sobre variables expresadas en puntos porcentuales y no en fracciones decimales. Al reescalar las cuatro variables por un factor de cien, el comportamiento del modelo se transforma: clasifica correctamente a las treinta empresas del grupo de control y detecta cuatro de los seis fracasos, con una exactitud balanceada que asciende del 50 % al 83 %. La mejora es de tal magnitud que resulta difícil atribuirla al azar. No obstante, bajo esa escala las probabilidades estimadas se saturan en los extremos (cero o cien por ciento en casi todos los casos), lo que sugiere que la escala correcta podría ser intermedia o que la especificación requiere un tratamiento adicional que la fuente disponible no permite determinar.

Se deja constancia de que el modelo se reporta en este trabajo con su implementación original, en fracciones decimales, y no con el reescalado propuesto. La razón es metodológica: modificar la escala de las variables sobre la base de que el resultado mejora constituiría un ajuste guiado por el desempeño en la muestra, exactamente el procedimiento que este trabajo cuestiona en el punto 5 del apartado 3.2. La resolución del punto exige el acceso al detalle de la especificación de variables en la publicación original y queda planteada como línea de trabajo pendiente.

El episodio tiene valor metodológico más allá del modelo puntual. La anomalía estuvo presente desde la primera implementación y permaneció inadvertida durante todo el desarrollo. No la detectó la verificación sintáctica del código, ni la revisión de los coeficientes contra el paper −que son correctos−, ni la aplicación a la muestra reducida de cinco empresas solventes, donde el modelo clasificaba a todas como fracaso y ese resultado se atribuyó a las limitaciones de validez externa ya conocidas de los otros modelos. Solo la ampliación del grupo de control hizo visible que un veredicto uniforme sobre treinta empresas heterogéneas no puede ser un problema de calibración. La lección es que un modelo implementado con coeficientes correctos puede estar mal cargado, y que la única verificación que detecta esta clase de error es la aplicación a una muestra suficientemente amplia y diversa.

5.5 Control de inaplicabilidad: entidades financieras

La muestra incorpora un grupo separado de seis entidades financieras cotizantes con obligaciones negociables vigentes, todas sin antecedentes de incumplimiento en el período. Este grupo no integra el cómputo del poder discriminante y cumple una función distinta: verificar qué sucede cuando modelos calibrados sobre una población se aplican a otra cuya estructura patrimonial responde a una lógica económica diferente.

El resultado es inequívoco. Sobre seis entidades y siete modelos (cuarenta y dos evaluaciones) no se registra un solo veredicto de empresa sana. Cinco de los siete modelos clasifican como fracaso a la totalidad de las entidades; el modelo de Altman las ubica sin excepción en zona gris. La tasa de falsos positivos del conjunto es del 100 %.

Tabla 10: Inaplicabilidad a entidades financieras

Todas las entidades cumplieron regularmente sus obligaciones durante el período analizado. Grupo Supervielle S.A. se omite por corresponder a la misma entidad económica que Banco Supervielle S.A.

Entidad	*Pasivo/PN*	*Activo/Ventas*	*Consenso*	*Veredictos*
Grupo Financiero Galicia	6,0	6,2	**85 %**	5 fracaso · 2 gris
Banco Macro	5,9	7,4	**85 %**	5 fracaso · 2 gris
Banco Patagonia	5,9	7,8	**85 %**	5 fracaso · 2 gris
Banco Hipotecario	6,3	6,6	**90 %**	6 fracaso · 1 gris
Banco Ciudad de Buenos Aires	6,5	6,5	**90 %**	6 fracaso · 1 gris
Banco Supervielle	6,0	7,2	**90 %**	6 fracaso · 1 gris
Referencia: Pampa Energía	0,7	4,3	**15%**	6 sana · 1 fracaso

1) Mecanismo del error

El resultado no obedece a deficiencias en los datos ni a un defecto de implementación. Obedece a que las variables de los siete modelos presuponen una estructura patrimonial que una entidad financiera no tiene ni puede tener. Un banco es un intermediario: su pasivo, fundamentalmente depósitos, es el insumo del negocio y no una carga sobre la capacidad de pago. Una relación Pasivo sobre Patrimonio neto del orden de seis a uno corresponde a un capital regulatorio en torno al 13 % o 14 % del activo, plenamente consistente con las exigencias de integración de capital mínimo del Banco Central. En una empresa no financiera, el mismo valor señala insolvencia inminente.

Pascale. La ecuación parte de una constante de -2,9511 y el único término con capacidad de compensarla es Patrimonio neto sobre Pasivo total. Para Grupo Financiero Galicia ese término aporta +0,438; para Pampa Energía aporta +3,878. Los tres términos restantes suman en conjunto 0,316 en el caso del banco. El puntaje final es de -2,197, dentro de la zona de fracaso, y ningún nivel de rentabilidad podría revertirlo: una entidad financiera no puede alcanzar una relación Patrimonio sobre Pasivo compatible con la zona de solvencia del modelo sin abandonar su actividad.

Dutto-Gallacher. Su coeficiente dominante penaliza el logaritmo de Activo sobre Ventas, y en una entidad financiera las ventas corresponden a ingresos financieros netos y no a facturación, de modo que la relación se ubica entre 6 y 8 frente a valores de 2 a 4 en empresas industriales. Para Galicia ese término aporta −2,788 y la probabilidad estimada de solvencia resulta de 1,46%, por debajo incluso del umbral corregido para empresas de capital intensivo. Es el fenómeno de la sección 5.3 llevado a su expresión extrema.

Alonso et al. El modelo pierde directamente uno de sus tres términos: la variable Créditos por ventas sobre Deuda total es nula por construcción en una entidad financiera, que no genera créditos comerciales. El término de mayor peso, Disponibilidades sobre Activo, tampoco captura la posición de liquidez de un banco, que se estructura en integración de efectivo mínimo, tenencia de títulos públicos y operaciones de pase.

Beaver y Altman. Ambos dependen de la apertura entre activo y pasivo corriente. Un estado contable bancario confeccionado bajo el marco normativo del Banco Central no presenta esa apertura: presenta una escalera de vencimientos. Los valores de activo y pasivo corriente utilizados para las entidades del grupo son estimaciones construidas a partir de esa escalera y no dato de balance, circunstancia que se declara expresamente. Es significativo que el modelo de Altman sea el menos distorsionado del conjunto: su variable de apalancamiento pesa apenas 1,05 sobre un puntaje que la constante de 3,25 sostiene, lo que impide que la estructura de intermediación domine el resultado. Esa relativa robustez es una propiedad de la forma funcional, no del ajuste del modelo a la población, y resulta consistente con su desempeño superior en el grupo de control del apartado 5.2.

2) Implicancias

El control de inaplicabilidad permite formular la limitación de validez externa en términos más precisos que los habituales en la literatura. El problema no se reduce a que un modelo pierda poder predictivo al alejarse de su población de calibración, que es la formulación estándar. El problema es que existen poblaciones en las que las variables del modelo no conservan el significado económico que el modelo les atribuye, y en esas poblaciones el resultado no es una predicción degradada sino una predicción sin contenido informativo.

La distinción tiene consecuencias prácticas. Una predicción degradada puede corregirse recalibrando los puntos de corte, que fue el intento realizado con el modelo Dutto-Gallacher. Una predicción sin contenido no admite corrección por umbrales: no existe un punto de corte que separe bancos solventes de bancos insolventes en un modelo cuya variable de apalancamiento asigna el mismo valor a la operación normal de intermediación que a la insolvencia industrial. El análisis de entidades financieras exige modelos construidos sobre variables regulatorias específicas (adecuación de capital, calidad de cartera, descalce de plazos y monedas, cobertura de liquidez), ajenos a la familia de modelos analizada en este trabajo.

De este resultado se desprende también una recomendación operativa para el uso de herramientas de scoring automatizadas. La herramienta desarrollada, como cualquier implementación de esta familia de modelos, produce un veredicto para todo conjunto de datos que reciba, sin advertir que la población de aplicación es inadecuada. La verificación de aplicabilidad es responsabilidad del analista y debe preceder a la lectura del resultado. Esta consideración adquiere particular relevancia cuando el desarrollo se apoya en asistencia de inteligencia artificial: el asistente implementa correctamente lo que se le solicita y no interpela el supuesto de aplicabilidad, que permanece enteramente a cargo del criterio profesional de quien lo utiliza.

5.6 Desempeño del índice de consenso ponderado

El resultado más llamativo del análisis es que el índice de consenso separa ambos grupos sin superposición alguna, pese a que cinco de los siete modelos que lo componen exhiben individualmente un desempeño deficiente.

Tabla 11: Indice de consenso ponderado

Grupo	Índice mínimo	Índice máximo	Casos
Fracasos	85 %	100 %	6
Grupo de control	15 %	75 %	30
Superposición entre grupos	—	—	**0**

Ninguna de las treinta empresas solventes alcanza el índice mínimo del grupo de fracasos. La brecha entre el máximo del grupo de control (75%, correspondiente a Telecom Argentina y AySA) y el mínimo del grupo de fracasos (85%, correspondiente a Generación Mediterránea) es de diez puntos porcentuales.

Este resultado exige cautela interpretativa por tres razones. En primer lugar, la separación perfecta sobre una muestra de treinta y seis casos, seis de los cuales son fracasos, no permite estimar tasas de error con precisión útil: el intervalo de confianza en torno a una tasa de falsos negativos de cero sobre seis observaciones es amplio. En segundo lugar, el buen desempeño del agregado se apoya en dos modelos (Milanesi y Altman) que concentran el 35 % de la ponderación y funcionan bien individualmente; el aporte de los cinco restantes es en gran medida un desplazamiento uniforme del índice que no altera el ordenamiento. En tercer lugar, y de manera más significativa, el modelo de Sandín-Porporato aporta un 15 % de peso emitiendo un veredicto constante, lo que equivale a sumar quince puntos porcentuales a todos los casos por igual; su contribución al poder discriminante del índice es nula por construcción, aun cuando no lo perjudique.

Corregida la anomalía de escala documentada en el apartado 5.4, o eliminado ese modelo del agregado, la separación entre grupos se ampliaría. El desempeño reportado debe entenderse, en consecuencia, como un piso y no como el resultado de una configuración óptima.

6. Discusión: límites y contribuciones

6.1. Limitaciones del trabajo

Tamaño del grupo de fracasos. La muestra contiene seis casos de fracaso. Aunque el grupo de control se amplió a treinta empresas, el número de eventos sigue siendo insuficiente para test estadísticos de significación o estimación precisa de tasas de error. Las medidas de sensibilidad reportadas en el apartado 5.2 tienen intervalos de confianza amplios y deben leerse como indicativas.

Desbalance de la muestra. La relación de cinco a uno entre control y fracasos se aparta del diseño balanceado de los estudios de calibración originales. La decisión fue deliberada, según se argumenta en el apartado 4.1, porque el objetivo del trabajo es evaluar el desempeño de los modelos en condiciones de uso reales, donde los fracasos son minoritarios. Como contrapartida, las medidas de sensibilidad y especificidad no son directamente comparables con las reportadas en los papers originales.

Homogeneidad temporal. Todos los casos pertenecen al período 2024-2025, caracterizado por alta inflación, un proceso de normalización macroeconómica en curso y un ciclo de crédito específico. La generalización a otros períodos no está garantizada.

Calidad de los datos. Una porción de los valores contables de la base son estimaciones, particularmente en empresas que discontinuaron la publicación de estados contables completos tras el evento de fracaso, y en los rubros de activo y pasivo corriente de las entidades financieras. El nivel de confianza de cada dato está documentado en la herramienta. Los parámetros de valuación del modelo de Milanesi son estimaciones sectoriales y no mediciones.

Desfase temporal. Los modelos operan sobre datos del último ejercicio cerrado antes del fracaso, que puede anteceder al evento entre seis y dieciocho meses. El desempeño reportado no equivale a la capacidad de anticipación con horizonte fijo.

Especificación pendiente del modelo Sandín-Porporato. La anomalía de escala documentada en el apartado 5.4 no pudo resolverse con las fuentes disponibles. El modelo se reporta con su implementación original, lo que probablemente subestima su desempeño real.

6.2. Contribuciones

Contribución metodológica. Se documenta un protocolo replicable de carga, implementación y verificación de modelos académicos con asistencia de IA generativa, con un criterio explícito de trazabilidad que sitúa la fuente de verdad en el artefacto de software y no en el registro conversacional. En el apartado 5.4 se documenta un caso concreto en que ese protocolo permitió detectar un error de transcripción que había sobrevivido a la verificación de coeficientes.

Contribución empírica. Se aplican siete modelos académicos simultáneamente a una muestra contemporánea del mercado argentino con un grupo de control de treinta empresas, lo que permite estimar tasas de falsos positivos que las muestras balanceadas de tamaño reducido no revelan. El resultado central (que cinco de los siete modelos exhiben especificidad inferior al 55 % sobre grandes empresas cotizantes argentinas) no había sido documentado con esta amplitud para el mercado local.

Contribución analítica. Se identifican formalmente tres limitaciones distintas: la de validez externa del modelo Dutto-Gallacher en empresas de capital intensivo, con la precisión de que la recalibración por umbrales resulta insuficiente; el requisito implícito de solvencia patrimonial

que el modelo de Pascale impone y que la mayoría de las empresas cotizantes no satisface; y la anomalía de escala del modelo Sandín-Porporato.

Contribución conceptual. Se propone distinguir entre predicción degradada y predicción sin contenido informativo, sobre la base del control de inaplicabilidad con entidades financieras. La primera admite corrección por recalibración; la segunda exige un cambio de familia de modelos.

Contribución pedagógica. La herramienta resultante es directamente utilizable en cursos de Administración Financiera, Análisis Crediticio y Mercado de Capitales, y la documentación del proceso de verificación tiene valor didáctico independientemente del resultado de los modelos.

6.3 El rol de la IA en el análisis financiero académico

La experiencia documentada permite extraer cuatro conclusiones sobre el uso de IA generativa en el análisis financiero académico.

La IA como acelerador de implementación. La traducción de fórmulas matemáticas a código ejecutable y la generación de interfaces de usuario son tareas en las que el asistente resulta altamente eficiente, con una reducción sustancial del tiempo de prototipado.

La IA como herramienta de diagnóstico. Identificado un resultado anómalo, el asistente ejecuta en minutos descomposiciones algebraicas y análisis de sensibilidad que manualmente demandarían horas. Los dos hallazgos analíticos de este trabajo se obtuvieron por esa vía.

La corrección sintáctica no es corrección sustantiva. El episodio del modelo Sandín-Porporato es la lección central del trabajo en esta materia. El código era sintácticamente correcto, los coeficientes coincidían con la publicación y el modelo devolvía resultados plausibles caso por caso. El error solo se hizo visible al aplicar el modelo a una muestra amplia y observar que su respuesta era constante. Un asistente de IA no advierte que un modelo carece de poder discriminante, porque esa determinación exige comparar su comportamiento contra una expectativa teórica que el asistente no formula por iniciativa propia. La verificación sustantiva es responsabilidad del investigador.

La IA no reemplaza el juicio analítico. En los dos hallazgos del trabajo, el reconocimiento del resultado como anómalo provino del autor, a partir de su incompatibilidad con expectativas teóricas fundadas en conocimiento del dominio. El asistente proveyó el diagnóstico matemático. La división del trabajo es estable y conviene explicitarla: la IA responde con eficiencia a la pregunta de por qué un resultado es el que es; formular la pregunta correcta sigue dependiendo del investigador.

7. Conclusiones

En este ensayo se ha demostrado la viabilidad de integrar herramientas de inteligencia artificial generativa en el desarrollo de aplicaciones analíticas financieras de nivel académico y profesional, y ha documentado tanto sus ventajas como sus riesgos específicos. El dashboard *ONcalc y Fracasos corporativos Argentina 2024-2025* es el resultado de un proceso de cocreación entre el conocimiento del dominio del autor y las capacidades de implementación y diagnóstico del modelo Claude (Anthropic).

La aplicación de los siete modelos a una muestra de treinta y seis casos, con un grupo de control de treinta empresas, permite formular cinco conclusiones.

Primera conclusión: la sensibilidad, aisladamente considerada, no constituye evidencia de poder predictivo. Cinco de los siete modelos detectaron la totalidad de los fracasos de la muestra. Tres de ellos lo hicieron clasificando también como fracaso a la mayoría de las empresas solventes. La evaluación de un modelo de scoring exige un grupo de control de tamaño suficiente, y las muestras balanceadas de escala reducida (habituales en la literatura de calibración) tienden a producir una imagen sistemáticamente optimista de su desempeño.

Segunda conclusión: solo dos de los siete modelos resultan utilizables individualmente sobre grandes empresas cotizantes argentinas. El modelo de opciones barrera de Milanesi alcanza una exactitud balanceada del 97,5 % sin producir falsos positivos, y el modelo Z' de Altman llega al 88,3 % con errores que se concentran en la zona gris. Ambos comparten una propiedad estructural: su forma funcional impide que una variable individual domine el puntaje. Los cinco restantes exhiben especificidades de entre el 0 % y el 53 %.

Tercera conclusión: la validez externa es sensible a la distancia entre la población de calibración y la de aplicación, y esa sensibilidad no siempre se corrige por recalibración de umbrales. El modelo Dutto-Gallacher conserva una especificidad del 37 % aun con el corte adaptativo en funcionamiento, porque el problema no es la ubicación del punto de corte sino la superposición de las distribuciones. El modelo de Pascale impone un requisito implícito de patrimonio superior al pasivo que solo cuatro de treinta empresas satisfacen. Una recalibración adecuada exigiría reestimar coeficientes, no ajustar umbrales.

Cuarta conclusión: existen poblaciones en las que estos modelos no producen predicciones degradadas sino predicciones sin contenido informativo. El control con entidades financieras arroja cuarenta y dos evaluaciones sin un solo veredicto de solvencia. En esas poblaciones las variables no conservan el significado económico que el modelo les atribuye, y ninguna elección de umbral restituye capacidad de separación. La verificación de aplicabilidad debe preceder a la lectura del resultado y es responsabilidad indelegable del analista.

Quinta conclusión: el desarrollo asistido por IA es eficiente, pero desplaza el riesgo desde el error de implementación hacia el error de verificación. El prototipo funcional completo se desarrolló en el equivalente a dos jornadas de trabajo, un plazo inalcanzable por medios convencionales. El costo de esa velocidad es que los errores que sobreviven no son sintácticos sino

sustantivos, y por lo tanto invisibles para las verificaciones habituales. El caso del modelo Sandín-Porporato, cuya ausencia total de poder discriminante permaneció inadvertida hasta que la ampliación del grupo de control la hizo evidente, ilustra la naturaleza del riesgo. El protocolo de trazabilidad propuesto en el apartado 3.2 es la respuesta metodológica que este trabajo ofrece.

El trabajo abre varias líneas de investigación futura. En primer término, la resolución de la anomalía de escala del modelo Sandín-Porporato mediante acceso al detalle de la especificación original, y la reevaluación de su desempeño con la escala corregida. En segundo término, la recalibración de los coeficientes del modelo Dutto-Gallacher sobre una muestra de grandes empresas cotizantes argentinas. En tercer término, la construcción de un modelo logístico que incorpore variables de estructura de deuda en moneda extranjera, dimensión que la experiencia de Generación Mediterránea sugiere determinante y que ningún modelo contable de los analizados captura adecuadamente.

Finalmente, la exploración de las capacidades diagnósticas de los modelos de lenguaje para la detección temprana de señales de deterioro financiero en información no estructurada (comunicados de hechos relevantes, notas a los estados contables, informes de auditoría con salvedades), terreno en el que su ventaja comparativa frente a los modelos cuantitativos podría resultar sustantiva.

REFERENCIAS

Altman, E. I. (1968). *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. Journal of Finance, 23 (4): 589-609

Altman, E. I. & Hotchkiss, E. (2006). *Corporate financial distress and bankruptcy*, 3rd ed. Wiley

Alonso, A. H., Barraza, S., Guglielmetti, G. y Legato, A. (2006). *Riesgo crediticio en las pyme: Desarrollo de un modelo de predicción de default*. XXVI Jornadas de Administración Financiera, SADAF, págs. 1-12

Beaver, W. H. (1966). *Financial ratios as predictors of failure*. Journal of Accounting Research, 4: 71-111.

Black, F. & Scholes, M. (1973). *The pricing of options and corporate liabilities*. Journal of Political Economy, 81 (3): 637-659

Brockman, P. & Turtle, H. (2003). *A barrier option framework for corporate security valuation*. Journal of Financial Economics, 67 (3): 511-529

Dutto, M. L. (2008). *La predicción de dificultades financieras: un análisis univariado aplicado a la industria de la construcción*. XXVIII Jornadas de Administración Financiera, SADAF, págs. 97-106

Dutto, M. y Gallacher, M. (2009). *Un modelo de pronóstico de insolvencia financiera de empresas*. XXIX Jornadas de Administración Financiera, SADAF, págs. 88-98

Gnecco, M. L. *et al*. (2020). *Aplicación de modelos de pronóstico de fracaso empresario a empresas con actividad en la ciudad de Mar del Plata*. 40 Jornadas de Administración Financiera, SADAF, págs. 114-137

Merton, R. C. (1974). *On the pricing of corporate debt: the risk structure of interest rates*. Journal of Finance, 29 (2): 449-470

Milanesi, G. S., Vigier, H. P. y Gzain, M. (2015). *Un modelo de predicción de fracasos financieros con opciones barrera*. XXXV Jornadas de Administración Financiera, SADAF, págs. 177-189

Milanesi, G. S. (2019). *Un modelo naíf de opciones knock out para la predicción de fracasos financieros*. 39 Jornadas de Administración Financiera, SADAF, págs. 139-159

Ohlson, J. A. (1980). *Financial ratios and the probabilistic prediction of bankruptcy*. Journal of Accounting Research, 18 (1): 109-131.

Pascale, R. (1988). *Discriminant analysis as a complement to ratio analysis in predicting financial failure*. Latin American Journal of Economics, 14 (2): 87-108

Sandín, M. & Porporato, M. (2007). *Corporate bankruptcy prediction models applied to emerging economies: evidence from Argentina in the 2000s*. International Journal of Commerce and Management, 17 (4): 295-311

Saunders, A. & Millon Cornett, M. (2003). *Financial institutions management: A risk management approach*. McGraw-Hill/Irwin.