



www.aileaders-project.eu

SIMULACIÓN - Retos éticos

en el análisis financiero



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

SIMULACIÓN

- Retos éticos en el análisis financiero

01



Resumen 3

02



Introducción 4

03



Presentación de herramientas 5

04



Ejecución de la simulación 6

05



Conclusiones 7

06



Referencias 8

• 01 Resumen



Tipo de REA

Demostración/simulación con Google Colab (clasificador RandomForest frente a FairGBM).

Objetivo/Finalidad

Proporcionar al alumnado una exploración práctica y crítica de cómo la toma de decisiones algorítmica en el análisis y la previsión financieros puede reproducir desigualdades estructurales mediante la comparación de los resultados de las métricas de equidad. Esta simulación fomenta la reflexión sobre las dimensiones éticas de la automatización financiera y promueve el desarrollo de modelos predictivos más justos e inclusivos.

Resultados de aprendizaje esperados

Al final de la simulación, el alumnado será capaz de:

- 01** Detectar e interpretar el sesgo algorítmico en la aprobación de créditos.
- 02** Reflexionar sobre las implicaciones éticas de las decisiones financieras automatizadas.
- 03** Utilizar métricas de equidad para evaluar los resultados del modelo.

Palabras clave

- Aprendizaje automático
- Clasificación
- Sesgos
- Equidad

Enfoque metodológico sugerido

Aprendizaje basado en problemas

NOTA

Se requieren conocimientos intermedios de **programación en Python** para comprender y trabajar con el contenido de este REA.

• 02 Introducción



Las herramientas de previsión financiera basadas en la inteligencia artificial se están adoptando cada vez más para automatizar las decisiones relacionadas con el crédito, como la aprobación de préstamos y la evaluación de riesgos (véase, por ejemplo, Chen, 2020; Dastile et al., 2020; Chen et al., 2024; Heß y Damásio, 2025). Estas tecnologías ofrecen considerables ventajas en términos de eficiencia, coherencia y escalabilidad. Sin embargo, cuando se implementan en contextos históricamente marcados por desigualdades estructurales, plantean importantes cuestiones éticas.

Una cuestión central es la posibilidad de que estos sistemas reproduzcan, o incluso amplifiquen, los sesgos existentes en los datos financieros. Los algoritmos que parecen neutrales pueden, cuando se entrenan con conjuntos de datos sesgados o incompletos, producir resultados injustos que afectan de manera desproporcionada a grupos ya marginados o vulnerables.

Un ejemplo particularmente convincente es el de las mujeres (Orser et al., 2006; Ongena y Popov, 2016; Moro et al., 2017; Beck et al., 2018; De Andrés et al., 2021, entre otros) y las personas negras (Chatterji y Seamans, 2012), que siguen enfrentándose a barreras agravadas para acceder al crédito de las instituciones financieras tradicionales debido a formas de discriminación que se entrecruzan.

Este ejercicio basado en la simulación invita a los estudiantes a explorar de forma crítica la intersección entre el aprendizaje automático, la previsión financiera y la equidad. Utilizando un conjunto de datos sintéticos que refleja solicitudes de crédito realistas, el alumnado investigará cómo atributos como el género, la ocupación, la educación y el estado civil influyen en los resultados de la aprobación de los préstamos. El conjunto de datos se ha construido deliberadamente para codificar patrones sutiles de sesgo, lo que ofrece una oportunidad práctica para detectar, analizar y mitigar las disparidades mediante la experimentación aplicada.

Basado en la literatura más amplia sobre inclusión financiera y brechas de financiación, este ejercicio desafía al alumnado a reflexionar sobre cómo los sistemas de IA pueden reforzar o confrontar la injusticia sistémica. Al hacerlo, no solo adquirirán competencias técnicas en auditoría de equidad y mitigación de sesgos, sino que también desarrollarán una comprensión crítica de las responsabilidades éticas asociadas con el diseño y la implementación de la IA en los ecosistemas financieros.

• 03 Presentación de herramientas



El **escenario principal** que se presenta en esta simulación se centra en un conjunto de datos sintético de solicitudes de préstamos.

Los datos incluyen características demográficas y socioeconómicas tales como:

- Género
- Raza
- Edad
- Ocupación
- Nivel educativo
- Estado civil

La variable objetivo es la **aprobación del préstamo**, modelada como un problema de clasificación binaria.

Aunque el conjunto de datos se ha construido para reflejar condiciones realistas, incorpora intencionadamente patrones ocultos de sesgo. Por ejemplo, las mujeres empresarias rurales pueden enfrentarse a tasas de rechazo más altas que sus homólogos masculinos o urbanos, incluso cuando se controlan las cualificaciones y los factores de riesgo. Esto simula cómo las barreras históricas y sistémicas pueden replicarse en los sistemas automatizados.

El objetivo es investigar dos cuestiones principales

- 01 ¿Los algoritmos de aprendizaje automático «recomiendan» de manera desproporcionada el rechazo de las solicitudes de crédito presentadas por personas negras y mujeres en los modelos predictivos binarios?
- 02 ¿Qué factores estructurales o algorítmicos pueden contribuir a esta disparidad?

• 04 Ejecución de la simulación



01 Acceda al cuaderno de simulación

Vaya a <https://tinyurl.com/2txdk9s>

02 Ejecutar el código

Ejecute todas las celdas del cuaderno de forma secuencial. Asegúrese de que no se produzcan errores y de que todos los resultados se muestren correctamente. Para ello, haga clic en el botón de reproducción situado en la parte superior izquierda de cada celda.

03 Explorar el conjunto de datos¹

Revise la estructura y el contenido del conjunto de datos. Preste especial atención a las variables demográficas clave, como el género, la ocupación, la educación y el resultado del préstamo. Observe cualquier posible desequilibrio o patrón que pueda indicar un sesgo.

04 Analizar la equidad

Examine los resultados. Compare las tasas de aprobación entre los diferentes grupos demográficos, centrándose especialmente en el género y la raza. Utilice las métricas de equidad proporcionadas en el cuaderno (por ejemplo, paridad demográfica, impacto desigual) para cuantificar el posible sesgo.

05 Compare los resultados

Compare cuidadosamente los resultados del modelo antes y después de aplicar las intervenciones de equidad, si se incluyen.

06 Reflexione sobre los resultados

Basándose en los resultados observados, reflexione sobre las implicaciones éticas y prácticas del uso de la IA en la toma de decisiones financieras. Considere las ventajas e inconvenientes entre precisión, eficiencia y equidad.

1. Tutorial de la herramienta «What If» (¿Qué pasaría si...?) - <https://www.youtube.com/watch?v=jHojeFCC5HE>

• 05 Conclusión



Los datos en los que nos basamos para esta simulación se utilizan ampliamente para estudiar la equidad y el sesgo en el aprendizaje automático, ya que incluyen atributos sensibles como el género y la raza. Se utilizan habitualmente para predecir si los ingresos de una persona superan los 50 000 dólares al año basándose en características como la edad, la educación, la ocupación y el estado civil.



Nuestra simulación demuestra cómo los sistemas de previsión financiera basados en la inteligencia artificial, si no se examinan, pueden perpetuar los sesgos estructurales incorporados en los datos históricos.

Revela que la precisión predictiva del rechazo de créditos en modelos binarios es mayor para las personas negras y las mujeres que para las personas blancas y los hombres.

Resumen de resultados



Precisión en la predicción de las decisiones de rechazo de crédito, es decir, «negativos» verdaderos (%)

Negr

97
MUJER

91 %
HOMBRE

Blanc

94
MUJER

86
HOMBRE

Incluso en modelos simplificados como el que nos ocupa, la presencia de resultados sesgados — como el menor porcentaje de falsos negativos en las tasas de rechazo previstas para las personas negras y las mujeres— pone de relieve la responsabilidad ética de los desarrolladores y analistas de auditar y perfeccionar los sistemas de IA. Estos resultados reavivan el debate sobre los puntos ciegos de los modelos de puntuación crediticia (véase, por ejemplo, Robb y Robinson, 2018).

Al trabajar directamente con métricas de equidad y estrategias de mitigación, los estudiantes no solo adquieren herramientas técnicas para identificar sesgos, sino que también desarrollan una comprensión más profunda de cómo se manifiestan dichos sesgos y cómo se pueden abordar. Este ejercicio refuerza el principio de que la equidad en la IA financiera no es una cuestión secundaria, sino un componente fundamental de la innovación responsable y sostenible.

• 06 Referencias



- Beck, T., Behr, P., & Madestam, A. (2011). Sex and credit: Is there a gender bias in lending?. European Banking Center Discussion Paper, (2011-027).
- Chatterji, A. K., & Seamans, R. C. (2012). Entrepreneurial finance, credit cards, and race. *Journal of Financial Economics*, 106(1), 182-195.
- Chen, T. H. (2020). Do you know your customer? Bank risk assessment based on machine learning. *Applied Soft Computing*, 86, 105779.
- Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, 312(1), 357-372.
- Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, 106263.
- De Andrés, P., Gimeno, R., & de Cabo, R. M. (2021). The gender gap in bank credit access. *Journal of Corporate Finance*, 71, 101782.
- Heß, V. L., & Damásio, B. (2025). Machine learning in banking risk management: Mapping a decade of evolution. *International Journal of Information Management Data Insights*, 5(1), 100324.
- Moro, A., Wisniewski, T. P., & Mantovani, G. M. (2017). Does a manager's gender matter when accessing credit? Evidence from European data. *Journal of banking & finance*, 80, 119-134.
- Ongena, S., & Popov, A. (2016). Gender bias and credit access. *Journal of Money, Credit and Banking*, 48(8), 1691-1724.
- Orser, B. J., Riding, A. L., & Manley, K. (2006). Women entrepreneurs and financial capital. *Entrepreneurship Theory and practice*, 30(5), 643-665.
- Robb, A., & Robinson, D. T. (2018). Testing for racial bias in business credit scores. *Small Business Economics*, 50, 429-443.





Sigue nuestro viaje



www.aileaders-project.eu



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.