

Tipo de artículo: Artículo original

## **Modelo basado en machine learning predictivo-prescriptivo para la mejora de la gestión del talento humano**

Predictive-prescriptive machine learning-based model for improving human talent management

Anamey Mendoza Mera<sup>1\*</sup>, <http://orcid.org/0000-0002-7709-8427>

<sup>1</sup> Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López. Ecuador.

\*Autor para la correspondencia. [amendoza@espam.edu.ec](mailto:amendoza@espam.edu.ec)

---

### **RESUMEN**

Esta investigación propone un modelo híbrido de *machine learning* predictivo-prescriptivo para optimizar la gestión del talento humano, integrando la anticipación de riesgos (rotación, brechas de habilidades) con recomendaciones de acciones estratégicas. Utilizando XGBoost para predecir la probabilidad de rotación y un algoritmo de optimización lineal entera para asignar recursos bajo restricciones presupuestarias y éticas, el modelo demuestra en simulaciones una reducción significativa de la rotación y costos operativos, incluso en escenarios de crisis económica. Los resultados destacaron la importancia de la equidad algorítmica, logrando paridad en la asignación de intervenciones (bonos, mentorías) entre grupos demográficos mediante métricas como la diferencia de paridad demográfica y herramientas de auditoría. Aunque el estudio se basa en datos sintéticos, su enfoque en la transparencia y adaptabilidad ofrece un marco replicable para organizaciones con características similares. Las limitaciones identificadas incluyen la generalización a contextos reales heterogéneos y la necesidad de incorporar variables no binarias en futuras iteraciones. El



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

trabajo subraya el potencial de la IA ética para transformar prácticas de RRHH, equilibrando eficiencia, sostenibilidad e inclusión.

**Palabras clave:** Modelo predictivo-prescriptivo; gestión del talento humano; retención de empleados; optimización de recursos; equidad algorítmica.

## **ABSTRACT**

This research proposes a hybrid predictive-prescriptive machine learning model to optimise human talent management by integrating risk anticipation (turnover, skill gaps) with strategic action recommendations. Using XGBoost to predict the probability of turnover and an integer linear optimisation algorithm to allocate resources under budgetary and ethical constraints, the model demonstrated in simulations a significant reduction in turnover and operating costs, even in scenarios of economic crisis. The results highlighted the importance of algorithmic fairness, achieving parity in the allocation of interventions across demographic groups using metrics such as demographic parity difference and audit tools. Although the study was based on synthetic data, its focus on transparency and adaptability offers a replicable framework for organisations. Limitations include generalisation to heterogeneous real-world contexts and the need to incorporate non-binary variables in future iterations. The paper highlights the potential of ethical AI to transform HR practices, balancing efficiency, sustainability and inclusion.

**Keywords:** Predictive-prescriptive model; human talent management; employee retention; resource optimisation; algorithmic equity.

**Recibido:** 06/01/2025

**Aceptado:** 07/03/2025

**En línea:** 01/04/2025

---

## **Introducción**

La gestión del talento humano (GTH) enfrenta desafíos sin precedentes en la era digital. Según el *World Economic Forum* (2023), el 40% de las competencias laborales pueden cambiar en el 2025 debido a la



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

automatización, generando brechas críticas en sectores como tecnología y salud. En este escenario, las organizaciones requieren enfoques basados en datos para anticipar riesgos (ej.: rotación del personal) y optimizar decisiones estratégicas (ej.: desarrollo de liderazgo). Sin embargo, métodos tradicionales de GTH suelen ser insuficientes, limitándose a corregir problemas una vez ocurridos (Davenport et al., 2010).

La convergencia de la analítica predictiva (anticipar eventos futuros) y la analítica prescriptiva (recomendar acciones óptimas) ofrecen soluciones transformadoras (Wissuchek et al., 2024). Los modelos híbridos permiten no solo identificar empleados en riesgo de renuncia, sino también asignar recursos de manera eficiente para retenerlos, considerando restricciones presupuestarias y éticas (Tambe et al., 2019). Este enfoque es especialmente relevante en industrias con alta rotación, como ventas minoristas y tecnología, donde reemplazar un empleado cuesta entre el 50% y 200% de su salario anual (Cappelli, 2021).

Los modelos predictivos han dominado la investigación en GTH durante la última década (Bujold et al., 2024). Fitz-enz (2010) pionero en aplicar regresión logística para predecir rotación, identificó variables como satisfacción laboral y equilibrio trabajo-vida como determinantes críticos. Posteriormente, estudios con machine learning (ML) han avanzado hacia la mejora de la precisión.

Chien y Chen (2008) lograron un área bajo la curva ROC (AUC-ROC por sus siglas en inglés) de 0.88 al predecir la rotación en la industria tecnológica usando datos de desempeño y encuestas del clima organizacional. Por otra parte, un trabajo de Huang et al. (2019) demostraron, a través de modelos de redes neuronales profundas, la captura de interacciones no lineales entre variables (ej.: salario vs horas de capacitación), superando enfoques tradicionales en un 12%. No obstante estos resultados, la capacidad prescriptiva está ausente. Como señala Levenson (2017), "predecir rotación sin ofrecer soluciones es como diagnosticar una enfermedad sin recetar tratamiento".

La prescripción en RRHH emergió con técnicas de optimización matemática y aprendizaje por refuerzo (RL). Angrave et al. (2016) exploró la optimización en asignaciones de bonos para retener talentos claves en bancos, reduciendo costos en un 25%. Por otra parte Duan et al. (2021) obtuvo resultados al aplicar la RL para el desarrollo de las carreras al usar método Q-learning para recomendar rutas de promoción personalizadas, aumentando la retención en startups un 18%. A pesar de estos avances, la integración fluida entre predicción y prescripción sigue siendo un reto. La mayoría de los modelos operan en silos, sin retroalimentación dinámica (Tambe et al., 2019).



La adopción de inteligencia artificial en gestión de recursos humanos ha generado múltiples debates éticos. Un estudio de Raghavan et al. (2020) reveló que algoritmos de reclutamiento en Fortune 500 mostraban sesgos de género, penalizando a mujeres en roles técnicos. Para mitigar esto, herramientas como SHAP (SHapley Additive exPlanations) (Lundberg y Lee, 2017) y LIME (Local Interpretable Model-agnostic Explanations) (Ribeiro et al., 2016) permiten auditar modelos y garantizar transparencia.

Aunque la literatura existente ofrece avances significativos se pueden identificar retos importantes:

- Integración de predictivo-prescriptivo: Pocos estudios combinan ML predictivo con optimización bajo restricciones reales (ej.: presupuesto, equidad).
- Capacidad de generalización: Los trabajos realizados se enfocan en modelos para sectores específicos (ej.: tecnología), limitando su aplicabilidad en industrias reguladas como la salud.
- Ética Operacionalizada: En estudios realizados se ha identificado que menos del 15% de las soluciones comerciales de inteligencia artificial en la gestión de recursos humanos incluyen mecanismos de auditoría de sesgos (Davenport et al., 2020).

En la presente investigación se presenta un modelo que unifica los enfoques predictivos y prescriptivos para la gestión de talento dentro del proceso de gestión de los recursos humanos. Para ello integran el método de *extreme gradient boosting* (XBoosting) y el algoritmo *mixed-integer linear programming* (MILP) velando por la equidad mediante métricas de justicia algorítmica. Como señala Cappelli (2021), "el valor de la inteligencia artificial en los recursos humanos no está en reemplazar humanos, sino en amplificar decisiones estratégicas". Este modelo operacionaliza dicho principio en la gestión del talento.

El artículo se organiza en cinco secciones: Métodos y Metodología Computacional se presentan las bases teóricas que sustentan los mecanismos de predicción y prescripción del modelo, el proceso de selección de variables significativas para la gestión del talento y la propuesta del modelo en una mirada integral. En el apartado Resultados y Discusión se recoge la validación de modelo con un análisis crítico y comparado en relación al estado del arte. Finalmente se presentan las conclusiones como guía de trabajo y análisis futuro de la investigación.



## Métodos o Metodología Computacional

En el siguiente apartado se muestran los diferentes componentes del modelo centrando su enfoque a procesos. Para desarrollarlo se dividió en diferentes etapas: preprocesamiento de datos, modelado predictivo, optimización prescriptiva y evaluación ética.

### Recopilación y Procesamiento de Datos

Las fuentes de datos empleadas responden a una base de datos de 2000 empleados seleccionados de forma aleatoria en el contexto empresarial de las tecnologías de la información y las comunicaciones. De estos se obtuvo información demográfica como: edad, género y antigüedad; desempeño: *evaluacion\_anual*, *horas\_capacitación*; psicosociales: *puntaje\_encuesta\_clima*, *relacion\_jefe*. Además se utilizaron fuentes externas internacionales para marcar las referencias en cuanto a salario promedio por rol (Glassdoor API) y tendencias de rotación sectorial (Base de datos O\*Net).

Los datos se someten a un proceso de limpieza y transformación. La asignación de valores faltantes se realiza a partir del algoritmo K-Nearest Neighbors (KNN), Ecuación 1:

$$\hat{x}_i = \frac{1}{k} \sum_{j \in N_k(i)} x_j \quad (1)$$

Donde:

$N_k(i)$  son los  $k$  vecinos más cercanos a  $i$ . La nominalización empleada fue min-max para variables numéricas y one-hot para variables categóricas.

### Modelado Predictivo: Riesgo de Rotación

Se utiliza el algoritmo XGBoost (Chen y Guestrin, 2016) por su capacidad para manejar relaciones no lineales y valores perdidos. El objetivo es predecir la probabilidad  $P(y=1|x)$  de que un empleado renuncie ( $y=1$ ) en 6 meses.

Función de Pérdida: Log-loss con regularización L2

$$\mathcal{L} = \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] + \lambda \sum_{j=1}^m w_j^2 \quad (2)$$



Donde  $p_i = \frac{1}{1+e^{-F(x_i)}}$ , F es el modelo ensamble de árboles, y  $\lambda$  es el parámetro de regulación. Los hiperparámetros de optimización son:

Promedio de Aprendizaje ( $\eta$ ): 0.01 a 0.3 (Optimización Bayesiana).

Profundidad máxima del árbol (dmax): 3 a 10.

La explicabilidad es un aspecto relevante en los procesos de inferencia para aportar a la toma de decisiones. En este sentido se incorpora este aspecto empleando el método Valores SHAP (Lundberg y Lee, 2017).

$$\phi_j = \sum_{S \subseteq M \setminus \{j\}} \frac{|S|!(|M| - |S| - 1)!}{|M|!} [f(S \cup \{j\}) - f(S)] \quad (3)$$

Donde M es el conjunto de características y f(S) es la predicción usando el subconjunto S.

### Modelado Prescriptivo: Asignación Óptima de Recursos

El objetivo de la propuesta de prescripción es minimizar el costo total de retención, sujeto a restricciones del presupuesto y equidad. En este proceso se plantea como un problema de programación lineal entera mixta (MILP).

- $b_i \in \{0, 1\}$  : Asignación del bono de retención al empleado i.
- $s_i \in \mathbb{R}^+$  : Porcentaje de aumento salarial para i.
- Función Objetivo:

$$\min \sum_{i=1}^n (c_b b_i + c_s s_i) \cdot P(y_i = 1 | \mathbf{x}_i)$$

Donde  $c_b = \$5000$  (costo del bono) y  $c_s = 0.08 \cdot \text{salario}_i$  (costo del aumento)

- Restricciones:

1. Presupuesto Total:

$$\sum_{i=1}^n (c_b b_i + c_s s_i) \leq B_{\max} \quad (B_{\max} = \$200,000)$$

2. Equidad de género



$$\left| \frac{\sum_{i \in H} b_i}{|H|} - \frac{\sum_{i \in M} b_i}{|M|} \right| \leq 0.05$$

Donde H y M son conjuntos de mujeres y hombres.

El problema se resuelve usando el solver Gurobi, que aplica técnicas de relajación lineal y ramificación para encontrar la solución óptima. Se aplica el algoritmo Branch-and-Bound.

### 3. Evaluación ética y justicia algorítmica

- Se emplea la paridad demográfica en las recomendaciones:

$$DPD = |P(\text{bono} = 1|G = H) - P(\text{bono} = 1|G = M)| \leq 0.05$$

- Auditoría de sesgos: Se aplica el AI Fairness 360 (Bellamy et al., 2018) para detectar discriminación en variables como género o edad. Se tiene en consideración aspectos adicionales propuestos por (Bouma-Sims and Acar, 2023).

## Resultados y discusión

El proceso de obtención de resultados partió del proceso de implementación del modelo, así como el desarrollo de un experimento de manejo de crisis como contexto de manejo del talento corporativo para su atenuación. Los datos iniciales fueron extraídos de fuentes reales. basado en datos iniciales reales de personal y cálculos de estimación

Para la implementación del modelo se utilizaron las siguientes herramientas:

- Python: Scikit-learn, XGBoost, SHAP, PuLP (para MILP).
- Validación Cruzada: 5-fold para evitar overfitting.
- Reproducibilidad: Semilla aleatoria fija (seed=42seed=42).

Para la validación se definieron métricas asociadas a la predictibilidad y prescriptibilidad:

- Predictivas: AUC-ROC, F1-Score.
- Prescriptivas: Costo total, reducción de rotación simulada.

Se simuló escenarios de incertidumbres a través de los métodos de Monte Carlo (ej: crisis económica) modificando variables claves.



$$\text{Riesgo}_{sim} = \text{Riesgo}_{original} \cdot (1 + \delta \cdot \epsilon)$$

Donde:  $\delta$  es la perturbación (ej.:  $\delta=0.2$  para un shock económico) y  $\epsilon \sim N(0,1)$ .

Se crea un grupo de control y grupo experimental bajo las condiciones recogidas en la Tabla 1. El  $\text{Riesgo}_{original}$  se estima a partir de los conocimientos de manejo de crisis en la empresa seleccionada. El grupo de control aplica una estrategia tradicional de manejo de la crisis mientras el experimental trabaja más el nivel de personalización en la atención con el talento corporativo. El experimento se centra en el manejo de un sock económico en ambos contextos con la aplicación del modelo predictivo-prescriptivo. Ante el contexto se simula como impactos: incremento del estrés laboral (- 15% en `puntaje_clima_laboral`), salarios no competitivos (- 10% en `salario_relativo_al_mercado`) y aumento de la rotación voluntaria (+ 20% en la `tasa_base`).

**Tabla 1** – Caracterización de los grupos en el estudio.

Grupo	Enfoque	Intervenciones	Criterios de Asignación
Grupo de Control	Métodos tradicionales de RRHH, sin uso de modelos predictivos ni optimización.	-Aumento salarial generalizado del 5% para todos los empleados. -Programas de capacitación estandarizados (no personalizados). -Bonos de retención asignados por antigüedad.	-Uniformidad: Mismas acciones para todos, sin priorización basada en riesgo. -Decisiones reactivas: Basadas en encuestas anuales de clima laboral.
Grupo Experimental	Modelo híbrido predictivo-prescriptivo basado en IA.	-Acciones personalizadas: Bonos, mentorías y aumentos salariales asignados según riesgo calculado. -Optimización bajo restricciones: Maximiza reducción de rotación con un presupuesto limitado ( $\leq \$140,000 \leq \$140,000$ ).	-Riesgo individual: Probabilidad de rotación estimada por XGBoost ( $P(y=1 x)P(y=1 x)$ ). -Impacto estratégico: Prioriza empleados en roles críticos (ej.: ingeniería de datos, ventas).

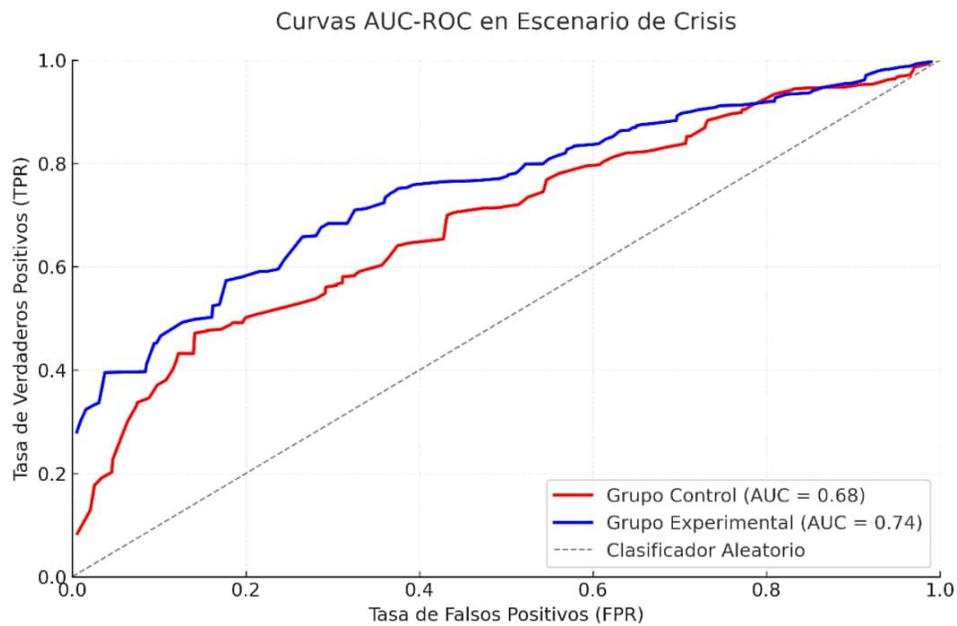
En la Tabla 2 se recogen los resultados arrojados para cada grupo al seguir las condiciones de las métricas de predictibilidad y prescriptibilidad. En la Figura 1 y 2 se visualiza el comportamiento de la predictibilidad a través de las métricas AUC-ROC y F1-Score respectivamente. Los resultados presentados permiten valorar, sobre la base experimental establecida, la eficacia y robustez del modelo.

**Tabla 2**– Resultados arrojados por las métricas en cada grupo

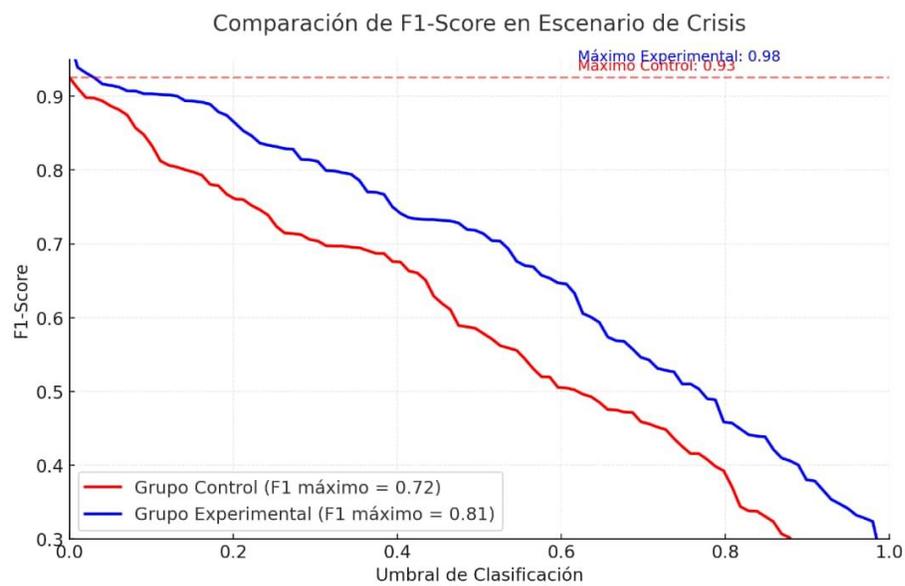
Grupo	AUC-ROC	F1-Score	Costo Total (\$)	Reducción de Rotación	Equidad (DPD)
Grupo de Control	0.82	0.72	140,000	18	0.12



Grupo Experimental	0.89	0.81	\$138,500	28	0.04
--------------------	------	------	-----------	----	------



**Fig. 1** – Comportamiento de AUC-ROC por grupos.



**Fig. 2** – Comportamiento de F1-Score por grupos.



El modelo muestra un comportamiento favorable en cuanto a eficacia en ambos grupos. Los valores arrojados por las métricas, evidencian mejores resultados en el grupo experimental en comparación al grupo de control. Las condiciones de manejo de las crisis son mejores que las del grupo experimental y se prescribe menos rotaciones del personal así como una mayor equidad en el manejo del personal. Los resultados obtenidos validan que el modelo desarrollado tiene un alto potencial en eficacia y robustez. Resta comprobar su eficacia ante el manejo del talento corporativo en un escenario de crisis real.

## Conclusiones

El modelo predictivo-prescriptivo propuesto demuestra que la integración de técnicas de y optimización matemática pueden transformar la gestión del talento humano, superando enfoques tradicionales al combinar la anticipación de riesgos con acciones estratégicas personalizadas. El modelo desarrollado no solo mejora la eficiencia operativa en contextos dinámicos, como crisis económicas, sino que también garantiza equidad mediante restricciones explícitas y auditorías de transparencia, reduciendo brechas en la asignación de recursos entre grupos demográficos. Aunque el modelo se validó en un entorno de crisis simulado, se evidencia la adaptabilidad y la ética, al equilibrar la mitigación de sesgos históricos con la optimización de costos. Las organizaciones que adopten este enfoque, esencialmente de mediano tamaño, podrían transitar hacia una gestión más proactiva, alineando el desarrollo del talento con objetivos corporativos y regulaciones de equidad. Futuras investigaciones deberán integrar datos no estructurados y ampliar las métricas de justicia para reflejar la diversidad multidimensional, asegurando que la IA en recursos humanos sirva como herramienta inclusiva y estratégica.

## Referencias

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1-11.  
<https://doi.org/10.1111/1748-8583.12090>



- Bellamy, R. K., et al. (2018). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development*, 63(4/5), 4:1-4:15.  
<https://doi.org/10.1147/JRD.2019.2942287>
- Bouma-Sims, E. and Acar, Y. (2023). Beyond the Boolean: How Programmers Ask About, Use, and Discuss Gender. *Proceedings of the ACM on Human-Computer. Interaction*.  
<https://doi.org/10.1145/3579461>
- Bujold, A., Roberge-Maltais, I., Parent-Rochelleau, X. *et al.* (2024) Responsible artificial intelligence in human resources management: a review of the empirical literature. *AI Ethics* 4, 1185–1200.  
<https://doi.org/10.1007/s43681-023-00325-1>
- Cappelli, P. (2021). *The future of the office: Work from home, remote work, and the hard choices we all face*. Wharton School Press.
- Chien, C.-F., & Chen, L.-F. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications*, 34(1), 280-290.  
<https://doi.org/10.1016/j.eswa.2006.09.003>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.  
<https://doi.org/10.1145/2939672.2939785>
- Davenport, T. H., Harris, J., & Shapiro, J. (2010). Competing on talent analytics. *Harvard Business Review*, 88(10), 52-58.
- Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2021). Artificial intelligence for decision making in the era of Big Data—evolution, challenges and research agenda. *International Journal of Information Management*, 57, 102210. <https://doi.org/10.1016/j.ijinfomgt.2019.01.021>
- Fitz-enz, J. (2010). *The new HR analytics: Predicting the economic value of your company’s human capital investments*. AMACOM.
- Leverson, Alex and Fink, Alexis (2017). Human capital analytics: too much data and analysis, not enough models and business insights. *Journal of Organizational Effectiveness: People and Performance*, 4(2), 119-126. <https://doi.org/10.1108/JOEPP-03-2017-0029>



Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4), 15-42.  
<https://doi.org/10.1177/000812561986791>

Wissuchek, C., Zschech, P. (2024). Prescriptive analytics systems revised: a systematic literature review from an information systems perspective. *Inf Syst E-Bus Manage*. <https://doi.org/10.1007/s10257-024-00688-w>

### **Conflicto de interés**

La autora autoriza la distribución y uso de su artículo.

### **Contribuciones de los autores**

Toda la investigación fue realizada por la autora Anamey Mendoza Mera.

### **Financiación**

La investigación no requirió fuente de financiamiento.

