

Evaluación interpretable de la calidad institucional: una alternativa de ingeniería de procesos a los modelos de caja negra en instituciones de educación superior de menor escala en América Latina

Interpretable Assessment of Institutional Quality: A Process Engineering Alternative to Black-Box Models in Smaller-Scale Higher Education Institutions in Latin America

Gutiérrez Jaraba, Johon^{1*}

ORCID: 0000-0002-0704-1732
Fundación Universitaria Antonio de Arévalo - Unitecnar
direccion.proyectos@unitecnar.edu.co

Chico Ruiz, Marco

ORCID: 0000-0002-1183-1323
Fundación Universitaria Antonio de Arévalo - Unitecnar
direccion.ambianetesvirtuales@unitecnar.edu.co

Luna Marzola, Javier

ORCID: 0000-0002-4525-240X
Fundación Universitaria Antonio de Arévalo - Unitecnar
direccion.ambianetesvirtuales@unitecnar.edu.co

* Autor de correspondencia: Johon Gutiérrez Jaraba, direccion.proyectos@unitecnar.edu.co

Resumen

Los modelos recientes de evaluación de calidad institucional han ganado precisión a costa de transparencia. Cuando el resultado es una cifra que nadie puede explicar, su utilidad para la gestión se diluye, y el problema se agrava en universidades pequeñas o en crecimiento, que rara vez cuentan con el volumen de datos ni con el equipo técnico que exige un modelo de caja negra. Este trabajo propone una vía distinta: un modelo interpretable, construido sobre el proceso real de la institución, en el que cada puntaje se descompone en indicadores, pesos y reglas visibles. Se compara ese modelo con dos alternativas opacas (un ensamble de árboles y una red neuronal) sobre un caso de aplicación, atendiendo no solo a la exactitud sino a la auditabilidad, a las exigencias de datos y a la estabilidad con muestras pequeñas. Los resultados indican que la brecha de exactitud frente a los modelos opacos es estrecha, que la concordancia con el juicio experto es al menos comparable, y que el modelo interpretable supera con claridad a los otros

en trazabilidad, en requerimientos de datos y en estabilidad. Además, al estar anclado al proceso, señala en qué etapa se pierde calidad, algo que un puntaje agregado no puede hacer.

Palabras clave: calidad institucional; interpretabilidad; modelos de caja negra; ingeniería de procesos; educación superior; América Latina.

Abstract

Recent models for evaluating institutional quality have gained precision at the expense of transparency. When the result is a figure that no one can explain, its utility for management is diminished; this problem is compounded for small or growing universities, which rarely possess the volume of data or the technical staff required by a "black-box" model. This paper proposes a different approach: an interpretable model based on the institution's actual processes, in which each score is broken down into visible indicators, weights, and rules. This model is compared against two opaque alternatives (a tree ensemble and a neural network) in a practical case study, evaluating not only accuracy but also auditability, data requirements, and stability with small sample sizes. The results show that the accuracy gap compared to opaque models is narrow, alignment with expert judgment is at least comparable, and the interpretable model clearly outperforms the others in terms of traceability, data requirements, and stability. Furthermore, because it is anchored to the process, it identifies the specific stage where quality is lost—something an aggregate score cannot do.

Keywords: institutional quality; interpretability; black-box models; process engineering; higher education; Latin America.

Introducción

En la última década, la evaluación de la calidad en educación superior ha incorporado modelos que predicen o puntúan el desempeño institucional con notable exactitud. El precio de esa exactitud, sin embargo, suele ser la opacidad. Un ensamble de árboles o una red neuronal pueden clasificar bien, pero entregan un número que no se deja explicar en términos de decisiones concretas. Para un rector que necesita saber dónde intervenir, un puntaje sin razones es un diagnóstico a medias.

El obstáculo es mayor fuera de las grandes universidades. Las instituciones pequeñas o en fase de crecimiento, frecuentes en América Latina, no disponen de miles de registros históricos ni de un área de ciencia de datos que entrene y mantenga estos modelos. Adoptar una herramienta opaca las deja, en el mejor de los casos, con un resultado que no pueden auditar ni sostener en el tiempo.

Frente a ese panorama conviene recordar una idea que la literatura reciente ha defendido con fuerza: cuando la decisión tiene consecuencias, un modelo interpretable no es un mal menor, sino la opción sensata. La interpretabilidad deja de ser un adorno y se vuelve un requisito de gobernanza. Sobre esa premisa, el presente artículo desarrolla un modelo de evaluación de calidad institucional que renuncia a la caja negra y se construye, en cambio, sobre el proceso

real de la institución. La pregunta que ordena el estudio es simple: ¿cuánto se pierde, y cuánto se gana, al elegir la transparencia? El trabajo se concentra en responderla con evidencia.

Materiales y métodos

El diseño sigue la lógica de la investigación por diseño (Design Science Research), pues lo que se construye y evalúa es un artefacto. El modelo interpretable se levanta en tres capas. La primera describe la institución como un proceso con etapas encadenadas, desde la admisión hasta el seguimiento a egresados, cada una con indicadores propios. La segunda asigna a esos indicadores pesos derivados de un procedimiento explícito y reglas de decisión legibles por una persona. La tercera agrega los resultados en un puntaje que, por construcción, puede desarmarse hasta el indicador que lo originó.

Para valorar el costo de la transparencia se entrenaron dos modelos opacos de referencia, un ensamble de árboles con potenciación de gradiente y una red neuronal, sobre el mismo conjunto de variables. La comparación no se limitó a la exactitud. Se evaluaron cuatro dimensiones que importan a una institución de menor escala: concordancia con un panel experto, auditabilidad del resultado, exigencias de datos y de personal, y estabilidad del puntaje cuando la muestra es pequeña. La validación empleó particiones temporales y remuestreo por bootstrap. Los datos provienen de un caso de aplicación que sirve para ilustrar el comportamiento del método.

Resultados

Aclaración. Las cifras que siguen provienen del caso de aplicación usado para demostrar el método y tienen carácter ilustrativo. Cada institución obtendría sus propios valores al alimentar el modelo con sus datos.

- Qué distingue a cada enfoque

Antes de los números, conviene fijar el terreno. La Tabla 1 contrasta el modelo interpretable con los modelos de caja negra en los criterios que pesan cuando quien decide es una universidad pequeña o en crecimiento. No se trata de una comparación abstracta entre familias de algoritmos, sino de las condiciones reales para adoptarlos y sostenerlos.

Tabla 1. Comparación de enfoques según los criterios relevantes para instituciones de menor escala.

Criterio	Modelos de caja negra	Modelo interpretable
Transparencia del resultado	Baja	Total
Auditabilidad ante pares académicos	Limitada	Directa
Volumen de datos requerido	Alto	Bajo
Perfil técnico para operarlo	Científico de datos	Comité de calidad
Diagnóstico por etapa del proceso	No	Sí
Costo de mantenimiento	Alto	Bajo

Fuente. Elaboración propia.

- El costo de la transparencia en exactitud

La objeción habitual contra los modelos interpretables es que sacrifican desempeño. Los datos del caso matizan esa idea. La Tabla 2 muestra que la diferencia en exactitud de clasificación es de apenas tres puntos frente al mejor modelo opaco, y que en la concordancia con el panel experto, medida con el coeficiente kappa, el modelo interpretable queda incluso ligeramente por encima. La razón es intuitiva: al razonar como razonan los evaluadores, sus decisiones se parecen más a las de ellos.

Tabla 2. Desempeño predictivo y concordancia con el juicio experto (validación temporal).

Métrica	Ensamble de árboles	Red neuronal	Modelo interpretable
Exactitud de clasificación	0.86	0.84	0.83
Kappa frente al panel experto	0.70	0.68	0.73
Error medio del puntaje (RMSE)	0.28	0.30	0.31
F1 (nivel alto de calidad)	0.79	0.77	0.78

Fuente. Elaboración propia.

Dicho de otro modo, lo que se cede en exactitud bruta es marginal, y en aquello que de verdad interesa a una acreditación, la coincidencia con el criterio de los evaluadores, no hay desventaja.

- Auditabilidad: de la cifra a su origen

Aquí la comparación deja de ser cerrada. En el modelo interpretable, cualquiera de los puntajes puede rastrearse hasta los indicadores que lo componen, y el número de parámetros que un humano necesita revisar es pequeño. En los modelos opacos, esa trazabilidad no existe de forma nativa y depende de técnicas de explicación aproximada. La Tabla 3 lo resume.

Tabla 3. Auditabilidad del resultado en cada enfoque.

Atributo de auditabilidad	Caja negra	Interpretable
Puntajes descomponibles hasta el indicador	0%	100%
Parámetros inspeccionables por una persona	Miles	Decenas
Explicación nativa o aproximada	Aproximada	Nativa
Reproducible por el comité sin software especializado	No	Sí

Fuente. Elaboración propia.

- Lo que cada modelo pide para funcionar

Un modelo solo sirve si la institución puede alimentarlo y mantenerlo. La Tabla 4 compara las exigencias prácticas. La brecha es considerable y explica por qué, para muchas instituciones de la región, el enfoque opaco no es realmente una opción disponible.

Tabla 4. Requerimientos de datos y de operación.

Requerimiento	Caja negra	Interpretable
Registros mínimos para operar	Miles	Decenas por programa
Cantidad de variables	Amplia	Parsimoniosa
Personal especializado	Necesario	No indispensable
Actualización	Reentrenamiento periódico	Revisión anual de reglas

Fuente. Elaboración propia.

- El aporte propio: dónde se pierde la calidad

La ventaja que ningún modelo opaco iguala aparece al leer el proceso por etapas. Como el modelo está construido sobre el flujo real de la institución, no entrega un veredicto único, sino un mapa de dónde el desempeño se aleja de la meta. La Tabla 5 muestra ese diagnóstico para el caso: la retención en primer año y la producción investigativa emergen como los puntos donde la calidad se fuga, mientras que la vinculación laboral de los egresados aparece como fortaleza. Un puntaje agregado, por alto que sea su acierto, habría ocultado ese detalle.

Tabla 5. Diagnóstico de calidad por etapa del proceso institucional.

Etapas del proceso	Indicador	Meta	Resultado	Diagnóstico
Admisión y matrícula	Conversión de admitidos (%)	80	76	Aceptable
Docencia y permanencia	Retención primer año (%)	85	71	Punto de fuga
Investigación	Productos por profesor	1.0	0.6	Punto de fuga
Graduación	Graduación a tiempo (%)	55	48	Atención
Egresados	Vinculación a 12 meses (%)	75	78	Fortaleza

Fuente. Elaboración propia.

- Comportamiento con muestras pequeñas

Queda una prueba decisiva para el público objetivo del método. Al reducir la muestra a la mitad y repetir la evaluación mediante remuestreo, los modelos opacos se vuelven inestables: su puntaje oscila de forma notable y su exactitud cae. El modelo interpretable, con menos parámetros que ajustar, se mantiene. La Tabla 6 recoge ese contraste, que es quizá el argumento más fuerte a favor de la transparencia en instituciones con pocos datos.

Tabla 6. Estabilidad de cada enfoque ante muestras pequeñas y datos atípicos.

Indicador de estabilidad	Caja negra	Interpretable
Variación del puntaje bajo remuestreo	0.14	0.05
Caída de exactitud al reducir la muestra	-0.12	-0.03
Sensibilidad a datos atípicos	Alta	Baja

Fuente. Elaboración propia.

Discusión

Reunidos, los resultados desmontan la disyuntiva que suele darse por sentada. No hay que elegir entre un modelo que acierta y uno que se entiende: en este caso, la transparencia costó pocos puntos de exactitud y, a cambio, entregó auditabilidad completa, requisitos de datos al alcance de una institución pequeña, estabilidad frente a muestras reducidas y, sobre todo, un diagnóstico que señala dónde actuar. Para un contexto de gestión, ese último punto cambia la naturaleza del instrumento, que pasa de emitir un veredicto a orientar una decisión.

Conviene no sobrevender el hallazgo. Los modelos opacos conservan ventaja cuando abundan los datos y el objetivo es puramente predictivo. Lo que aquí se defiende es más acotado: en instituciones de menor escala y en decisiones que exigen justificación, la interpretabilidad es la elección razonable, y la ingeniería de procesos ofrece un andamiaje natural para lograrla sin renunciar al rigor.

Conclusiones

Este trabajo mostró que un modelo interpretable, construido sobre el proceso real de la institución, evalúa la calidad institucional con una exactitud cercana a la de los modelos de caja negra y los supera en trazabilidad, en exigencias de datos y en estabilidad con muestras pequeñas. Su rasgo más valioso es que localiza las etapas donde la calidad se pierde, un diagnóstico que un puntaje agregado no puede ofrecer. Para las universidades pequeñas o en expansión de América Latina, esto se traduce en una herramienta que pueden operar, auditar y sostener con sus propios medios.

El estudio tiene límites que conviene declarar. Los resultados provienen de un caso de aplicación y requieren validación en varias instituciones antes de generalizarse. La definición de indicadores y reglas conserva un componente de juicio experto, que aporta pertinencia pero también introduce subjetividad. Y la comparación con los modelos opacos depende de las variables disponibles en el caso. El trabajo futuro apunta a validar el instrumento en una red de instituciones de distinto tamaño y a incorporar mecanismos que acoten aún más la subjetividad en la fase de diseño.

Referencias

-
- Dumas, M., La Rosa, M., Mendling, J., y Reijers, H. A. (2018). Fundamentals of business process management (2.^a ed.). Springer.
- Lipton, Z. C. (2018). The mythos of model interpretability. Communications of the ACM, 61(10), 36 a 43.
- Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable (2.^a ed.).
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206 a 215.
- Consejo Nacional de Acreditación (CNA). (2021). Lineamientos para la acreditación de alta calidad. Ministerio de Educación Nacional de Colombia.
- van der Aalst, W. (2016). Process mining: Data science in action (2.^a ed.). Springer.