



UNIVERSIDAD NACIONAL DE CHIMBORAZO
VICERRECTORADO DE INVESTIGACIÓN, VINCULACIÓN Y
POSGRADO
DIRECCIÓN DE POSGRADO

Desarrollo de un modelo de árboles de decisión para evaluar y optimizar el cumplimiento de estándares de aseguramiento de la calidad en el Instituto Superior Universitario Carlos Cisneros

Trabajo de Titulación para optar al título de
Magíster en Matemática Aplicada con mención en Matemática Computacional

AUTORA:

Arq. Carolina Valeria Llanga Cruz, Mgs.

TUTORA:

Ing. Guillermo Edvin Machado Sotomayor, PhD.

Riobamba, Ecuador. 2025

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

Yo, **Carolina Valeria Llanga Cruz** con número único de identificación **0603358516**, declaro y acepto ser responsable de las ideas, doctrinas, resultados y lineamientos alternativos realizados en el presente trabajo de titulación denominado: “Desarrollo de un Modelo de Árboles de Decisión para Evaluar y Optimizar el Cumplimiento de Estándares de Aseguramiento de la Calidad en el Instituto Superior Universitario Carlos Cisneros.” previo a la obtención del grado **de Magíster en Matemática Aplicada con mención en Matemática Computacional**.

Declaro que mi trabajo investigativo pertenece al patrimonio de la Universidad Nacional de Chimborazo de conformidad con lo establecido en el artículo 20 literal j) de la Ley Orgánica de Educación Superior LOES.

Autorizo a la Universidad Nacional de Chimborazo que pueda hacer uso del referido trabajo de titulación y a difundirlo como estime conveniente por cualquier medio conocido, y para que sea integrado en formato digital al Sistema de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor, dando cumplimiento de esta manera a lo estipulado en el artículo 144 de la Ley Orgánica de Educación Superior LOES.

Riobamba, 22 de julio de 2025



Arq. Carolina Valeria Llanga Cruz Mgs.

C.I.: 0603358516

 UNIVERSIDAD NACIONAL DE CHIMBORAZO	NOMBRE DEL FORMATO		 SGC <small>SISTEMA DE GESTIÓN DE LA CALIDAD UNIVERSIDAD NACIONAL DE CHIMBORAZO</small>
	CÓDIGO:	VERSIÓN:	
	FECHA:		
	MACROPROCESO:		
PROCESO:			
SUBPROCESO:			

Riobamba, 22 de julio de 2025

ACTA DE CULMINACIÓN DEL TRABAJO DE TITULACIÓN

En calidad de miembros del Tribunal designados por la Comisión de Posgrado, CERTIFICAMOS que una vez revisado el Trabajo de titulación bajo la modalidad Proyecto de Investigación y/o desarrollo denominado **“Desarrollo de un Modelo de Árboles de Decisión para Evaluar y Optimizar el Cumplimiento de Estándares de Aseguramiento de la Calidad en el Instituto Superior Universitario Carlos Cisneros”**, dentro de la línea de investigación de ciencia de datos y sistemas inteligentes, presentado por el maestrante **Carolina Valeria Llanga Cruz** portador de la CI. **0603358516**, del programa de Maestría en matemática aplicada con mención en matemática computacional, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,



Firmado electrónicamente por:
GUILLERMO EDVIN MACHADO SOTOMAYOR
Validar únicamente con FirmasC

PhD. Guillermo Edvin Machado Sotomayor

TUTOR



Firmado electrónicamente por:
CIRO DIEGO RADICELLI GARCIA
Validar únicamente con FirmasC

PhD. Ciro Diego Radicelli García

MIEMBRO DEL TRIBUNAL

1



Firmado electrónicamente por:
ESTALIN FABIAN MEJIA HIDALGO
Validar únicamente con FirmasC

Mgs. Estalin Fabián Mejía Hidalgo

MIEMBRO DEL TRIBUNAL

2



Riobamba, 30 de julio del 2025

CERTIFICADO

De mi consideración:

Yo Guillermo Edvin Machado Sotomayor, certifico que **Carolina Valeria Llanga Cruz** con cédula de identidad No. 0603358516 estudiante del programa de **Maestría en Matemática Aplicada con Mención en Matemática Computacional**, cohorte Segunda presentó su trabajo de titulación bajo la modalidad de Proyecto de titulación con componente de investigación aplicada/desarrollo denominado: **“Desarrollo de un Modelo de Árboles de Decisión para Evaluar y Optimizar el Cumplimiento de estándares de Aseguramiento de la Calidad en el Instituto Superior Universitario Carlos Cisneros”**, el mismo que fue sometido al sistema de verificación de similitud de contenido COMPILATION identificando el porcentaje de similitud 2% en el texto y el 10% en inteligencia artificial.

Es todo en cuanto puedo certificar en honor a la verdad.

Atentamente,



PhD. Guillermo Edvin Machado Sotomayor

CI: 0602166027

Adj.-

- Resultado del análisis de similitud(Compilation)

IV

DEDICATORIA

Dedico este trabajo de titulación a mi hermano, quien, aunque ya no esté físicamente, permanece como una fuente constante de inspiración en mi vida profesional, por su ejemplo de constancia, dedicación y firmeza en cada meta que se proponía.

AGRADECIMIENTO

Agradezco profundamente a mi tutor y a los miembros del tribunal por su paciencia, apoyo y ejemplo a lo largo de este proceso. A mi familia, a mi esposo y a todas las personas que, dentro de sus posibilidades, me brindaron su ayuda y respaldo incondicional, les expreso mi más sincera gratitud.

ÍNDICE GENERAL

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS	II
ACTA DE CULMINACIÓN DEL TRABAJO DE TITULACIÓN	III
CERTIFICADO DEL CONTENIDO DE SIMILITUD.....	IV
DEDICATORIA.....	V
AGRADECIMIENTO.....	VI
ÍNDICE GENERAL	VII
ÍNDICE DE ILUSTRACIONES.....	XIII
ÍNDICE DE TABLAS.....	XIV
ÍNDICE DE ANEXOS	XV
RESUMEN	- 1 -
ABSTRACT	- 2 -
Introducción.....	- 3 -
Capítulo 1	- 4 -
Generalidades	- 4 -
1.1 Planteamiento del Problema	- 4 -
1.2 Justificación de la investigación	- 4 -
1.3 Objetivos	- 5 -
1.3.1 Objetivo General:	- 5 -
1.3.2 Objetivos Específicos:	- 5 -
1.3.3 Población y Muestra	- 5 -
Capítulo 2	- 7 -
2 Estado del Arte y la Práctica	- 7 -
2.1 Antecedentes investigativos	- 7 -

2.1.1	Antecedentes	- 7 -
2.2	Fundamentación Legal Nacional	- 10 -
2.3	Ciencia de datos, para la toma de decisiones en marcos evaluativos del CACES	- 12 -
2.3.1	<i>El rol del CACES en la evaluación de indicadores de calidad</i> - 12 -	
2.3.2	<i>Aplicación de modelos predictivos en el marco de evaluación del CACES</i>	- 13 -
2.3.3	<i>Importancia de la acreditación de los Institutos Superiores Universitarios y la oportunidad del uso de machine learning</i>	- 13 -
Ilustración 1	<i>Estructura del modelo de evaluación</i>	- 14 -
2.4	Fundamentos de la Minería de Datos y Aprendizaje Automático ...	- 16 -
2.4.1	<i>Minería de Datos</i>	- 16 -
2.4.2	<i>Árboles de Decisión</i>	- 18 -
2.4.3	<i>Random Forest</i>	- 19 -
2.5	Preprocesamiento y Transformación de Datos	- 21 -
2.5.1	<i>Limpieza de datos: valores faltantes e inconsistencias</i> - 21 -	
2.5.2	<i>Mapeo de escalas cualitativas a valores numéricos</i>	- 22 -
2.5.3	<i>Escalado y normalización de variables</i>	- 22 -
2.5.4	<i>One-hot encoding para variables categóricas</i>	- 23 -
2.5.5	<i>Uso de variables dummy</i>	- 24 -
2.6	Árboles de Decisión: Construcción y Aplicación	- 25 -
2.6.1	<i>Algoritmo de árbol de decisión</i>	- 25 -
Ilustración 2	<i>Árbol de decisión Salary</i>	- 26 -
2.6.2	<i>Evaluación de indicadores de calidad</i>	- 27 -

2.6.3	<i>Modelos agregados vs. modelos granulares</i>	- 28 -
2.7	Evaluación de Modelos Predictivos	- 29 -
2.7.1	<i>Validación cruzada</i>	- 29 -
2.7.2	<i>Métricas: accuracy, precisión, recall y F1-score</i>	- 30 -
2.7.3	<i>Curvas ROC y análisis de sensibilidad</i>	- 31 -
2.8	Importancia de Variables e Indicadores	- 31 -
2.8.1	<i>Cálculo de importancia de características</i>	- 31 -
2.8.2	<i>Indicadores mediante variables dummy</i>	- 32 -
2.8.3	<i>Comparación de relevancia: enfoque agregado y granular</i> -	
	33 -	
2.9	Interpretación de la Importancia de Variables y Análisis de	
	Contribución de Indicadores	- 34 -
Capítulo 3		- 36 -
3	Diseño Metodológico	- 36 -
3.1	Enfoque de la Investigación	- 36 -
3.2	Origen de datos y preprocesamiento	- 37 -
3.2.1	<i>Autoevaluación 2024:</i>	- 37 -
3.2.2	<i>Series históricas 2023-2024</i>	- 37 -
3.3	Entorno de desarrollo y librerías	- 38 -
Ilustración 3 <i>Librerías usadas para la investigación</i>		- 38 -
Tabla 1 <i>Códigos usados según su propósito</i>		- 38 -
3.4	Flujo metodológico general	- 40 -
3.5	Procedimientos de modelado	- 41 -
3.5.1	<i>Árbol de decisión global</i>	- 41 -
3.5.2	<i>Árbol de decisión granular</i>	- 42 -

3.5.3	Selección de hiperparámetros.....	- 43 -
3.5.4	<i>Estimación de importancia de variables (Random Forest).</i>	- 44 -
3.5.5	<i>Comparación histórica 2023-2024</i>	- 45 -
Capítulo 4	- 46 -
4	Análisis y Discusión de Resultados	- 46 -
4.1	Lectura y descripción inicial del conjunto de datos	- 46 -
Tabla 2	<i>Indicadores utilizados en la autoevaluación ISUCC 2024</i>	- 47 -
Tabla 3	<i>Resumen de características del DataFrame</i>	- 50 -
4.2	Resultados del Árbol de Decisión Global	- 52 -
Ilustración 4	<i>Diagrama del árbol de decisión global</i>	- 53 -
Tabla 4	<i>Resumen de umbrales y distribución de clases en cada nodo</i>	- 54 -
4.3	Validación Cruzada de Árbol Global	- 55 -
Tabla 5	<i>Exactitud por pliegue (K-Fold)</i>	- 55 -
Tabla 6	<i>Métricas por pliegue (accuracy, precisión, recall, F1, ROC-AUC)</i>	- 56 -
4.4	Reentrenamiento final y ajustes operativos	- 57 -
Ilustración 5	<i>Diagrama del árbol de decisión global final</i>	- 58 -
Tabla 7	<i>Comparación de métricas y distribución de clases entre modelos 80/20 y modelo final</i>	- 58 -
4.5	Resultados del Árbol de Decisión Granular	- 60 -
Ilustración 6	<i>Diagrama del árbol de decisión granular</i>	- 60 -
4.6	Validación del Árbol Granular	- 61 -
Tabla 8	<i>Exactitud por pliegue- árbol granular</i>	- 62 -
Tabla 9	<i>Métricas por pliegue (precisión, recall, f1)</i>	- 62 -
4.7	Importancia de Variables con Random Forest	- 62 -

Tabla 10	<i>Importancia de variables (Top 15)</i>	- 63 -
Ilustración 7	<i>Gráfico de barras horizontales variables de mayor importancia</i>	- 64 -
4.8	Comparación Histórica 2023-2024	- 65 -
Tabla 11	<i>Evolución 2023 vs 2024 por criterio</i>	- 65 -
Ilustración 8	<i>Gráfico de barras horizontales variables de delta interanual</i>	- 66 -
Capítulo 5	- 68 -
5	Marco Propositivo	- 68 -
5.1	Sistema de alerta temprana	- 68 -
Tabla 12	<i>Diagrama de semáforo</i>	- 68 -
5.2	Priorización según importancia de variables	- 68 -
Tabla 13	<i>Priorización de objetivos según su importancia</i>	- 69 -
Tabla 14	<i>Evidencias y oportunidades</i>	- 70 -
5.3	Plan de Acción	- 72 -
5.3.1	Implementación de semáforo de cumplimiento	- 72 -
5.3.2	Focalización de recursos en indicadores críticos	- 73 -
5.3.3	Validación continua y refinamiento.	- 73 -
5.3.4	Agrupación de indicadores y poda de modelos	- 73 -
5.3.5	Capacitación y rendición de cuentas	- 73 -
5.3.6	Paquete de intervención en el bloque “Profesores”	- 73 -
Tabla 15	<i>Paquete de intervención en el bloque “Profesores”</i>	- 73 -
5.3.7	Consolidar el bloque de I+D e Innovación	- 74 -
5.3.8	Revisión fina de estándares exigentes	- 74 -
5.3.9	Fortalecimiento del sistema de monitoreo	- 74 -
5.3.10	Validación continua de los modelos	- 75 -
5.4	Matriz de diagnóstico y plan de acción (riesgo/oportunidad)	- 75 -

Tabla 16 <i>Matriz de diagnóstico y plan de acción</i>	- 75 -
Tabla 17 <i>Impacto esperado</i>	- 76 -
5.5 Limitaciones del estudio.....	- 77 -
5.6 Trabajos a futuro.....	- 78 -
Conclusiones	- 79 -
Recomendaciones	- 81 -
Referencias Bibliográficas.....	1
Anexos.....	6

ÍNDICE DE ILUSTRACIONES

Ilustración 1 <i>Estructura del modelo de evaluación</i>	- 14 -
Ilustración 2 <i>Árbol de decisión Salary</i>	- 26 -
Ilustración 3 <i>Librerías usadas para la investigación</i>	- 38 -
Ilustración 4 <i>Diagrama del árbol de decisión global</i>	- 53 -
Ilustración 5 <i>Diagrama del árbol de decisión global final</i>	- 58 -
Ilustración 6 <i>Diagrama del árbol de decisión granular</i>	- 60 -
Ilustración 7 <i>Gráfico de barras horizontales variables de mayor importancia</i>	- 64 -
Ilustración 8 <i>Gráfico de barras horizontales variables de delta interanual</i>	- 66 -

ÍNDICE DE TABLAS

Tabla 1 <i>Códigos usados según su propósito</i>	- 38 -
Tabla 2 <i>Indicadores utilizados en la autoevaluación ISUCC 2024</i>	- 47 -
Tabla 3 <i>Resumen de características del DataFrame</i>	- 50 -
Tabla 4 <i>Resumen de umbrales y distribución de clases en cada nodo</i>	- 54 -
Tabla 5 <i>Exactitud por pliegue (K-Fold)</i>	- 55 -
Tabla 6 <i>Métricas por pliegue (accuracy, precisión, recall, F1, ROC-AUC) ...</i>	- 56 -
Tabla 7 <i>Comparación de métricas y distribución de clases entre modelos 80/20 y modelo final</i>	- 58 -
Tabla 8 <i>Exactitud por pliegue- árbol granular</i>	- 62 -
Tabla 9 <i>Métricas por pliegue (precisión, recall, f1)</i>	- 62 -
Tabla 10 <i>Importancia de variables (Top 15)</i>	- 63 -
Tabla 11 <i>Evolución 2023 vs 2024 por criterio</i>	- 65 -
Tabla 12 <i>Diagrama de semáforo</i>	- 68 -
Tabla 13 <i>Priorización de objetivos según su importancia</i>	- 69 -
Tabla 14 <i>Evidencias y oportunidades</i>	- 70 -
Tabla 15 <i>Paquete de intervención en el bloque “Profesores”</i>	- 73 -
Tabla 16 <i>Matriz de diagnóstico y plan de acción</i>	- 75 -
Tabla 17 <i>Impacto esperado</i>	- 76 -

ÍNDICE DE ANEXOS

ANEXO A: Capturas del Código realizado

ANEXO B: Enlace al código de la investigación en Google Colab

ANEXO C: Enlace a las bases de datos usadas para la investigación

ANEXO D: Encuesta utilizada para la recolección de datos

ANEXOS E: Arte desarrollado para la campaña de Facebook Ads

RESUMEN

El objetivo de la presente investigación fue diseñar y validar un sistema analítico para diagnosticar el cumplimiento de los 42 indicadores del Modelo 2024 del CACES en el Instituto Superior Universitario Carlos Cisneros. Como punto de partida se utilizó un DataFrame que ya incorporaba, las valoraciones cualitativas convertidas en los campos “peso ideal” y “peso real”. A partir de este conjunto se generó la etiqueta binaria “cumple” ($\text{ratio} \geq 0,60$). Sobre esta base de datos se entrenó un árbol de decisión global (profundidad máxima = 3) mediante una división entrenamiento–prueba 80 %/20 % con `train_test_split`. La visualización del árbol identificó umbrales operativos en PESOS DE LA IES (0,009 y 0,012) y en B PESOS SIN TP (0,012 y 0,017), formulados como reglas condicionales de alta interpretabilidad. La validación cruzada estratificada en cinco pliegues (StratifiedKFold) arrojó una exactitud media del 97 % y un recall del 100 %, con falsos positivos puntuales en un pliegue.

Para capturar la identidad de cada indicador se aplicó one-hot encoding y se construyó un árbol granular (profundidad = 5) que alcanzó 100 % de exactitud sobre la muestra y reveló excepciones para indicadores docentes como “Formación académica en curso y capacitación”. Para cuantificar la relevancia predictiva de las variables se entrenó un Random Forest (50 árboles, profundidad = 5) cuantificó importancias, destacando PESOS DE LA IES (37 %) y B PESOS SIN TP (10 %), seguidos por la formación práctica y la remuneración docente. Finalmente, un árbol de criterio comparó las valoraciones de 2023 y 2024, validando un avance global de +12,7 puntos porcentuales y señalando el retroceso en “Profesores”. Estos hallazgos sustentan un plan de acción basado en semáforos, validación continua y focalización de indicadores críticos para la mejora de la calidad institucional.

Palabras clave: aprendizaje supervisado, aseguramiento de la calidad, árbol de decisión, Random Forest, evaluación institucional.

ABSTRACTS

The objective of the present study was to design and validate an analytical system for diagnosing compliance with the 43 indicators of the *2024 model* at the “*Instituto Superior Universitario Carlos Cisneros*.” To this end, a pre-existing DataFrame—into which qualitative assessments had already been converted under the fields “peso ideal” and “peso real”—served as the point of departure. From this dataset, a binary label, *cumple* (compliance), was generated based on a threshold ratio of ≥ 0.60 . Using an 80/20 train–test split via `train_test_split`, a global decision tree (maximum depth = 3) was trained. Visualization of this tree revealed operational thresholds for “**PESOS DE LA IES**” (0.009 and 0.012) and “**B PESOS SIN TP**” (0.012 and 0.017), which were expressed as highly interpretable conditional rules. Five-fold stratified cross-validation (StratifiedKFold) yielded a mean accuracy of 97% and a recall of 100%, with only isolated false positives occurring in one fold.

To preserve the identity of each indicator, one-hot encoding was applied, and a more granular decision tree (maximum depth = 5) was constructed. This model achieved 100% accuracy on the sample and uncovered exceptions among teaching-related indicators, such as “*Formación académica en curso y capacitación*.” With **PESOS DE LA IES** (37%) and **B PESOS SIN TP** (10%) emerging as the most influential features—followed by practical training and teaching compensation—a Random Forest comprising 50 trees (maximum depth = 5) was employed to quantify feature importances. Finally, an *árbol de criteria* (criteria tree) confirmed an overall improvement of +12.7 percentage points when comparing the 2023 and 2024 evaluations, while also highlighting a regression in the *Profesores* indicator. These results support a traffic-light signalling system and a validation-based, ongoing action plan that prioritizes critical indicators in order to enhance institutional quality.

Keywords: supervised learning; quality assurance; decision tree; Random Forest; institutional evaluation.



Revised by
Mario N. Salazar
0604069781

Introducción

Es de suma importancia garantizar la excelencia dentro de la educación superior tanto en el ámbito académico como operativo, es por ello, que se ha desarrollado protocolos de evaluación que requieren un monitoreo y mejora continua, abriendo una oportunidad para el uso de instrumentos para la toma de decisiones basada en datos para su sofisticación. En Ecuador, el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES, 2021) ha dispuesto un conjunto específico que evalúa la calidad institucional de una "academia con adelanto de cuatrimestre". Criterios de evaluación 41(CACES, 2021). El Instituto Universitario Carlos Cisneros de Riobamba ha llevado a cabo capacitaciones corporativas alineadas a esta orden, enfocándose en la integración de una unidad dedicada a la garantía de calidad, cuyo objetivo principal es la gestión y el control sistemático del cumplimiento de los indicadores. (ISU CARLOS CISNEROS, 2024).

El objetivo de la investigación es el diagnóstico a través de un modelo analítico usando aprendizaje de máquina. En cumplimiento de los criterios definidos por el CACES se ha planteado a la institución el desarrollo y mejora de procesos. La fuente de datos utilizados proviene de instituciones en procesos de autoevaluación con elementos codificables en verificación para esos puntos de pesos evaluativos elementales confrontados tipo bilitankue KAP. Se incorporarán aspectos cuantitativos y cualitativos en las evaluaciones previas realizadas por la entidad. Estos incluyen pesos ideales y reales. Evaluaciones categoriales transformadas en datos numéricos también forman parte.

Los modelos predictivos son capaces de permitir una evaluación de desempeño más efectivo para el reconocimiento crítico de patrones, impactando de manera integral los procesos de monitoreo y evaluación, y proporcionando pautas poderosas y eficientes para fomentar decisiones institucionales estratégicas impulsadas por la gestión (Shalev-Shwartz & Ben-David, 2020).

Capítulo 1

Generalidades

1.1 Planteamiento del Problema

A lo largo de la historia del ISU Carlos Cisneros, ha tenido problemas sustanciales que le han impedido un rendimiento adecuado en las pruebas de acreditación propuestas por el CACES. Recientes informes al interior de la institución revelan preocupantes problemas en la práctica formativa, compensación docente, y en la evaluación docente con respecto al desempeño de la institución y su imagen externa (ISU Carlos Cisneros, 2024).

La falta de atención a dichos indicadores conlleva algunos riesgos importantes que pueden causar la pérdida de la acreditación institucional, afectando la confianza de los estudiantes, docentes y de entidades externas a la institución. El problema central está en la ausencia de un buen marco analítico robusto y basado en datos que señalen las causas que provocan los incumplimientos considerados críticos.

Pregunta de investigación: ¿ Cuáles son los elementos que afectan al cumplimiento de los estándares educativos del Instituto Superior de Educación Carlos Cisneros y cuál es la forma de optimizar estos elementos usando técnicas de decisión del árbol y Random Forest?

1.2 Justificación de la investigación

Este estudio es importante debido que proporciona al Instituto de Educación Superior Carlos Cisneros una herramienta analítica para evaluar el cumplimiento continuo con los criterios del marco de calidad educativa CACES. Este estudio predictivo permite a la institución identificar áreas prioritarias e implementar soluciones específicas para apoyar una estrategia basada en datos para la mejora continua. (Géron, 2019).

El uso de tecnologías avanzadas en análisis predictivo fortalece la gestión institucional, tarea que este análisis realiza. Desde una perspectiva académica, este trabajo permite un avance importante en la implementación de modelos de aprendizaje automático

dentro de entornos de educación superior, y en consecuencia, da origen a estudios análogos en otras instituciones universitarias del país.

1.3 Objetivos

1.3.1 Objetivo General:

Desarrollar un sistema analítico fundamentado en técnicas de aprendizaje automático, árboles de decisión permitiendo el diagnóstico, explicación y optimización de los estándares educativos establecidos por el CACES en el Instituto Superior Universitario Carlos Cisneros.

1.3.2 Objetivos Específicos:

- Realizar un tratamiento previo de los datos históricos y actuales relacionados con el cumplimiento de los estándares educativos del ISU Carlos Cisneros.
- Validar modelos predictivos que faciliten la evaluación y explicación de los factores determinantes del cumplimiento institucional, fundamentados en los árboles de decisión y Random Forest.
- Desarrollar un marco propositivo con estrategias específicas orientadas hacia la mejora constante y el cumplimiento de los estándares educativos del ISU Carlos Cisneros.

1.3.3 Población y Muestra

Dentro del alcance del estudio, se toman en consideración todos los procesos evaluativos establecidos en el ISU Carlos Cisneros en el curso académico 2023-2024, concretizando a los 42 indicadores establecidos por el CACES. La muestra explícita se conformará tal como se han ido recogiendo durante el periodo bajo estudio, todos los datos cuantitativos y cualitativos generados por las evaluaciones internas llevadas a cabo en este periodo, es decir, la totalidad de los registros institucionales oficiales conformando así un

conjunto completo y representativo de los registros institucionales oficiales que van referidos a los estándares de calidad científica de los 42 indicadores descritos.

Capítulo 2

2 Estado del Arte y la Práctica

2.1 Antecedentes investigativos

2.1.1 Antecedentes

Varios estudios recientes han explorado la implementación de técnicas de aprendizaje automático para abordar diversos desafíos enfocados a la calidad de la educación. Ruiz Martínez (2020), en su trabajo **"El aprendizaje automático en la educación superior como herramienta para optimizar los índices de deserción y detectar los factores que la ocasionan"**, examinó la capacidad que tienen los algoritmos de aprendizaje automático para predecir la deserción estudiantil a nivel universitario. Para este estudio, Ruiz Martínez (2020) optó por la revisión de documentos de manera detallada, en donde recopiló, clasificó y analizó un número considerable de investigaciones que se han llevado a cabo referidas al abandono escolar, otros enfoques y técnicas de análisis de datos de sector educativo.

La gran cantidad de datos conseguidos le llevan a pensar que las técnicas predictivas en su mayoría (árboles de decisión, redes neuronales, bosques aleatorios) pueden ofrecer un manto de mayor comprensión que envuelve a los alumnos que están en situación de riesgo de abandonar la escuela. Así, su trabajo permite proponer que las técnicas de aprendizaje automático pueden prever, con mayor solvencia que los métodos tradicionales, factores que giran en torno a la deserción escolar (rendimientos académicos previos, contexto social y familiar. En el día a día, dichos modelos logran aportar a los orientadores académicos la capacidad de identificar señales prematuras de desinterés que frecuentemente son desatendidas. Convertir los resultados en mensajes ejecutables para las familias y los docentes les permite apoyar la asistencia y mejorar así la retención de estudiantes.

Contrastando con la investigación actual, este trasfondo ofrece evidencia que apoya la utilidad del análisis de datos en el examen de indicadores predictivos de una educación de calidad, ya que Ruiz se centró en la deserción a nivel individual, mientras que el presente estudio evalúa integralmente el cumplimiento institucional de estándares establecidos por el CACES usando árboles de decisión.

El estudio realizado por Gavilanes Simbaña y Rojas Cevallos (2023), titulado **"Predicción del rendimiento académico de los estudiantes en la Universidad Central del Ecuador"**, es otra referencia pertinente. El principal objetivo fue el desarrollo de un modelo predictivo que permite reconocer a los estudiantes universitarios con mayor riesgo de sufrir de bajo rendimiento académico. Se obtuvieron datos sobre la situación académica, socioeconómica y otros aspectos personales de 111 estudiantes a través de una encuesta. La predicción del bajo rendimiento académico se logró con un 85 % de precisión usando árboles de decisión, una herramienta de minería de datos. Los hallazgos muestran que el rendimiento académico está profundamente afectado por la motivación, la financiación y la institución educativa por el tipo de institución. Gavilanes y Rojas concluyeron que estos modelos predictivos permiten realizar intervenciones en etapas tempranas.

Satalaya (2024) realizó una revisión sistemática titulada **"Aprendizaje automático para abordar la abandono escolar: una revisión de los modelos más innovadores"**, en donde tuvo como objetivo analizar en qué medida se utiliza el aprendizaje automático como un mecanismo predictivo de deserción educativa. Empleando la metodología PRISMA, llevó a cabo una revisión de 20 investigaciones actualizadas, detectando algoritmos altamente potentes, incluyendo los árboles de decisión o redes neuronales. Los resultados de la investigación destacaron la habilidad del aprendizaje automático para anticipar la deserción escolar adaptándose a las necesidades específicas de los alumnos por medio de programadas personalizadas. En este caso, la investigación a la que aludimos sirve como

soporte del marco teórico de la presente investigación porque demuestra la mejora importante alcanzada por las técnicas de aprendizaje automático para la evaluación predictiva de los indicadores de la calidad educativa. Esta investigación se destaca por utilizar estos modelos específicamente para evaluar los criterios institucionales del CACES.

Además, la investigación realizada por Marín-Rodríguez et al. (2023) llamado **“Modelo de clasificación para la deserción estudiantil en una universidad pública del Perú”** ha creado un sistema predictivo fundamentado en árboles de decisión para identificar a los alumnos que presentan alto riesgo de abandonar sus estudios universitarios. El método permite demostrar que la precisión del árbol de decisión está exactamente por encima del 80% a través de distintos algoritmos predictivos y la exhibición de los datos históricos o datos del pasado. Cualitativamente, se pone en evidencia que cuanto mayor es la probabilidad de que el estudiante abandone su compromiso académico, menor es la asistencia a clase y la nota media. La recogida de datos en el ámbito de la educación mejora la gestión institucional y potencia la puesta en marcha de proyectos adicionales orientados hacia la prevención, como se indica en la conclusión.

Este caso deja claro que los análisis de datos que aplican aprendizaje automático pueden ser adaptados a procesos orientados al aseguramiento de la calidad de educación, ofreciendo pruebas empíricas adecuadas para esta investigación. La principal diferencia es que la investigación más reciente se centra en la estricta adhesión a los requisitos institucionales del CACES, mientras que los estudios peruanos se centran solo en las tasas de abandono estudiantil activo.

El estudio de Mendoza y Castillo de 2023 finalizó con un **“Análisis comparativo entre los estándares de calidad del Ministerio de Educación y el CACES”**. El proyecto consistió en un análisis de los estándares educativos del CACES, en relación con los niveles de educación básica, usando la comparación entre estándares, dejando a un lado las técnicas

analítico-predictivas típicas de la investigación actual, para acentuar la exactitud y el alcance de los estándares del CACES. El informe sirve como documento de referencia muy completa para las evaluaciones institucionales, enfocándose a partir de los estándares del CACES en el marco del país.

2.2 *Fundamentación Legal Nacional*

La arquitectura normativa que garantiza la excelencia en la educación superior ecuatoriana se basa en la Constitución de la República. El artículo 346 dicta la presencia de una entidad pública que posea autonomía dedicada a la evaluación interna como a la externa con el fin de fomentar la calidad educativa. Esta normativa facilitó la capacitación de entidades especializadas, como el CACES, que realizan procedimientos sistemáticos de acreditación y evaluación (Asamblea Constituyente del Ecuador, 2008).

La Disposición Transitoria Vigésima de la Constitución impone la obligatoriedad que algunas instituciones de educación superior (IES) sean evaluadas y acreditadas en los cinco años subsiguientes a la entrada en vigor del texto constitucional. Este requisito, vigente entre 2014 y 2016, legitimó la decisión de las instituciones que no cumplían con los criterios requeridos y respaldó legalmente las exhaustivas evaluaciones de universidades e institutos.

La LOES fue reformada en el año 2018 para ampliar el marco legal. El artículo 94 del código constitucional establece el Sistema Interinstitucional de Aseguramiento de la Calidad, compuesto por el CES, el CACES y las Instituciones de Calidad. El artículo 93 define el concepto de calidad como la responsabilidad colectiva del Estado, las instituciones y la sociedad, así como el afán inquebrantable de progreso. El artículo 95 obliga al CACES a crear modelos evaluativos que incorporen criterios e indicadores cualitativos y cuantitativos, garantizando así que el objetivo final sea la excelencia académica y no solo la acreditación (Asamblea Nacional del Ecuador, 2018).

Por otro lado, CACES ha emitido disposiciones complementarias que incluyen el Reglamento de Evaluación Externa con Fines de Acreditación. Esta norma contiene los amparos legales de sus decisiones, los ciclos evaluativos y el sistema de evaluación. Revisitar, por ejemplo, la publicación del documento normativo en noviembre de 2023 actualizó los requisitos para la clasificación de títulos sobre la pertinencia y vigencia de las normas aplicables (CACES, 2023).

La institución de aseguramiento de la calidad ha experimentado una evolución progresiva desde el CEAACES (que aún estaba vigente en 2018) hasta el CACES con el objetivo de promover la calidad. En el sentido actual, el Centro de Educación Superior (CES) aporta al sistema a través de la creación de leyes y estándares académicos de parte inferior.

Sin embargo, es responsabilidad del CACES llevar a cabo la evaluación global del cumplimiento de los estándares mediante procedimientos de evaluación externa. El primer y segundo ciclo evaluativos, llevados a cabo de 2014 a 2016 y de 2020 a 2021, respectivamente por el CEAACES y el CACES, constituyen dos procesos evaluativos contemporáneos donde el marco legal está siendo efectivamente aplicado.

Ambos tendieron a ser evaluaciones estructuradas basadas en modelos creados para categorizar las instituciones y sus programas en una escala de desempeño. Estos ciclos dieron como resultado acreditaciones, recomendaciones y en unos pocos casos la suspensión de operabilidad de instituciones que no cumplieron con los requisitos, así como visitas en territorio y análisis documental.

La Constitución integra secciones que incorporan cláusulas fundamentales como el Artículo 350, que define los objetivos de los sistemas de educación superior centrados en los estándares de la productividad académica y científica. El Artículo 351 establece además que los institutos técnicos y tecnológicos también constituyen una parte del sistema nacional de educación.

Además, los Centros Regionales e Internacionales de Educación de Duke también son partes del sistema nacional que, en colaboración con la Universidad de Duke, ofrecen el programa combinado de grado MD-MPH. Por ejemplo, durante la revisión nacional de medio ciclo de 2020 a 2021, se evaluaron los programas prioritarios y se concedieron nuevas acreditaciones o extensiones de validez.

En conjunto, la legislación nacional da un fuerte apoyo normativo a la investigación propuesta, cuyo objetivo es el desarrollo de un modelo analítico de aprendizaje automático fundamentado en árboles de decisión para diagnosticar el cumplimiento de los estándares del CACES.

La propuesta cumple con los requisitos de la actual ley nacional en cuanto a verificabilidad, utilización de métodos técnicamente sólidos y objetivos, para modelos de evaluación que promueven la excelencia en educación superior.

2.3 Ciencia de datos, para la toma de decisiones en marcos evaluativos del CACES

2.3.1 El rol del CACES en la evaluación de indicadores de calidad

La responsabilidad del CACES es garantizar la excelencia de las IES en Ecuador, su propósito principal es diseñar e implementar modelos de evaluación que incorporen parámetros y estándares de tipo cuantitativos y cualitativos, los cuales deben cumplir las instituciones, carreras y programas para obtener la acreditación correspondiente.

En términos generales de estándares de calidad, los modelos establecidos por CACES contemplan criterios cualitativos cuyo uso de evaluación y valoración son aplicables a componentes clave de las IES.

De este modo, la aplicación efectiva de estos modelos de evaluación es el objetivo principal y fundamental, que implica la mejora sostenible y continua, así como el desarrollo a través del impulso concreto de la calidad académica y la persistencia de la transparencia en la gestión institucional. Por otro lado, la exigencia para la creación de criterios precisos

y cuantificables propone la participación activa y fácil identificación de los cauces que requieren mayor potencia e inyección.

2.3.2 Aplicación de modelos predictivos en el marco de evaluación del CACES

La elaboración de modelos predictivos como los Árboles de Decisión y el Bosque Aleatorio o Random Forest tiene una función en la evaluación de la calidad educativa aportando herramientas para la predicción y el análisis del cumplimiento de las normas dispuestas por el CACES; para la toma de decisiones en torno a la detección de patrones complejos y la identificación de relaciones interdependientes entre las variables son deseables. Estos modelos estadísticos tienen un papel importante en este proceso (CACES, 2021).

La aplicación de uno de los algoritmos de modelado de aprendizaje automático permite hacer visible, de forma eficaz y precisa, la calidad del rendimiento académico de una institución educativa, alcanzando este objetivo mediante la evaluación de variables de tipo cuantitativo como la cantidad de docentes con posgrados, la infraestructura educativa y el rendimiento académico, lo que permite a las instituciones evitar deficiencias y realizar acciones de corrección antes de las evaluaciones del CACES(CACES, 2021).

Por otra parte, la utilización de estos modelos en procesos de autoevaluación institucional puede ayudar al incremento de la eficacia y la eficiencia de los planes de desarrollo. La investigación basada en datos fortalece la capacidad de las instituciones para cumplir con los estándares de calidad y abordar adecuadamente las demandas del CACES.

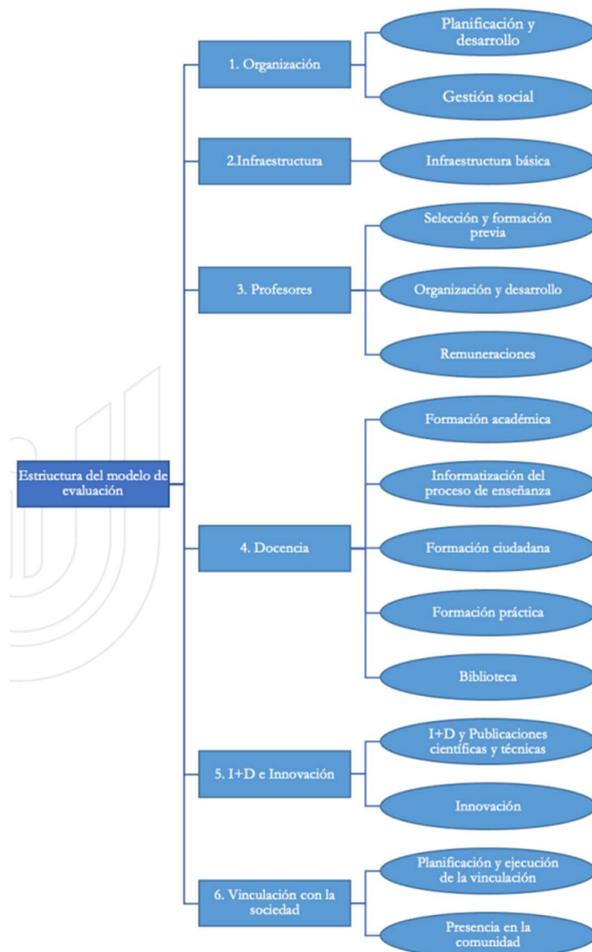
2.3.3 Importancia de la acreditación de los Institutos Superiores Universitarios y la oportunidad del uso de machine learning

Los modelos de evaluación creados por el CACES combinan criterios e indicadores específicos para evaluar el desempeño institucional. Estos modelos tienen como objetivo

asegurar que las instituciones cumplan con los estándares de excelencia en las áreas académica, administrativa y de infraestructura.

Ilustración 1

Estructura del modelo de evaluación



Nota. Extraída del modelo de evaluación externa 2024 del CACES

La página 10 del Modelo de Evaluación Externa 2024 para Institutos Superiores Técnicos y Tecnológicos es la estructura jerárquica del instrumento del CACES. El nodo "Estructura del modelo de evaluación" se encuentra en la parte superior y de él derivan seis ejes rectangulares (criterios).

- Organización
- Infraestructura
- Profesores
- Docencia
- I+D e Innovación
- Vinculación con la sociedad.

Cada uno de estos parámetros se especifica a través de subindicadores ovalados que implementan el estándar: por ejemplo, “Planificación y desarrollo” y “Gestión social” corresponden a la organización, mientras que “Infraestructura básica” se refiere a la infraestructura. Para los profesores, aplican “Formación académica”, “Remuneraciones” y “Organización y desarrollo”.

Este esquema en cascada permite ver la relación padre-hijo entre criterios e indicadores, así como la cobertura integral del modelo en lo que respecta al gobierno institucional, recursos físicos, talento humano, procesos formativos, producción científica y relevancia social. Además, CACES ofrece una representación gráfica clara de los índices estructurales, lo que permite a los equipos de autoevaluación comprender la importancia estructural de cada elemento y adaptar sus pruebas a los niveles de análisis estipulados dentro del marco de evaluación (CACES, 2022, p. 10).

El uso de los métodos de aprendizaje automático recomendados en los procesos de evaluación para la certificación de instituciones de educación superior presenta una oportunidad notable en términos de efectividad y eficiencia en la precisión al tratar con procesos evaluativos. Los criterios evaluativos del CACES relacionados con los datos de un bosque de progresión institucional y árboles de decisión estocásticos permiten el análisis de una gran cantidad de datos institucionales, discerniendo patrones notables y proyectando

tendencias futuras. La capacidad predictiva, por tanto, permite a las universidades identificar áreas problemáticas y corregirlas de forma preventiva. La correcta reutilización del aprendizaje automático proporciona el diseño de estrategias de mejora personalizadas para cada universidad y promueve la cultura de la excelencia y del compromiso institucional con respecto a la calidad educativa.

2.4 Fundamentos de la Minería de Datos y Aprendizaje Automático

2.4.1 Minería de Datos

La minería de datos, un subcampo de la ciencia de datos, busca patrones y asociaciones ocultas en conjuntos de datos masivos. Con el paso de los años esta técnica ha cobrado importancia junto con la disponibilidad de datos por motivos informáticos. En este sentido, Jiménez Galán (2024) menciona que la minería de datos es útil para identificar patrones relevantes y conocimiento en grandes volúmenes de datos, resaltando su función en la tecnología actual como el análisis de datos contemporáneo.

La minería de datos no solo es la etapa final del proceso de Descubrimiento de Conocimiento en Bases de Datos (KDD); también incluye la recolección, limpieza, integración, transformación, minería, interpretación y representación gráfica de la información. En cada uno de esos compases del proceso es esencial que los patrones encontrados sean comprobados en relación a su validez, originalidad, utilidad y entendimiento. Para esto, la minería de datos no tiene en cuenta únicamente los métodos estadísticos más habituales; ciertas afirmaciones exigirán métodos que provendrán de algoritmos basados en inteligencia artificial y del aprendizaje automático que permitan que el análisis del dato y su extrapolación a cifras, que a través de métodos tradicionales resultarían imposibles de resolver, sí puedan ser físicamente accesibles (Jiménez Galán, 2024).

Uriz Martín (2021) en su tesis doctoral acerca de cómo optimizar los algoritmos en minería de datos menciona 'que el preprocesamiento del proceso como un paso muy importante del algoritmo debe ser tenido en consideración. El autor cita así el trabajo de los métodos de selección de funciones y transformación de variables adecuadas, ya que: "La calidad de los datos de entrada es tan importante como los algoritmos para la modificación efectiva del clasificador.

Por otro lado, Basgall (2022) señala acerca de una técnica que se basa en la tecnología de preprocesamiento a escala con tecnología para un enfoque del desbalance de los datos en el Big Data. Ella convirtiendo todo este conocimiento en la crisis humanitaria muestra que: "La minería de datos puede y debe ser sensible a las situaciones de extrema gravedad. Según la tesis, el preprocesamiento debe tener en cuenta el tamaño y el tipo de datos y la disposición de los datos relacionados con el problema para dar resultados pertinentes y significativos.

El empleo en situaciones de crisis humanitaria muestra la versatilidad de la minería de datos en cuanto a la adaptación a situaciones de extrema urgencia (Basgall, 2022). Para lograr resultados que realmente se puedan calificar como útiles, su tesis propone que el preprocesamiento debe considerar no solo la escalabilidad del problema, sino también su naturaleza y la disposición de los datos.

Particularmente en el ámbito de la educación, estos temas son relevantes debido a que las instituciones generan datos estructurados y no estructurados a partir de evaluaciones, cuestionarios, registros administrativos y observaciones cualitativas. Antes de que los datos puedan procesarse automáticamente, deben someterse a procedimientos de limpieza, normalización y transformación. Algunos de los pasos más críticos para asegurarse de que los modelos futuros, como los árboles de decisión o los bosques aleatorios, funcionen de

forma eficiente, son transformar las escalas cualitativas en valores numéricos, derivar variables ficticias y codificar las variables con codificación *one-hot*.

2.4.2 Árboles de Decisión

Los árboles de decisión son muy valorados por su simplicidad interpretativa, así como por su capacidad para manifestar decisiones complejas dentro de marcos lógicos. Según Díaz Martínez y colaboradores (2021), un árbol de decisión es una prueba estadística que proporciona un modelo visual y probabilístico en un marco de decisión, el cual simplifica la evaluación de los resultados, por medio de una serie de observaciones y construcciones lógicas.

En cuanto a la estructura, este tipo de modelo se organiza jerárquicamente alrededor de nodos internos que se dividen por atributos, clases en niveles y juicios de decisión final y con nodos hoja que representan clases en evolución o juicios finales. Los nodos internos se subdividen por criterios de segmentación utilizados en este método como la entropía o el índice de Gini con el fin de aumentar la pureza de cada subgrupo resultante. La entropía, por ejemplo, es una medida de incertidumbre; una reducción de la entropía indica que se ha hecho un progreso en la agrupación de datos (Álvarez Polo, 2023).

La naturaleza no paramétrica de los árboles de decisión, que les permite ajustarse a distribuciones de datos desconocidas sin definir su forma, es una de sus características más distintivas. Debido a ello, los árboles de decisión resultan muy beneficiosos en programas educativos que recolectan información de diferentes fuentes con variados niveles de calidad. Por ejemplo: pueden emplear cuestionarios cualitativos para medir los grados de satisfacción del personal, anticipar el desempeño de los estudiantes y de esta manera anticipar la observancia de las políticas de la institución (Álvarez Polo, 2023).

Jiménez Galán (2024) destaca la gran versatilidad de los árboles de decisión, razonando que, al resolver ambos problemas de clasificación y regresión, su flexibilidad

tiene el potencial de adaptarse a prácticamente cualquier industria. Dado que la variabilidad institucional es un elemento crucial en el marco de referencia del modelo de evaluación, este modelo se adapta a la evaluación de criterios de calidad según diversas combinaciones de indicadores.

Los modelos de aprendizaje automático como este son ideales para no expertos, como administradores institucionales o evaluadores de calidad académica, que necesitan analizar resultados y tomar decisiones basadas en criterios precisos, ya que son fáciles de visualizar e interpretar. Por tanto, los árboles de decisión son herramientas útiles para la comunicación y la investigación en el marco de la evaluación institucional (Jiménez Galán, 2024)

2.4.3 *Random Forest*

El modelo de aprendizaje automático supervisado Random Forest construye múltiples árboles de decisión y agrega pronósticos para crear un modelo en cadena más robusto y generalizable. En contraste con un único árbol, Random Forest construye muchos árboles a partir de subconjuntos aleatorios del conjunto de datos de entrenamiento, fusionando los resultados mediante votación para clasificación o pre-promediado para regresión; esto mejora la precisión del modelo y lo hace más robusto. Jiménez Galán (2024) está de acuerdo en que este método aumenta el poder predictivo y reduce el riesgo de sobreajuste al incrementar la capacidad de ensamblaje de innumerables árboles.

Sus ventajas significativas radican en su capacidad para manejar conjuntos de datos complejos con completitud y simplicidad; sin embargo, puede incluir datos irrelevantes o colineales. la función también es relevante en el sector educativo tomando la afirmación de Jiménez Galán (2024) que este modelo ofrece información cualitativa institucional, satisfacción del estudiante, rendimiento académico, y cumplimiento regulatorio en conjunto

con datos adecuados. su aplicación es práctica, ya que maneja datos directrices y dispares sin disminuir su eficiencia.

Según Álvarez Polo (2023), Random Forest considera la importancia de todas las variables involucradas en el conjunto de datos con respecto a ciertas métricas internas, como la reducción en la pureza y la precisión en la clasificación. Random Forest capta elementos destacados y significativos que son importantes para la formulación de decisiones estratégicas, generando así un modelo claro, lo que se traduce en decisiones basadas en evidencia, es importante entender su importancia en la evaluación cuando el objetivo supera la mera predicción.

Otra de las ventajas que tiene el uso del Random Forest es que plantea un comportamiento adecuado frente a problemas de clasificación de clases desiguales, algo frecuente en tareas dentro del ámbito institucional, dado que la mayoría de los indicadores cumplen con los requisitos marcados, y hay una pequeña proporción que exhibe casos problemáticos. El algoritmo Random Forest permite combinar las predicciones que realizan varios modelos de aprendizaje de tipo automático, esto se realiza a través de un proceso de combinación de distintos resultados con la finalidad de obtener predicciones más precisas y más robustas, ayudando de esta manera a tener una mejor comprensión sobre la distribución de los datos y de las relaciones que existen entre las variables.

A pesar de que la estructura del árbol de decisión pueda tener algunas dudas en su uso, hay que indicar que es un método muy eficiente y útil para mejorar y potenciar la calidad en el ámbito educativo. Random Forest es el siguiente paso posterior a los árboles de decisión individuales. Además de combinar las ventajas de los árboles de decisión, también minimiza sus desventajas.

2.5 Preprocesamiento y Transformación de Datos

2.5.1 Limpieza de datos: valores faltantes e inconsistencias

La limpieza de los datos es un proceso muy necesario para que el modelo predictivo se ejecute de manera efectiva. El objetivo fundamental en este caso es el mantenimiento de la consistencia del conjunto de datos y evitar resultados inesperados durante la ejecución de los modelos. Como dice Gonzales León (2020), la detección de los posibles errores y la realización de las correcciones a tiempo son aspectos básicos para que el análisis no tenga problemas. Algunos de los errores potenciales serían, por ejemplo, la existencia de valores perdidos, registros duplicados y la existencia de discrepancias en el conjunto de datos. Se irá construyendo una base de datos cuidando hasta el extremo los detalles y los registros. Romero-Duque et al. (2023) indican que una base de datos sólida debe tratar esos valores vacíos mediante métodos de imputación.

Los sistemas de aseguramiento de calidad en las IES de educación, como el que implica la evaluación institucional, son sistemas que tienen vacíos en cuanto a la cobertura; la limpieza se convierte en un requisito metodológico para asegurar que existen diferentes algoritmos de análisis, como por ejemplo Random Forest y árboles de decisión.

A menudo, los sistemas modernos de aseguramiento de calidad para instituciones de educación superior, como los relacionados con la evaluación institucional, carecen de datos subyacentes. Esto hace que, en el contexto de la ingeniería, la limpieza de datos de algoritmos de análisis como Random Forest y Árboles de Decisión se convierta en un paso metodológico imprescindible para asegurar que funcionen a partir de conjuntos de datos verificados. Estas técnicas, aunque informadas por alguna deficiencia, requieren cierta coherencia en la información para operar de manera proactiva.

Además, la literatura académica también subraya la necesidad de establecer estándares de calidad de datos para organizaciones responsables de gestionar y recopilar

información educativa. Desde tácticas de entrada de datos, hasta sistemas de comprobación por automatización anticipada, la identificación temprana de irregularidades simplifica la acción proactiva y garantiza la legitimidad de los futuros modelos basados en las interpretaciones de precisión de estos datos para evaluar el cumplimiento de las regulaciones institucionales.

2.5.2 Mapeo de escalas cualitativas a valores numéricos

Para variables de naturaleza de naturaleza categórica, el cambio de escalas jerárquicas a valores numéricos permite la aplicación de distintos modelos cuantitativos. En este sentido, Arias González (2021) asegura que una operacionalización adecuada se basa en equivalencias numéricas que reflejen de forma proporcional las propiedades semánticas de las categorías. Al usar algoritmos que gestionan magnitudes, como la regresión lineal o los ejes de decisión, dichos valores implican estar espaciados de forma proporcional.

Para algoritmos sensibles a la magnitud, como modelos de regresión lineal, árboles de decisión y otros, el impacto de cada variable a lo largo del modelo debe ser constante. Como sugieren Pérez Peña et al. (2023), el mapeo juega una función especial en la aplicación de métodos mixtos, porque vuelve los datos cualitativos a cuasi números, lo que no sólo facilita su espacio con diferentes modelos estadísticos, sino que les regala un fenómeno cualitativo sin romper la solidez del modelo.

2.5.3 Escalado y normalización de variables

La escalabilidad y la normalización son dos técnicas que se utilizan para mejorar la eficiencia de los algoritmos de autoaprendizaje al dar rangos aproximados a los datos numéricos. En particular, los datos estandarizados tienen una media de cero y una desviación estándar de uno. Una de las técnicas de normalización más comunes es la normalización mínima-máxima, que define la transformación de los valores de una variable a un rango dado, frecuentemente entre 0 y 1.

Este método permite la comparación de diferentes variables con escalas dispares. De acuerdo a su investigación, Rodríguez González y Ugalde Saborío (2021) proponen que, con una adecuada selección de estas técnicas de análisis de datos, la precisión y la efectividad de algunos modelos complejos, tales como las redes neuronales artificiales y los árboles de toma de decisiones, pueden mejorar considerablemente.

Rivera Díaz (2024) desarrolla un estudio sobre los deslizamientos en Guayaquil y utiliza este concepto, apoyándose en sus argumentos, pues la normalización permite comparar variables de diferentes unidades. Este procedimiento se vuelve necesario cuando empleamos indicadores de distinta procedencia, tal como en valoraciones del rendimiento institucional, en cuyo caso es común encontrar porcentajes, conteos y categorías de valores modificados. En estos casos, la estabilidad numérica del modelo se basa en que los datos adolecen de una escala uniforme.

Desde el punto de vista técnico, la normalización de los datos permite, entre otros elementos, hacer frente a la pertenencia de variables con escalas más elevadas, sobre todo en algoritmos que se apoyan en la distancia o en funciones de separación de datos erróneos. Esta adecuación previa propicia, además, modelos repetitivos más rápidos y una simplificación de la representación de datos.

2.5.4 *One-hot encoding para variables categóricas*

Para el tratamiento de variables nominales se suele utilizar la técnica *one-hot encoding*. El procedimiento implica convertir cada categoría en una variable binaria autónoma, que asume el valor de 1 si la observación es de esa categoría, y de 0 si no es así. La CNMC 2020 destaca que, si se convierten variables categóricas a numéricas, se podrán usar datos cualitativos en análisis estadísticos sin crear jerarquías falsas.

Sin la ejecución de este procedimiento, varios modelos pensarán que una categoría es mayor que otra solo porque se le asignó un número, aunque no hay un orden real entre

las variables. Transformar de la forma correcta entrega a los modelos como árboles de decisión y Random Forest predicciones más certeras y exactas. Una mala codificación puede hacer que el modelo descubra relaciones entre clases que no existen. En los sistemas de aseguramiento de la calidad institucional, podría permitir añadir variables del tipo de carrera, la duración y forma de titulación sin prejuicios.

Desde su perspectiva computacional, el uso de la codificación *one-shot* genera una mayor dimensionalidad en el conjunto de datos, el eventual efecto de lo cual puede ser negativo en algoritmos con muchas variables categóricas. Sin embargo, cuando se hace correctamente, mejora la capacidad de interpretar los coeficientes, analizar la importancia de las variables y se integran sin problemas en las estructuras de matriz requeridas por la mayoría de las bibliotecas de Machine Learning hoy en día.

2.5.5 *Uso de variables dummy*

Las variables *dummy* son un procedimiento a través del cual se presentan categorías para los modelos de aprendizaje automático y los estadísticos. Las mismas traducen cada categoría en una variable binaria con valores de 0 o 1 dependiendo de que la característica esté considerada en el caso de estudio. Esta estrategia es especialmente útil si tenemos la intención de usar variables nominales en modelos lineales o en modelos de árbol, sin necesidad de secuencias jerárquicas definidas. Según lo que dictaba la CNMC (2020), la implementación de esta serie de variables *dummy* evita los errores de interpretación de convertir atributos cualitativos en datos numéricos de entrada.

Las variables *dummy* colaboran con la supresión de la colinealidad de las variables utilizadas en los modelos lineales, dado que solo hay una categoría de referencia a eliminar en lugar de la codificación de tipo dummy que genera tantas columnas como categorías; en evaluación institucional los aspectos pueden ser el proceso de acreditación, el impacto del sector educativo en los indicadores de calidad, etcétera.

En modelos como el Random Forest, las variables *dummy* no solo colaboran con la introducción de datos categóricos sino también, colaboran en el establecer la relevancia de forma relativa de los atributos.

2.6 Árboles de Decisión: Construcción y Aplicación

2.6.1 Algoritmo de árbol de decisión

También se puede considerar el árbol de decisiones como un algoritmo de aprendizaje supervisado que clasifica las instancias de los datos y predice comportamientos en torno a las mismas a partir de una serie de condiciones de causa-efecto que se disponen en el árbol mediante un orden temporal. En palabras de Álvarez Polo (2023), algoritmos como ID3 y CART utilizan medidas como la entropía o el índice de Gini para encontrar la mejor característica para dividir los datos en cada nodo. Estas medidas evalúan la pureza de las clases dentro de un subconjunto, lo cual es absolutamente crucial para reducir la impureza en las ramas del árbol. Este tipo de modelo define reglas claras que dicen si se cumplen o no los indicadores de calidad en los entornos institucionales. Por ejemplo, se puede determinar que la dotación de profesores con posgrado por debajo de un cierto límite no cumple el criterio. Una de las fortalezas del modelo es que este tipo de interpretabilidad permite justificar las decisiones y su trazabilidad ante auditores o comités de evaluación. Los algoritmos de árboles de decisión se caracterizan principalmente por sus capacidades para trabajar con datos numéricos y categóricos y por ser muy capaces de identificar esos registros atípicos que trastocarían el modelo. Estas características los convierten en modelos muy adecuados para análisis exploratorios y para aquellas places donde la interpretabilidad del modelo reviste tanta importancia como su precisión.

2.6.1.1 Estructura de un árbol de decisión

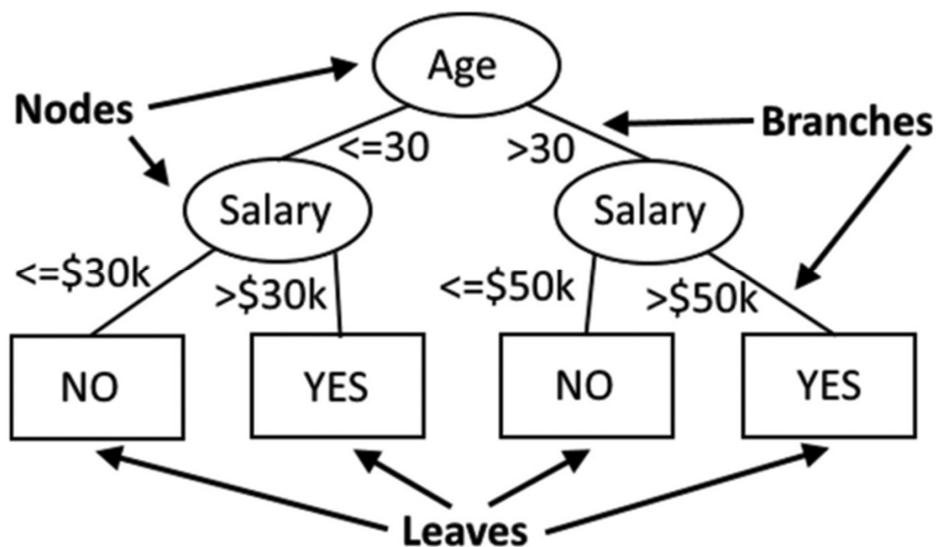
Un árbol de decisión tiene su origen en el nodo raíz haciendo el análisis de ese atributo que mayor capacidad tiene para separar el conjunto de datos: capacidad que se mide

mediante una métrica de elección del atributo, como la ganancia de información (ID3 y C4.5) o el índice Gini (CART). El nodo raíz del ejemplo ilustrado en la Figura 1 examina la variable Age: las instancias se bifurcan en dos ramas dependiendo de si la edad es inferior o igual a 30 años o superior a 30. En consecuencia, cada rama simboliza el resultado potencial del examen aplicado en el nodo (Yates, Islam & Gao, 2019).

El procedimiento de particionamiento avanza sobre un nuevo atributo cuando una rama posee otro nodo interno, igualmente ovalado. En la representación gráfica, el algoritmo solicita nuevamente la rama " ≤ 30 ", esta vez por Salary; de manera análoga, la rama " > 30 " conduce a otro nodo Salary. El secreto radica en esta recursividad: la métrica de separación sobre los registros que llegan a un nodo se recalcula en cada paso. De este modo, el árbol "afina" las preguntas hasta que la impureza de clase baja se encuentre por debajo de un umbral (Fayyad & Irani, 1992).

Ilustración 2

Árbol de decisión Salary



Nota: Árbol de decisión extraído de Yates et al. (2019)

El estudio concluye en los nodos de hoja (rectángulos). Una hoja contiene la clase o resolución que se otorga a cualquier etapa que continúe con ese camino. El esquema categoriza el caso como "NO" si $(Age \leq 30)$ y $(Salary \leq 30 \text{ k\$})$; en caso contrario, el sistema responde "YES". Estas hojas tienen la capacidad de almacenar un valor numérico medio de los registros recibidos, una probabilidad o una etiqueta categórica en regresión (Díaz Martínez et al., 2021).

Siempre se inicia la lectura del árbol desde el nodo raíz y se progresa respetando la condición de la instancia en cada nodo. De este modo, la estructura convierte un complejo conjunto de reglas en una secuencia lineal de preguntas lógicas: cada camino raíz-hoja es una regla IF... AFTER... Esta imagen facilita la comprensión de especialistas y usuarios no cualificados, una ventaja subrayada como uno de los rasgos distintivos de los árboles en relación a modelos de "caja negra" (Yates et al., 2019).

Finalmente, tanto el tamaño como la profundidad del árbol se determinan por las distintas características y valores del conjunto de datos, así como por los criterios de parada (número mínimo de registros por nodo, profundidad máxima, podas subsiguientes). Mientras que un árbol extremadamente superficial puede perder predictividad, un árbol demasiado profundo puede recordar el ruido de entrenamiento (*overfitting*). Los métodos normales para abordar este equilibrio entre sesgo y varianza incluyen validación cruzada a lo largo de nodos (Yates et al., 2019).

2.6.2 Evaluación de indicadores de calidad

El modelo de árbol de decisión al igual que varios modelos de aprendizaje automático son muy beneficiosos para determinar el valor de diversos indicadores de calidad, estas herramientas facilitan la simulación y análisis del cumplimiento de los rigurosos criterios que la institución ha fijado mediante regulaciones precisas y claras. Luna y Marín (2018) emplearon este novedoso método para elaborar rutinas de mantenimiento

fundamentadas en indicadores clave de rendimiento (KPI), demostrando su eficacia en circunstancias donde estos indicadores requieren un análisis ordenado y transparente.

El enfoque del modelado basado en decisiones permite en la posibilidad de que cada nodo del árbol represente un estado de evaluación determinado, bien sea un score o un umbral indicando el cumplimiento. Esta lógica contribuye a la verificación automatizada del cumplimiento de la regulación establecida por organizaciones como el CACES.

Este tipo de evaluación también favorece el avance en el sistema dentro de la organización, y puede ser aplicada a mediciones anuales o por departamentos. Permitiendo así encontrar con más facilidad los factores que influyen de una forma más significativa, el cual resulta importante para la planificación del comportamiento en referencia al fortalecimiento de las debilidades que se evidencian en los indicadores institucionales.

2.6.3 Modelos agregados vs. modelos granulares

Las herramientas específicas para la toma de decisiones son muy aprovechables para la evaluación de indicadores de áreas particulares de calidad. Estas herramientas le permiten simular y evaluar el cumplimiento de criterios institucionales designados a partir de regulaciones bien definidas y no ambiguas. Luna y Marín (2018) hicieron uso de esta idea rompedora para generar rutinas de mantenimiento mediante la utilización de indicadores clave de rendimiento (KPI), evidenciando su eficiencia en escenarios en los que este tipo de métricas requieren una revisión cuidada y transparente.

El modelado sustentado en la adopción de decisiones permite que cada nodo dentro de un árbol se vea como una capa evaluativa diferente según un umbral de puntuación mínima o aún una capa de cumplimiento. El modelado granular, por su parte, aborda cada señal por separado lo que ayuda, entonces, a la identificación de patrones para el incumplimiento concreto distinto del diseño de estrategias atendiendo al incumplimiento de cada uno de los elementos en la evaluación de desempeño.

Las evaluaciones pueden ser aprovechadas en el seno de unas comparativas en la medida que se querían aplicar comparaciones interdepartamentales o intradepartamentales (large annuals) que son reveladoras en relación a una visión prospectiva para la autoevaluación, orientada a la mejora continua. Ayuda a identificar factores que influyen más directamente en la puntuación final, de modo que se puedan implementar acciones para abordar las debilidades determinadas a partir de los puntos de referencia institucionales.

2.7 Evaluación de Modelos Predictivos

2.7.1 Validación cruzada

La validación cruzada ayuda a realizar una evaluación integral sobre el rendimiento del modelo en el aprendizaje sistemático, ya que es el caso cuando el tamaño volumétrico de los datos plantea un desafío para reservar un único subconjunto fijo para la prueba. Como destaca Basgall (2022), el método de k-pliegues es el más común. Aquí, el conjunto de datos se divide en k porciones de igual tamaño, el modelo se entrena k veces en k-1 secciones y verifica la sección que cambia en cada ciclo. Se obtiene una estimación de rendimiento más estable y menos sesgada que la de una única partición de entrenamiento y prueba promediando las métricas obtenidas en las k iteraciones.

Álvarez Polo (2023) muestra cómo la validación cruzada constante genera resultados precisos en métricas como exactitud, recuerdo, el puntaje F1 y precisión. Cada pliego considera los resultados y marca cada una de las métricas, de manera que no solo permite observar el rendimiento medio del modelo, sino también su variabilidad. Esto es particularmente relevante para el caso de modelos de árbol de decisión o Random Forest, donde la consistencia en relación con la perforación de muestras es crítica para la estabilidad del modelo. Ultra técnico, la validación cruzada permite personalizar la sintonización de hiperparámetros ya sea con búsqueda en cuadrícula o búsqueda aleatoria, lo que representa un impulso significativo en el rendimiento del modelo. El procedimiento garantiza que la

adaptación se realice utilizando un conjunto rico y diverso de escenarios para validación en lugar de un pequeño subconjunto homogéneo, mejorando así enormemente la confiabilidad y calidad del modelo con respecto a predicciones complejas del mundo real.

2.7.2 Métricas: *accuracy*, *precisión*, *recall* y *F1-score*

La **precisión** permite conocer el número de verdaderos positivos de todos los positivos que el modelo ha categorizado, es decir, da a entender qué tan confiables son las afirmaciones de que las predicciones de nuestro modelo de casos positivos son confiables. Esta medida es de gran importancia cuando el costo de tener falsas alarmas es alto, como en los sistemas de diagnóstico, pues una clasificación errónea puede dar lugar a una respuesta innecesaria. Zapeta et al. (2022) hacen énfasis en la importancia de dicha métrica para disminuir la posibilidad de falsas alarmas.

El **recall** (exhaustividad o sensibilidad) es el porcentaje de verdaderos positivos reconocidos correctamente de todos los verdaderos positivos, dicho de otra manera, mide la habilidad que posee el modelo para descubrir todos los casos relevantes y resulta especialmente significativo cuando se intenta evitar un caso positivo que no nos conviene perder, como la manera de detectar todos los incumplimientos en un sistema regulatorio. Un modelo puede pasar por alto contextos relevantes cuando la cobertura es baja.

El **F1-score** se calcula como el promedio armónico entre recall y precisión. Ofrece un valor balanceado entre las dos métricas. Es favorable que haya alguna clase de compensación entre cómo el modelo identifica a los verdaderos positivos y su tasa de aciertos. Para Álvarez Polo (2023), en escenarios con clases desbalanceadas, esta métrica es interesante porque si se intenta optimizar una única métrica, el resultado obtenido no refleja con exactitud el desempeño global.

2.7.3 Curvas ROC y análisis de sensibilidad

Las curvas de recepción operativa, denominadas como ROC, son gráficos que permiten trabajar la evaluación del trabajo de distintos modelos de clasificación binaria. A lo que nos referimos es a la relación entre la tasa de verdaderos positivos (sensibilidad) y la tasa de falsos positivos (1 - especificidad) para diferentes umbrales. Martínez-Pérez y Pérez-Martín (2023) indican que las curvas de recepción operativa muestran el funcionamiento del modelo ante las distintas situaciones de decisión y ayudan a encontrar el punto que ideal que compense entre la sensibilidad y la exactitud.

Desde un enfoque técnico, el área bajo la curva ROC (AUC) es un indicador que muestra cómo de bien puede el modelo diferenciar entre las clases. Si bien los valores cercanos a 1 muestran una gran capacidad para diferenciar entre clases, un AUC de 0,5 muestra que el modelo no mejora aleatoriamente. Es decir, esta medida es muy útil para evaluar clasificadores sin depender de la elección de un único umbral.

El análisis de sensibilidad mejora las curvas ROC al investigar cómo se altera el desempeño del modelo al cambiar los valores límite empleados para la clasificación. Esto es de relevancia en contextos en donde se crean grupos de productividad institucional considerando un umbral fijo, ya sea “cumple” o “no cumple”. Esta estrategia facilita el hallazgo de limitantes operativos con mayor precisión y soporte estadístico.

2.8 Importancia de Variables e Indicadores

2.8.1 Cálculo de importancia de características

Permite establecer cuáles de las variables impactan en mayor medida la predicción en modelos como en los árboles de decisión y el Random Forest a partir del cálculo de la importancia de las características.

Álvarez Polo (2023) señala que una de las técnicas más utilizadas es la reducción de impurezas, que evalúa en qué medida cada variable contribuye a mejorar la partición en los

nodos del árbol. Esta significancia se acumula en todo el modelo para generar clasificaciones que orientan la comprensión del fenómeno investigado.

El uso de una característica en todos los árboles determina la reducción de impurezas promedio utilizada para calcular esta medida en los modelos de Bosques Aleatorios. Este enfoque proporciona una estimación más precisa que el uso de árboles de decisión individuales, ya que reduce la variabilidad de los datos y refuerza las interacciones entre diferentes factores. Los analistas de decisiones pueden analizar los resultados del estudio utilizando gráficos y diagramas que representen múltiples atributos de manera cohesiva (Álvarez Polo, 2023).

Desde un punto de vista computacional, calcular la importancia no solo ayuda a aclarar el modelo, sino que también permite utilizar algunas técnicas de selección de características. Dentro de conjuntos de datos de alta dimensión, este enfoque es útil porque aumenta la eficiencia computacional al mismo tiempo que disminuye el ruido del modelo. En contextos administrativos o de diseño instruccional, este enfoque ayuda a identificar qué variables requieren un examen minucioso debido a su participación directa en los resultados del rendimiento.

2.8.2 *Indicadores mediante variables dummy*

Los modelos lineales y los árboles de decisión, por su naturaleza, hacen uso de las variables dummy no solo en la etapa de preprocesamiento, sino que sus aportes resultan invaluable en el momento de interpretar los resultados del modelo. Las variables dummy, si se definen y estructuran adecuadamente, permiten analizar la contribución de un componente en una situación particular con un control granular. Esto asegura que todas las otras variables relevantes permanecen inalteradas y controladas en relación a la variable de interés. Esto ayuda a aclarar los efectos resultantes basados en la naturaleza del estudio o la disciplina académica involucrada con alta precisión. En el caso de modelos de Random

Forest, donde no se utilizan los coeficientes en regresiones para construir el modelo, se considera sustancia en base a la frecuencia y en la ganancia de impureza en los nodos del modelo.

Desde otra óptica, estos factores posibilitan diagnósticos más precisos por el análisis de sensibilidad a variables discretas. También permiten interpretación post-hoc a través de SHAP que calculan el efecto de cada variable dummy sobre una predicción particular. Esto contribuye a clarificar la explicación de sistemas predictivos o auditoría algorítmica.

2.8.3 Comparación de relevancia: enfoque agregado y granular

Entender cómo se interrelacionan las variables bajo la figura general del modelo depende de la comparación entre modelos agregados y granulares. La FAO (1996) afirma que, aunque los datos combinados pueden enmascarar algunas variaciones importantes a nivel de indicador individual, también proporcionan una impresión significativamente excesivamente optimista del desempeño. Si se necesita comprender los detalles de los factores explicativos en los resultados predictivos del sistema, tal falta de claridad es perturbadora. Esta claridad reducida se vuelve problemática cuando se necesita entender los factores que influyen en los resultados de un modelo predictivo.

Por el contrario, los métodos granulares se centran en analizar cada variable por separado, lo que facilita la identificación de las más impactantes. Álvarez Polo (2023) muestra que, en los procesos de toma de decisiones, las medidas de importancia de las características ayudan a distinguir algunas contribuciones al emplear un árbol de decisión o un bosque aleatorio. La capacidad de rastrear el modelo mejora porque este nivel de detalle es posible para que se puedan realizar auditorías y ajustes precisos.

El enfoque de la discusión técnica que describe uno u otro enfoque necesita estar basado en el propósito del análisis. Mientras que los granulares son imprescindibles para acciones correctivas y diseño de políticas específicas, los modelos agregados son

convenientes para reportes ejecutivos y visualización simplificada. Una estrategia combinada que use ambos niveles deja en claro la precisión técnica con la utilidad práctica.

2.9 Interpretación de la Importancia de Variables y Análisis de Contribución de Indicadores

Uno de los beneficios más notables de los modelos de árboles de decisión y bosques aleatorios en el aprendizaje automático es su capacidad para investigar y establecer la relevancia relativa que posee cada variable en el proceso de predicción, a veces llamada "evaluación de la relevancia de las características". Este tipo de estudio apoyado hace posible identificar con mucha más claridad los elementos que contribuyen positiva o negativamente al resultado del modelo. De esta manera, proporciona una herramienta favorable para la toma de decisiones y la planificación que se basa en modelos concretos derivados de información confiable.

Como afirma Álvarez Polo (2023), la evaluación de la importancia de las variables en los modelos de los árboles se realiza a menudo utilizando una técnica conocida como alivio, que calcula la disminución de impureza por característica, o utilizando el índice de Gini dentro del ámbito del árbol. Este enfoque genera una lista clasificada de las variables, ordenadas en lo que respecta a la relación con el resultado del modelo, ayudando así a elucidar componentes más fundamentales del fenómeno bajo consideración. En el caso de las instituciones educativas, esta habilidad permite la selección de los indicadores más relevantes para alcanzar los estándares de calidad.

El análisis gráfico y comparativo de características es posible al formar un array de datos lineal que contiene las variables más importantes y sus valores (Álvarez Polo, 2023). Esto resulta especialmente útil en las instituciones donde la información ha de ser comunicada de forma compacta a ejecutivos y partes interesadas no técnicas. Esta

representación gráfica es fundamental para los procesos correctivos interpretativos de los modelos predictivos orientados al avance institucional.

Valorar la relevancia de diferentes indicadores ayudará a contrabalancear el rendimiento de los modelos agregados que ponderan promedios o criterios de resumen para con los modelos a nivel índice, que tratan a cada indicador de forma granular. La forma de índices compuestos puede resultar útil para tener una visión general, aunque la FAO (1996) sostiene que a menudo ocultan información importante necesaria para una investigación comprensiva, mientras que los modelos a nivel índice permiten disponer de una trazabilidad amplia en los resultados, lo que a su vez da lugar a corregir de manera más exhaustiva. Esta forma de afrontar los modelos resulta muy interesante en determinados ámbitos de aseguramiento de la calidad en los servicios educativos. En este sentido, un análisis diagnóstico institucional exhaustivo exige simultáneamente perspectivas macroscópicas y granuladas a la vez. Así, mientras que el modelo agregado permite calificaciones generalizadas sobre lo bueno, el modelo a nivel índice puede ofrecer síntomas concretos que deberían ser atendidos a la voz de ya, permitiendo que la gestión sea más contundente y, por lo tanto, más óptima.

En términos generales, el análisis de la relevancia de los factores, además de optimizar la interpretación de los modelos de predicción, propicia la claridad y la mejora del proceso en la toma de decisiones dentro del marco educativo. Por su capacidad para transformar algoritmos estadísticas complejas en evidencias claras y comprensibles, dicha herramienta constituye otro de los elementos esenciales del análisis institucional contemporáneo.

Capítulo 3

3 Diseño Metodológico

Se utilizó algoritmos supervisado de machine learning para modelar e investigar los indicadores del modelo CACES 2024 en la toma de decisiones, por la naturaleza de los datos obtenidos el estudio fue de carácter cuantitativo y explicativo-predictivo. Con la meticulosidad requerida y a la vez siguiendo las mejores prácticas en Knowledge Discovery Database (KDD) (Fayyad, Piatetsky y Shapiro, 1996), el procedimiento de trabajo de un patrón preconstitutivo se fue realizando a través de los objetivos del rendimiento y la presentación de resultados como, por ejemplo, la selección de variables relevantes, la estratificación de los datos, la categorización de patrones, las métricas de evaluación y, finalmente, la representación gráfica de las reglas descubiertas. En este estudio, solo se utilizó Python 3.10 y la conocida biblioteca de aprendizaje automático scikit-learn para garantizar la reproducibilidad y la transparencia en los experimentos científicos.

3.1 Enfoque de la Investigación

Este estudio ha seguido un enfoque cuantitativo explicativo-predictivo donde el objetivo principal es detectar los patrones que clasificarían cada indicador institucional como “cumple” o “no cumple” siguiendo los lineamiento operacionales de CACES. Se da por hecho que la combinación de dos variables, el peso asignado y el peso real observados del indicador, contienen la información necesaria para poder construir modelos interpretables y robustos.

El desarrollo se realizó en su totalidad en Google Colab, que permite la integración con GitHub para el control de versiones y la compartición de cuadernos. Además de tener acceso gratuito a GPU ligeras que aceleran la ejecución de validaciones cruzadas y ensamblajes, este entorno permite la colaboración remota y ofrece la ventaja de una trazabilidad completa (registro automático de versiones de bibliotecas y dependencias).

3.2 Origen de datos y preprocesamiento

Dos conjuntos de datos primarios constituyen la base del estudio:

3.2.1 Autoevaluación 2024:

Se construyó un DataFrame que resume los 42 indicadores definidos por CACES para el año 2024. Cada fila tiene tres columnas principales que corresponden a un indicador:

- INDICADORES: identificador alfanumérico del indicador.
- PESO_SIN_TP: valor porcentual perfecto dado siguiendo el Modelo CACES.
- PESO_DE_LA_IES: porcentaje real adquirido siguiendo la autoevaluación institucional.

A partir de estos campos se elaboraron dos variables derivadas: “**ratio_cumplimiento**”, determinada por la relación entre el peso real y el peso ideal. “**cumple**”, etiqueta binaria que valora 1 si $\text{ratio_cumplimiento} \geq 0,60$ y 0 en caso contrario, en base al umbral aconsejado en la Guía Metodológica CACES 2022. La comprobación de tipos, la identificación, la eliminación de duplicados y el tratamiento de posibles valores atípicos formaron parte del preprocesamiento. Utilizando pandas y NumPy, todas estas actividades se registraron en celdas Colab con comentarios que seguían el progreso del proceso.

3.2.2 Series históricas 2023-2024

Se combinó otro DataFrame con valoraciones de criterios institucionales en 2023 y 2024 para la comparación interanual. Además de las columnas **Valuation_2023** y **Valuation_2024** (codificadas primero como cadenas con "%"), se incluyó el peso relativo de cada criterio (*Weight_model*). Una vez convertidas las cadenas a valores decimales, se calculó una nueva columna delta igual a $\text{Valoración_2024} - \text{Valoración_2023}$. De este modo se preparó el conjunto de datos de entrada para la modelización comparativa

3.3 Entorno de desarrollo y librerías

Su progresión, documentación y amplia aceptación en la elección guiada de instrumentos dentro de la academia. La ilustración 3 presentada a continuación contiene todas las librerías usadas en esta investigación:

Ilustración 3

Librerías usadas para la investigación

```
import pandas as pd          # Lectura/escritura de datos y manejo de DataFrames
import numpy as np          # Operaciones numéricas y álgebra lineal
import matplotlib.pyplot as plt # Creación y personalización de gráficos 2D
import seaborn as sns       # Visualizaciones estadísticas avanzadas (barras, calor, etc.)

# ----- Modelado, validación y evaluación -----
from sklearn.model_selection import (
    train_test_split, # División entrenamiento-prueba
    KFold,             # Validación cruzada K-Fold simple
    StratifiedKFold,  # K-Fold estratificado (mantiene proporción de clases)
    cross_val_score,  # Métrica única en CV
    cross_validate    # Múltiples métricas en CV
)

from sklearn.tree import DecisionTreeClassifier, plot_tree
# DecisionTreeClassifier: construcción de árboles de decisión
# plot_tree: visualización de la estructura del árbol

from sklearn.ensemble import RandomForestClassifier
# RandomForestClassifier: ensamble de árboles para mayor robustez

from sklearn.metrics import (
    accuracy_score, # Exactitud
    precision_score, # Precisión
    recall_score,   # Sensibilidad
    f1_score,       # F1-score
    roc_auc_score,  # Área bajo la curva ROC
    classification_report # Resumen de métricas por clase
)
```

La tabla 1 a continuación resume los códigos usados según su propósito:

Tabla 1

Códigos usados según su propósito

Propósito	Librería	Componentes clave
Manipulación de datos	pandas, NumPy	DataFrame, Series, operaciones vectorizadas, read_excel, to_excel, funciones de álgebra lineal

Visualización exploratoria	matplotlib.pyplot, seaborn	plot, barh, heatmap, personalización de ejes y leyendas
Modelado y validación	scikit-learn 1.4	DecisionTreeClassifier, RandomForestClassifier, train_test_split, KFold, StratifiedKFold, cross_validate, plot tree
Cálculo de métricas	scikit-learn.metrics	accuracy_score, precision_score, recall_score, f1_score, roc_auc_score, classification_report
Persistencia de artefactos	openpyxl, joblib	Serialización de modelos (dump/load), exportación de DataFrames enriquecidos

El análisis y modelado de datos se realizó con las siguientes bibliotecas de Python:

- **pandas (import pandas as pd):** lectura y escritura de archivos Excel, manejo de estructuras tipo DataFrame para limpieza, transformación y resumen de datos. Operaciones numéricas avanzadas y manejo de arrays, esenciales para cálculos de métricas y manipulación de vectores.
- **matplotlib.pyplot (exportar matplotlib.pyplot en plt):** creación de gráficos en dos dimensiones, cuentos como la representación de árboles de decisión y comparaciones históricas.
- **seaborn (import seaborn as sns):** desarrollo de gráficos de barras, gráficos de calor y otras representaciones estadísticas que simplifican la lectura de hallazgos exploratorios.

Se utilizaron componentes de Scikit-learn para el modelado, validación y valoración de los algoritmos.

- **train_test_split:** segmentación de la información reproducible en conjuntos de entrenamiento y pruebas.
- **KFold y StratifiedKFold:** Activación de métodos de validación cruzada estratificada y simple.

- **cross_val_score** y **cross_validate**: cálculo de métricas de rendimiento en cada etapa de validación cruzada.
- **DecisionTreeClassifier** and **plot_tree**: construcción y presentación de árboles de decisión.
- **RandomForestClassifier**: conjunto de árboles para evaluar la relevancia de variables.
- **accuracy_score**, **precision_score**, **recall_score**, **f1_score**, **roc_auc_score** y **classification_report**: cálculo y resumen de métricas de clasificación para evaluar la eficacia de los modelos.

3.4 Flujo metodológico general

El proceso de modelado se organizó en cinco etapas principales, aplicadas de manera recurrente para cada uno de los cuatro ejes de trabajo:

- **Elección y desarrollo de variables**: definición de matrices de entrada (X) con las variables predictoras de cada bloque y vectores objetivo binarios (y) que garantizan la coherencia con el umbral de cumplimiento.
- **División de los datos**: El árbol de decisión global utilizó un hold-out (80% entrenamiento / 20% prueba); para los demás modelos se configuraron esquemas de validación cruzada K-Fold o estratificada en función del grado de desequilibrio.
- **Entrenamiento de los modelos**: Se utilizaron hiperparámetros propuestos para afinar cada clasificador se mantuvo la interpretabilidad y se evitó el sobreajuste mediante una profundidad controlada (**max_depth**). En el caso de Random Forest, un número moderado de estimadores (50 árboles).

- **Métricas y evaluación:** Se obtuvieron valores adecuados de exactitud, precisión, recuperación, F1 y ROC-AUC; se hizo hincapié en la recuperación y la F1 en caso de desequilibrio de clases. Los resultados se resumieron en tablas y gráficos.
- **Visualización e interpretación:** Diagramas de árbol (`plot_tree`) con anotaciones sobre la pureza de Gini y los umbrales. Gráficos de barras para las variaciones interanuales o la importancia de las variables.

Los conjuntos de datos y modelos mejorados se conservaron finalmente para su validación y referencia externas.

3.5 Procedimientos de modelado

3.5.1 *Árbol de decisión global*

Este bloque crea un clasificador que depende únicamente de los dos factores de ponderación:

- **Definición de cumple**

La columna `cumplimiento_ratio` se calcula internamente y se convierte en la etiqueta `cumple` según el umbral 0.60.

- **Partición hold-out**

Con `train_test_split(test_size=0.20, random_state=42)`, se separan aleatoriamente 7 observaciones (20 %) para prueba externa, dejando 32 para ajuste.

- **Entrenamiento**

Se crea un árbol de decisión con `profundidad_máxima = 3`; su semilla consistente garantiza resultados repetibles. El límite de profundidad ayuda a extraer reglas obvias y a reducir el sobreajuste.

- **Validaciones de los modelos**

- **Validación K-Fold:** cinco particiones barajadas (`KFold(n_splits=5, shuffle=True, random_state=42)`) para obtener la media y la varianza de la precisión
- **Validación estratificada:** frente al desequilibrio de la clase "No cumple", uso de `StratifiedKFold(5)` y `cross_validate` para evaluar recall y F1.

- **Reentrenamiento final**

Una vez confirmada la solidez del modelo (rendimiento medio ≈ 0.98), se ajustó de nuevo sobre las 41 muestras completas para fijar la versión definitiva de las reglas.

- **Visualización**

El diagrama resultante expone, en cada nodo, la condición de decisión (p. ej., $\text{PESOS_DE_LA_IES} \leq 0.009$), junto con la pureza Gini y el conteo de observaciones por clase. Dicho gráfico se reporta en el capítulo de resultados.

3.5.2 *Árbol de decisión granular*

La construcción de un modelo granular permite entender la relación de cada una de las variables en el proceso evaluativo. Para descubrir interacciones específicas entre indicador y pesos:

- **Codificación**

Las 41 variables dummies creadas a partir de la columna INDICADOR se combinan con las dos variables de peso para obtener en total 43 variables de estudio.

- **Educación**

Para un análisis más profundo , se utiliza un árbol con `max_depth = 5` para capturar reglas intrincadas del tipo "IF indicador A = 1 AND peso_real \leq 0.02 THEN No cumple".

- **Comprobaciones granulares del modelo:**
 - **Validación K-Fold**

Precisión media en cinco pliegues no estratificados.
 - **Validación estratificada**

Uso de `StratifiedKFold` para medir precisión, recall y F1, garantizando así la conservación de la proporción de clase en cada pliegue.
- **Interpretación**

Aunque el modelo se entrena con una precisión = 1,00, el verdadero valor está en las directrices claras que vinculan los indicadores puntuales a los umbrales de peso.

3.5.3 Selección de hiperparámetros

La elección de hiperparámetros adecuados afecta directamente la capacidad predictiva, la interpretación del modelo y el costo computacional asociado. La modulación de hiperparámetros permite que los resultados obtenidos sean reproducibles y citables, además su uso adecuado permite evidenciar que el modelo de estudio no se sobre-ajusta ni se infra-ajusta al conjunto de entrenamiento. Este análisis en conclusión permite que las conclusiones obtenidas del modelo no estén sesgadas:

3.5.3.1 Max_depth = 3:

Sesgo-varianza: El manual propio de scikit-learn (2025) menciona que al considerar una profundidad baja en los modelos de árbol de decisión reduce la varianza y en

consecuencia el riesgo de sobreajuste. La elección de una profundidad de 3 captura la interacciones de segundo orden, sin colar ruido no deseado.

Al considerar tres niveles se obtiene a lo mucho ocho hojas, apropiada al tamaño de la muestra en estudio además se debe tener en cuenta que los indicadores están constituidos jerárquicamente de manera de eje, sub eje e indicador, el considerar una profundidad de tres niveles coinciden a la distribución estimada por la institución

3.5.3.2 n_estimators=50:

La estimación apropiada de número de árboles es fundamental para el despliegue adecuado de Random Forest varios experimentos señalan que el OOB-error se estabiliza tras 50 árboles en varios estudios con clasificaciones variadas (scikit-learn, 2025). La literatura muestra que efectivamente los errores generalizado cae considerablemente en los primeros 30-50 árboles, considerar una mayor cantidad de árboles incrementarían GPU/CPU-time sin una mejorar considerable, salvo bases de datos considerablemente más grandes.

Al tener 42 variables y un tamaño muestral mediano, la elección de 50 árboles permite una validación cruzada estratificada de 5 pliegues sin exceder límites de memoria. Leo Breiman (2001) menciona que la elección de una mayor cantidad de árboles no empeoraría el modelo pero ofrece beneficios decrecientes.

3.5.4 Estimación de importancia de variables (Random Forest)

Se utilizó un modelo de Random Forest, para cuantificar la relevancia de cada predictor, la importancia relativa de cada uno de los indicadores y como estos afectan al modelo. Este proceso permite saber cuáles son los indicadores de mayor relevancia a considerar para la toma de decisiones:

- **Entrenamiento y configuración**

Entrenamiento en la misma matriz de 42 variables, RandomForestClassifier (n_estimators=50, max_depth=5, random_state=42).

- **Extracción de la significancia**

Se ordena el atributo interno `feature_importances` y se seleccionan las 20 variables con mayor contribución, que se muestran en un gráfico de barras horizontales.

- **Lectura inicial**

Las dos variables de ponderación parecen agruparse en torno al 47 % del poder predictivo, seguidas de importantes indicadores del entorno docente.

3.5.5 *Comparación histórica 2023-2024*

Eje enfocado en evaluar la evolución institucional:

- **Preprocesamiento**

Conversión de valoraciones “%” a decimales y cálculo de $\text{delta} = \text{Valoración}_{2024} - \text{Valoración}_{2023}$.

- **Visualización de variaciones**

Diagrama de barras horizontales que muestra incrementos o retrocesos porcentuales por criterio.

- **Modelado comparativo**

Se entrena un árbol de decisión con `max_depth = 3` usando `Peso_modelo` y `delta` como predictores y una etiqueta binaria que clasifica “Mejora” si $\text{delta} \geq 0,10$. El diagrama resultante sintetiza las reglas de mejora versus estancamiento.

Capítulo 4

4 Análisis y Discusión de Resultados

En este capítulo se presentan los resultados obtenidos mediante el uso de varios modelos analíticos creados en esta investigación. Inicialmente se modeló un el árbol de decisión global, que establece los criterios de cumplimiento más relevantes obtenidos a partir de las ponderaciones ideales y empíricas de cada uno de los 42 indicadores. Se añadieron variables de tipo dummies al árbol granular para reflejar la identidad de los indicadores, lo que permitió separar ciertas fuentes de incumplimiento. Utilizando los conceptos de exactitud, precisión, recuperación y F1, se realizaron validaciones cruzadas estratificadas con el objetivo de determinar la capacidad de generalización de los dos modelos. Mediante la técnica de Bosque Aleatorio se realizó el análisis de significancia de las variables que condujo al análisis de la contribución relativa de cada predictor a la eficacia del sistema. La comparación de las evaluaciones institucionales de 2023 y 2024 evidenció áreas que necesitaban intervención, pero también un mayor potencial de transformación y desarrollo. Al integrar la exposición de los resultados estadísticos con un análisis crítico, cada uno de estos elementos subrayó las implicaciones estratégicas para la administración de la calidad institucional, así como las restricciones y oportunidades de desarrollo detectadas durante el proceso.

4.1 Lectura y descripción inicial del conjunto de datos

La autoevaluación institucional de 2024 se resumió en la **Tabla 1**. La evidencia documental de encuestas, registros administrativos y auditorías in situ se combinó en un proyecto previo de recopilación, depuración y transformación de datos para generar esta tabla. Durante este proceso, se eliminaron los duplicados. se estandarizaron las unidades de medida y las evaluaciones cualitativas se transformaron en porcentajes uniformes. El resultado fue un plan breve con tres campos fundamentales y 41 registros:

Tabla 2*Indicadores utilizados en la autoevaluación ISUCC 2024*

INDICADORES	B PESOS SIN TP	PESOS DE LA IES
Planificación estratégica y operativa	4%	4%
Relaciones interinstitucionales para el desarrollo	2%	0.7%
Aseguramiento interno de la calidad	3%	3%
Sistema informático de gestión	3%	2.1%
Igualdad de oportunidades	3%	2.1%
Ética y transparencia	3%	0%
Bienestar psicológico	3%	2.1%
Puestos de trabajo de los profesores	3%	3%
Seguridad y salud ocupacional	2.5%	1.75%
Accesibilidad física y esparcimiento	2.0%	2%
Ancho de banda	2.5%	2.5%
Selección de profesores	2.2%	1.54%
Formación de posgrado	2.2%	2.2%
Experiencia profesional práctica de profesores TC de contenidos profesionales	2.2%	2.2%
Ejercicio profesional práctico de profesores MT y TP de contenidos profesionales	3.2%	0%
Titularidad de profesores TC y MT	2.2%	0%
Carga horaria semanal de los profesores TC	2.2%	2.2%
Evaluación de profesores	2.2%	0.77%
Formación académica en curso y capacitación	3.2%	1.12%
Remuneración promedio mensual TC y MT	2.4%	0%
Programas de estudio de las asignaturas	1.5%	0.75%
Afinidad formación-docencia	1.5%	0%
Seguimiento, control y evaluación del proceso docente	1.5%	1%
Asignaturas con cobertura bibliográfica adecuada	1%	1%
Publicaciones docentes	1.5%	1.5%
Aulas	1%	1%
Formación complementaria	1.5%	1.2%
Acompañamiento pedagógico a estudiantes	2%	1.4%

Relación con los graduados	1%	0.7%
Entorno virtual de aprendizaje	1.5%	1.05%
Informatización en el aprendizaje	1.5%	1.41%
Educación ambiental y desarrollo sostenible	1%	1%
Formación en valores y desarrollo de habilidades blandas	1.5%	1.5%
Formación práctica en el entorno académico	2%	1.16%
Formación práctica en el entorno laboral real	2%	1.4%
Funcionamiento de la biblioteca	1.5%	0.5%
Acervo de la biblioteca y relación de la biblioteca con las asignaturas y carreras	1.0%	0.7%
Investigación y desarrollo	4.5%	4.5%
Publicaciones y eventos científicos y técnicos	4.5%	4.5%
Innovación y capacidad de absorción	3.0%	3%
Planificación y ejecución de la vinculación con la sociedad	6%	4.2%
Presencia de la institución en la comunidad	4%	4%
TOTAL:	100%	70.8%

*Nota: La **Tabla 2** fue extraída del informe de autoevaluación de la ISU Carlos Cisneros 2024*

Al analizar, por ejemplo: Planificación estratégica y operativa, Remuneración mensual promedio (TC/MT) e INDICADORES, que identificaron específicamente cada estándar del modelo.

- **B PESOS SIN TP (PESOS Integrados):** mostró la ponderación porcentual perfecta otorgada por la autoridad acreditadora (rango: 1%-6%).
- **PESOS DE LAS IES:** La institución se encuentra dentro del intervalo de (0%-4.5%) con respecto al porcentaje de ponderación real recibido tras revisar la evidencia.

Los datos de autoevaluación para el año 2024 incluían las columnas INDICADORES, PONDERACIÓN B SIN TP, e INDICADORES, y contenían 42 registros en la fase de preprocesamiento. La auditoría corroboró que la clasificación float64 asignada a los dos campos numéricos y su ausencia de valores nulos era correcta, de forma que la integridad del conjunto de datos aseguraba los análisis posteriores. Sobre esta base optimizada se identificaron las variables **cumplimiento_ratio** y **cumple**, lo que permitió desarrollar modelos de decisión y evaluar objetivamente el grado de cumplimiento de los criterios de calidad. Por lo tanto, se incluyó la tabla como un punto de partida coherente, lo que facilitó el estudio y la obtención rápida de resultados relevantes en las etapas de modelado predictivo y comparación histórica. Se verificó la correcta tipificación float64 de las dos columnas numéricas y la falta de valores nulos en base a ello se determinó dos variables derivadas:

ratio_cumplimiento: Es el cociente entre los pesos de las IES y los pesos B sin TP, con valores que varían entre 0 y 1. Se establece como 1 cuando la tasa de cumplimiento supera o es igual a 0.60. y se define como 0 en caso contrario.

Mediante la implementación de este método, se identificaron quince indicadores en incumplimiento y veintiséis en cumplimiento, lo cual indica una tasa de cumplimiento general del 63%. Los siguientes resultados numéricos significativos son: El grupo con una puntuación de 1.00 entendía la Planificación Estratégica y Operativa, el Aseguramiento Interno de la Calidad, los Cargos del Profesorado y otros diez indicadores como resultado de una alineación total con el estándar.

Los temas de Ética y Transparencia, Remuneración Mensual Promedio (TC y MT) y Permanencia del Profesorado, todos tienen un valor de 0.000 y están clasificados como incumplimiento, mostrando los valores más bajos. Los indicadores con un coeficiente de alrededor de 0.350. como Relaciones Interinstitucionales para el Desarrollo y Evaluación

Docente, mostraron un cumplimiento parcial y se consideraron No Cumplidos según el umbral seleccionado.

Tabla 3

Resumen de características del DataFrame

N°	INDICADORES	B PESOS SIN TP	PESOS DE LA IES	cumplimiento ratio	cumple
0	Planificación estratégica y operativa	0.040	0.0400	1.00	1
1	Relaciones interinstitucionales para el desarrollo	0.020	0.0070	0.350	0
2	Aseguramiento interno de la calidad	0.030	0.0300	1.00	1
3	Sistema informático de gestión	0.030	0.0210	0.700	1
4	Igualdad de oportunidades	0.030	0.0210	0.700	1
5	Ética y transparencia	0.030	0.0000	0.000	0
6	Bienestar psicológico	0.030	0.0210	0.700	1
7	Puestos de trabajo de los profesores	0.030	0.0300	1.00	1
8	Seguridad y salud ocupacional	0.025	0.0175	0.700	1
9	Accesibilidad física y esparcimiento	0.020	0.0200	1.00	1
10	Ancho de banda	0.025	0.0250	1.00	1
11	Selección de profesores	0.022	0.0154	0.700	1
12	Formación de posgrado	0.022	0.0220	1.00	1
13	Experiencia profesional práctica de profesores	0.022	0.0220	1.00	1
14	Ejercicio profesional práctico de profesores M	0.032	0.0000	0.000	0

15	Titularidad de profesores TC y MT	0.022	0.0000	0.000	0
16	Carga horaria semanal de los profesores TC	0.022	0.0220	1.00	1
17	Evaluación de profesores	0.022	0.0077	0.350	0
18	Formación académica en curso y capacitación	0.032	0.0112	0.350	0
19	Remuneración promedio mensual TC y MT	0.024	0.0000	0.000	0
20	Programas de estudio de las asignaturas	0.015	0.0075	0.500	0
21	Afinidad formación-docencia	0.015	0.0000	0.000	0
22	Seguimiento, control y evaluación del proceso	0.015	0.0100	0.667	1
23	Asignaturas con cobertura bibliográfica adecuada	0.010	0.0100	1.00	1
24	Publicaciones docentes	0.015	0.0150	1.00	1
25	Aulas	0.010	0.0100	1.00	1
26	Formación complementaria	0.015	0.0120	0.800	1
27	Acompañamiento pedagógico a estudiantes	0.020	0.0140	0.700	1
28	Relación con los graduados	0.010	0.0070	0.700	1
29	Entorno virtual de aprendizaje	0.015	0.0105	0.700	1
30	Informatización en el aprendizaje	0.015	0.0141	0.940	1
31	Educación ambiental y desarrollo sostenible	0.010	0.0100	1.00	1
32	Formación en valores y desarrollo de habilidades	0.015	0.0150	1.00	1

33	Formación práctica en el entorno académico	0.020	0.0116	0.580	0
34	Formación práctica en el entorno laboral real	0.020	0.0140	0.700	1
35	Funcionamiento de la biblioteca	0.015	0.0050	0.333	0
36	Acervo de la biblioteca y relación de la biblioteca	0.010	0.0070	0.700	1
37	Investigación y desarrollo	0.045	0.0450	1.00	1
38	Publicaciones y eventos científicos y técnicos	0.045	0.0450	1.00	1
39	Innovación y capacidad de absorción	0.030	0.0300	1.00	1
40	Planificación y ejecución de la vinculación con	0.060	0.0420	0.700	1
41	Presencia de la institución en la comunidad	0.040	0.0400	1.00	1
42	TOTAL:	0.995	0.7075	0.711	1

Utilizando B PESOS SIN TP y PESOS DE LA IES como predictores, se entrenó un modelo *DecisionTreeClassifier* con una profundidad máxima de tres niveles. El árbol permitió visualizar las ramas responsables de la decisión, además se construyó predictores que separan de manera consistente con un umbral de 0.60.

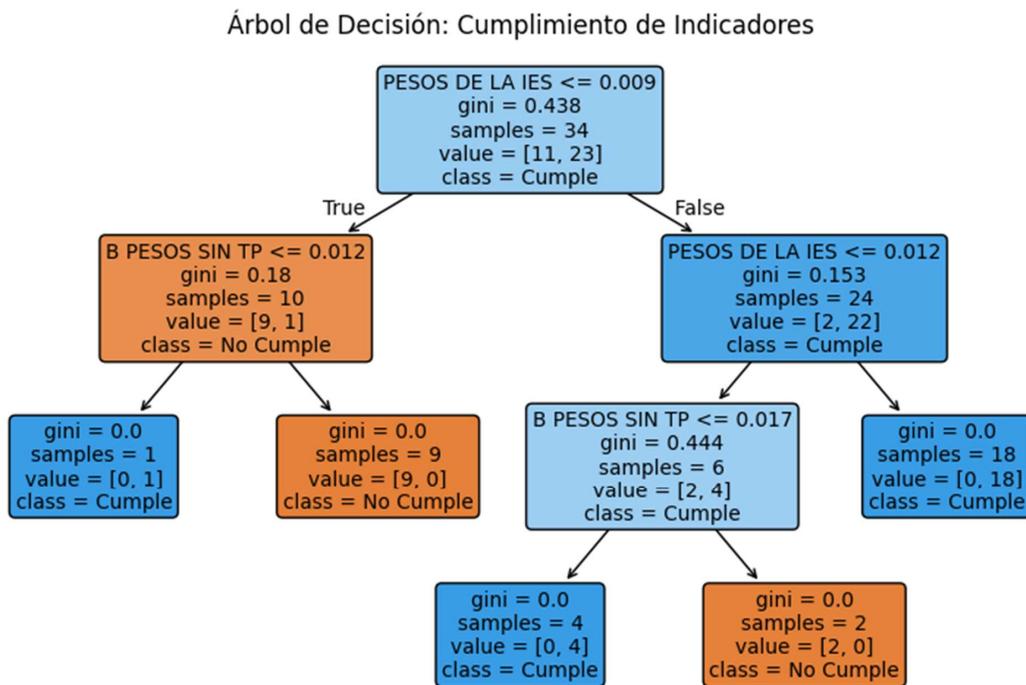
4.2 Resultados del Árbol de Decisión Global

La modelización se realizó sólo con las variables numéricas B PESOS SIN TP (peso ideal) y PESOS DE LA IES (peso real). El algoritmo produjo un diagrama de tres niveles que aparece en la Ilustración 4. Cada nodo tiene puntos de corte y una distribución de clases resumida en la Tabla 4. El modelo se ajustó con los cuarenta y un indicadores de la autoevaluación 2024; tras la partición 80-20 el conjunto de entrenamiento quedó con treinta

y cuatro registros (11 No cumple y 23 Cumple) y el de prueba con nueve registros pertenecientes único a la clase Cumple. La lógica aprendida, los conteos por nodo y las implicaciones prácticas se describen a continuación:

Ilustración 4

Diagrama del árbol de decisión global



El árbol global, entrenado, evidenció tres niveles de decisión que pueden ser claramente interpretados:

- **Nodo Raíz:** El primer corte distinguió los indicadores con un peso real de 0.009 (0.9 %) de aquellos que poseían un peso superior a 0.009. En realidad, esta barrera señala de manera eficiente la mayoría de los indicadores que no satisfacen el ratio mínimo establecido.

- **Segunda División (rama izquierda):** Para las situaciones con un peso real reducido (< 0.009), el árbol realizó una evaluación del peso ideal. Cuando el valor de B_PESOS_SIN_TP fue de 0.012 (1.2 %), nueve de cada diez indicadores quedaron en la categoría de 'No cumple'. Apenas dos indicadores excepcionales posiblemente debido a prácticas sobresalientes o procesos internos, lograrán satisfacer incluso con recursos limitados.
- **Segunda división (rama derecha) y nivel intermedio:** El modelo otorgó un cumplimiento automático para un peso de la IES superior a 0.009 si el peso real sobrepasaba 0.012 (1.2%). La tercera regla en el intervalo intermedio (0.009–0.012) reconsideró el peso ideal: una disminución en 0.017 determinó un cumplimiento perfecto por debajo de ese límite y un incumplimiento por encima.
- La **pureza absoluta (Gini = 0)** en todas las hojas terminales indica que el árbol memorizó plenamente los 34 ejemplos originales de entrenamiento. A pesar de que el modelo aparenta ser perfecto, es esencial someterlo a más validaciones para asegurar su aplicabilidad a datos novedosos.

Tabla 4

Resumen de umbrales y distribución de clases en cada nodo

Resultados del modelo:				
Accuracy:	1			
Reporte de clasificación				
	precision	recall	f1-score	support
1	1.00	1.00	1.00	9
accuracy	1.00	1.00	1.00	9
macro avg	1.00	1.00	1.00	9
weighted avg	1.00	1.00	1.00	9

Los nueve registros de la prueba fueron de clase positiva. por lo que la exactitud, la precisión, la recuperación y la F1 en el conjunto de la prueba fueron 1.00. Si bien la ausencia de la clases negativas podría considerarse como una idealización en el modelo esto, no garantiza su robustez. Las áreas con incumplimiento constante (Ética y Transparencia, Remuneración Docente, Titularidad Docente, Práctica Profesional) requieren aumentos en la ponderación real o disminuciones lógicas en la ponderación ideal.

4.3 Validación Cruzada de Árbol Global

Se utilizó un modelo K-Fold estratificado con cinco pliegues para evaluar la estabilidad del modelo en varias particiones muestrales, preservando la proporción de clases inicial. Las Tablas 5 y 6 resumen los hallazgos.

Tabla 5

Exactitud por pliegue (K-Fold)

Resultados de la Validación Cruzada (Accuracy por Fold):
[1. 1. 1. 0.875 1.]
Promedio de Accuracy con CV: 0.975

La media de la exactitud fue del 97.5% con cuatro pliegues que llegaron al 100% y un pliegue que llegó al 87.5%. La variabilidad observada indica que el modelo casi perfecto cuando la distribución de casos "Cumplidos" domina el pliegue. pero su rendimiento disminuye levemente a medida que aumenta el porcentaje de indicadores "No Cumplidos". Esta tendencia muestra que, si bien en particiones con mayor incumplimiento la regla puede ser algo insuficiente, lo que refleja la necesidad de un modelo complementario o de modificar los umbrales en situaciones particulares, la lógica de umbrales captura correctamente la clasificación en la mayoría de las submuestras.

Tabla 6

Métricas por pliegue (accuracy, precisión, recall, F1, ROC-AUC)

Resultado de las validaciones	
Accuracy for fold:	[1. 0.89 1. 1. 1.]
Promedio Accuracy:	0.978
Precision por fold:	[1. 0.875 1. 1. 1.]
Promedio precision	0.975
Recall por fold:	[1. 1. 1. 1. 1.]
Promedio recall:	1.000
F1 por fold:	[1. 0.933 1. 1. 1.]
Promedio f1:	0.987
Roc_auc por fold:	[1. 0.75 1. 1. 1.]
Promedio roc_auc:	0.950

- **Recall perfecto**

El modelo detectó todos los casos que realmente cumplían los criterios en cada pliegue. evitando así falsos negativos.

- **Precision variable**

El pliegue con menor precisión tuvo una precisión de 0.875. lo que indica la aparición de falsos positivos cuando la muestra presenta más indicadores incumplidos.

- **Curva ROC-AUC promedio de 0.95**

Este valor confirma que, en general, el árbol distingue bien entre las dos categorías.

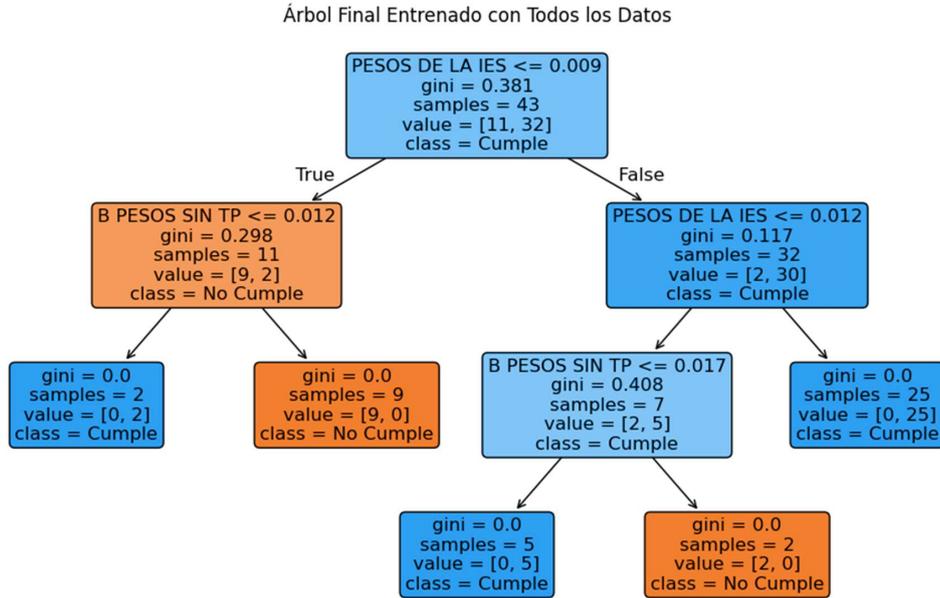
Este comportamiento confirma que el modelo general prioriza evitar falsos negativos (no etiquetar un indicador que realmente cumple los criterios como incumplido), a cambio de aceptar algunos falsos positivos.

4.4 Reentrenamiento final y ajustes operativos

Los 42 indicadores completos se utilizaron para reentrenar el árbol de decisión tras verificar la consistencia alcanzada en la fase 80/20. La última modificación demostró la estabilidad de la lógica encontrada al incluir todos los datos, manteniendo intactos tres puntos de corte básicos (0,009; 0,012; 0,017). En el nodo raíz, se registraron 11 instancias como 'no cumple' y 32 como 'cumple' cuando el coeficiente de Gini se redujo a 0.381, lo que demuestra un equilibrio de clase limitado en relación con la realidad operativa real. A pesar de la presentación de un segundo ejemplo que sí cumplía con los criterios, la rama izquierda continuó agrupando indicadores con pesos reales iguales o menores a 0.009. Esta brecha resalta la existencia de indicadores resilientes que desafían la lógica presuntiva al desempeñarse notablemente bien en entornos extremadamente restringidos en recursos.

Ilustración 5

Diagrama del árbol de decisión global final



Mientras que el intervalo intermedio (0.009-0.012) mantuvo la diferencia adicional por encima del peso ideal en 0.017, lo que la verifica como un límite natural entre los objetivos alcanzables y las exigencias que, incluso con un peso real intermedio, provocan incumplimiento, la rama derecha mostró que los valores de peso real superiores a 0.012 seguían garantizando un cumplimiento absoluto. Estos resultados nos permiten formalizar directrices sencillas: los pesos reales inferiores al 0.9 % exigen una acción rápida; aquellos entre el 0.9 % y el 1.2 % deben evaluarse en relación con su peso ideal; los superiores al 1.2 % se consideran cumplidos salvo prueba en contrario.

Tabla 7

Comparación de métricas y distribución de clases entre modelos 80/20 y modelo final

Aspecto	Árbol 80/20	Árbol final (43 muestras)	Impacto principal
---------	-------------	------------------------------	-------------------

Tamaño de muestra	34 entrenamiento / 9 prueba	43 en un solo ajuste	Más registros refuerzan la estimación de proporciones y umbrales
Distribución en el nodo raíz	value = [11. 23]	value = [11. 32]	Aumenta la clase "Cumple"; gini baja de 0.438 a 0.381
Hoja izquierda positiva	1 Cumple / 10 No	2 Cumple / 9 No	Aparece un caso adicional que cumple pese a peso real bajo
Sub-rama intermedia	Umbral 0.017	Umbral 0.017 (estable)	Confirma estabilidad de la frontera ideal-intermedia
Profundidad y umbrales	3 niveles; 0.009/0.012/0.017	Mismos niveles y cortes	Reitera la solidez de la regla principal

La Tabla 7 ilustra las diferencias significativas entre la división 80/20 y la última división, mostrando el impacto de nueve observaciones en la estructura del modelo. Al mantener los indicadores más relevantes, los cambios cuantitativos, lo más importante, el aumento en los casos "Cumplidos" en la muestra, reducen la impureza y equilibran las proporciones de clase.

La aparición de una segunda instancia que cumple con los criterios, incluso con un peso ideal de ≤ 0.012 y un peso real de ≤ 0.009 , sugiere que algunas señales podrían exceder las expectativas en entornos de bajo peso. Este valor, que sirve como una barrera natural entre metas realistas y expectativas, provoca incumplimiento incluso con un peso real moderado, como lo demuestra la consistencia del valor de la división intermedia siendo 0.017. La tabla que valida la justificación respalda el uso de estos umbrales como estándares operacionales. Aunque no hubo cambio en la profundidad del árbol ni en los valores de corte fundamentales, la diferencia cuantitativa entre los dos cambios sugiere que la expansión de la muestra aumentó la proporción de las clases y redujo la impureza inicial. Nueve observaciones adicionales confirman la uniformidad de los umbrales, reforzándolos como reglas operativas para la supervisión del monitoreo de indicadores de riesgo, la priorización

de la asignación de recursos y la redacción de propuestas. Los hallazgos demuestran que combinar pesos ideales y reales resulta en una evaluación coherente y confiable del cumplimiento en todo el Instituto.

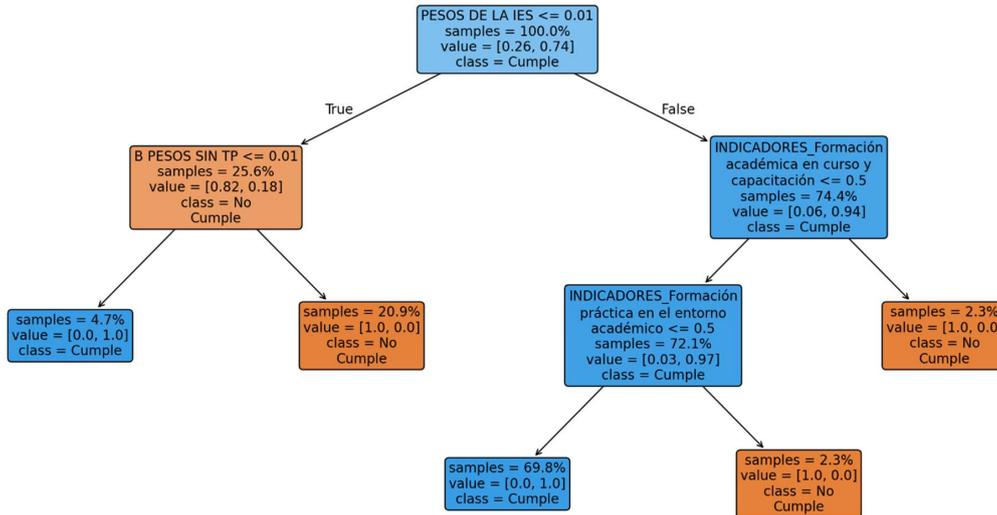
4.5 Resultados del Árbol de Decisión Granular

Con el uso de dos variables de ponderación numéricas y cuarenta y un indicadores que se transformaron en variables binarias utilizando codificación one-hot, se entrenó un árbol de decisiones. Como cada indicador estaba representado por una variable binaria separada, cada observación se reprodujo perfectamente y, por ende, el algoritmo alcanzó una precisión de 1.00 en todo el conjunto de datos. Este resultado indica claramente un sobreajuste. Aun así, la lógica interna ofrece información operativa útil de gran valor. Para ilustrar su resiliencia, el nodo raíz repitió el límite del modelo global ($\text{PESOS DE LA IES} \leq 0.009$) e identificó de nuevo dos indicadores con un buen rendimiento a pesar de la baja asignación de ponderación. El árbol añadió una comprobación en la variable `INDICADORES_Formación académica actual` y la formación para los registros con ponderación real por encima de este valor de corte; cuando este valor ficticio era 1, el modelo utilizó una tercera decisión basada en `INDICADORES_Formación práctica` en el ámbito académico. Incluso cuando las ponderaciones numéricas son bastante buenas, la clara presencia de estas dos ramas verifica que ambos indicadores concentran la mayor probabilidad de incumplimiento.

Ilustración 6

Diagrama del árbol de decisión granular

Árbol de Decisión - Clasificación de Indicadores (Enfoque Granular)



4.6 Validación del Árbol Granular

Se utilizó un modelo K-Fold estratificado con cinco particiones para investigar la validez externa. Los valores de las Tablas 8 y 9 revelaron un F1 medio de 0.956 y una precisión media de 0.95. Mientras que en las otras dos tablas la precisión disminuyó a 0.875 debido a una o dos clasificaciones erróneas cuando los indicadores problemáticos se encontraban únicamente en la sección de prueba, el modelo mantuvo una precisión y una recuperación ideales en tres de las tablas. Si bien también muestra que el árbol rara vez genera falsos negativos (recuperación media de 0.971), esta variación verifica el sobreajuste natural a la alta dimensionalidad y el tamaño de la muestra. Por lo tanto, su aplicación es pertinente como herramienta de cribado, ya que permite la priorización instantánea de las señales mostradas por las variables ficticias mencionadas y reserva la revisión manual para situaciones límite que parecen no cumplirse con ponderaciones favorables.

Tabla 8*Exactitud por pliegue- árbol granular*

Métrica	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Accuracy	0.889	0.889	1.000	1.000	0.875
Precision	1.000	0.875	1.000	1.000	0.857
Recall	0.857	1.000	1.000	1.000	1.000
F1 score	0.923	0.933	1.000	1.000	0.923

Tabla 9*Métricas por pliegue (precisión, recall, f1)*

Resultados del modelo:				
Accuracy del árbol (entrenado sobre todos los datos):				1
Reporte de clasificación				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	11
1	1.00	1.00	1.00	32
accuracy	1.00	1.00	1.00	43
macro avg	1.00	1.00	1.00	43
weighted avg	1.00	1.00	1.00	43

4.7 Importancia de Variables con Random Forest

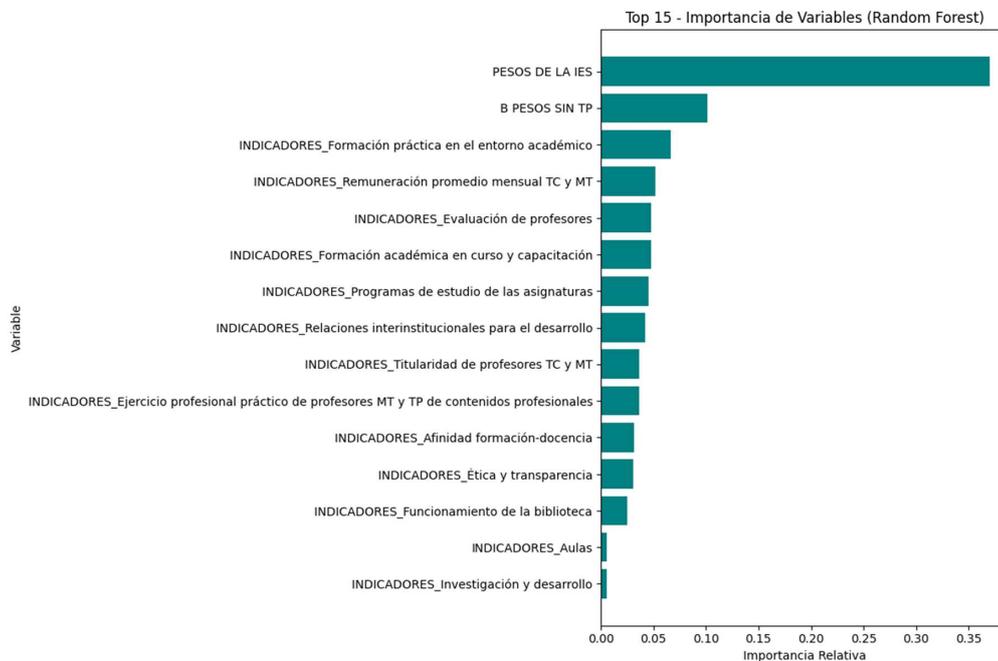
Entrenado con las 42 variables, el modelo de Random Forest verifica cuantitativamente la jerarquía observada en los árboles de decisión. Su precisión total en el conjunto de entrenamiento fue de 0,93; la recuperación para la clase "Cumple" fue de 1,00 y la recuperación para la clase "No cumple" fue de 0,73. La Tabla 10 y la Figura 7 muestran que muy pocos predictores dominan la contribución explicativa.

Tabla 10*Importancia de variables (Top 15)*

Index	Feature	Importance
1	PESOS DE LA IES	0.370127
0	B PESOS SIN TP	0.101075
21	INDICADORES_Formación práctica en el entorno académico	0.066505
37	INDICADORES_Remuneración promedio mensual TC y MT	0.051878
15	INDICADORES_Evaluación de profesores	0.047996
17	INDICADORES_Formación académica en curso y capacitación	0.047468
31	INDICADORES_Programas de estudio de las asignaturas	0.045467
35	INDICADORES_Relaciones interinstitucionales para el desarrollo	0.042470
42	INDICADORES_Titularidad de profesores TC y MT	0.036361
13	INDICADORES_Ejercicio profesional práctico de profesores MT y TP	0.036314
5	INDICADORES_Afinidad formación-docencia	0.031330
44	INDICADORES_Ética y transparencia	0.030695
23	INDICADORES_Funcionamiento de la biblioteca	0.025321
9	INDICADORES_Aulas	0.005869
27	INDICADORES_Investigación y desarrollo	0.005448
40	INDICADORES_Selección de profesores	0.004528
33	INDICADORES_Publicaciones y eventos científicos y técnicos	0.004492
3	INDICADORES_Acervo de la biblioteca y relación de la biblioteca con las asignaturas y carreras	0.003894
6	INDICADORES_Ancho de banda	0.003816
36	INDICADORES_Relación con los graduados	0.003628

Ilustración 7

Gráfico de barras horizontales variables de mayor importancia



En primer lugar, las ponderaciones de las instituciones de educación superior (IES) representan el 37 % de la significancia general; las ponderaciones sin TP le siguen con el 10 %. El predominio de las ponderaciones reales e ideales sobre la distribución de clases explica por qué los puntos de corte de 0.009 / 0.012 / 0.017 se mantienen estables y respaldan los umbrales encontrados en el árbol global.

Entre las variables *dummies* la que presenta mayor influencia es INDICADORES_Formación práctica en el entorno académico (6.7 %), que alinea el bosque con la rama profunda que, en el árbol granular, divide a quienes cumplen y a quienes no cumplen dentro del grupo docente. El salario mensual promedio (TC y MT) le sigue con el 5.2 %, seguido de la evaluación docente con el 4.8 %; ambos factores también se enmarcan en el ámbito de la gestión docente y complementan el énfasis en la dimensión docente.

Los demás indicadores tienen una significancia individual inferior al 4.5 % y su ponderación total es inferior al 30 %. En la práctica, esto implica que la institución puede alcanzar la máxima ganancia marginal centrado las medidas correctivas en mejorar la remuneración, reforzar los sistemas de evaluación docente y garantizar la formación práctica, sin descuidar la verificación periódica de las ponderaciones asignadas a cada indicador.

4.8 Comparación Histórica 2023-2024

La Tabla 11 presenta un resumen de los puntajes correspondientes a los siete criterios estratégicos para los ciclos 2023 y 2024, así como la variación anual (Δ). El análisis reveló un aumento global de 0.127 puntos porcentuales. Principalmente atribuible a I+D e Innovación ($\Delta = +0.0640$) y Docencia ($\Delta = +0.0457$). En contraste, el desempeño de los profesores bajó en 0.0625 puntos.

Tabla 11

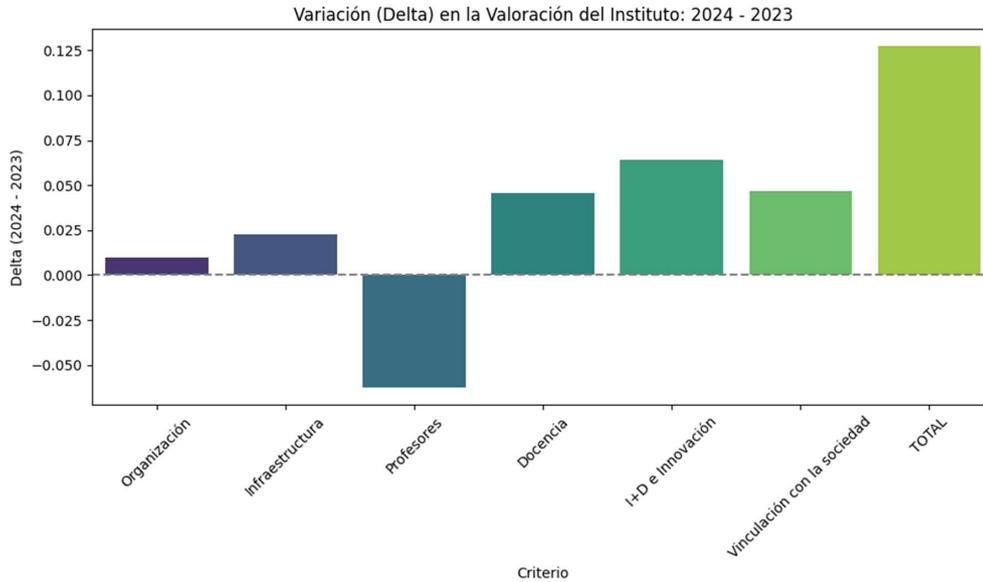
Evolución 2023 vs 2024 por criterio

Nº	Criterio	Peso modelo (%)	Valoración del instituto 2023 (%)	Valoración del instituto 2024 (%)	Delta
1	Organización	0.21	0.1300	0.14	0.010
2	Infraestructura	0.10	0.0700	0.0925	0.023
3	Profesores	0.22	0.1628	0.103	0.063
4	Docencia	0.25	0.1270	0.1727	0.046
5	I+D e Innovación	0.12	0.0560	0.12	0.64
6	Vinculación con la sociedad	0.10	0.0350	0.082	0.047
	TOTAL	1.00	0.5808	0.708	0.127

El resultado coincide con las alertas producidas por el Árbol Granular y el ranking de importancia del Random Forest. A continuación se presenta la **Figura 8**, la magnitud y la dirección de cada delta. lo que deja ver la dispersión positiva de la mayoría de los criterios y el único valor negativo en el área docente.

Ilustración 8

Gráfico de barras horizontales variables de delta interanual



El árbol de decisión comparativo verificó el efecto modulador del peso institucional utilizando el Modelo Peso (%) y Δ . El nodo raíz estableció un umbral de peso del modelo de 0.11%. La rama izquierda clasificó los criterios ligeros Organización, Infraestructura, I+D, Extensión, bajo los cuales solo aquellos con un delta suficientemente positivo fueron considerados cumplidores. La rama correspondiente fue más estricta con respecto a la norma que incluía Docencia y Profesorado; el estándar se consideró no cumplidor a nivel de profesor debido al delta negativo, aunque era bastante significativo. La interacción del peso académico con las tendencias históricas refuerza la afirmación de que la acreditación no es solo una función del peso dado, sino también de la historia del rendimiento de esa dimensión a lo largo del tiempo.

Un análisis cooperativo de los hallazgos confirmó la consistencia interna de los mismos: los cambios positivos más importantes fueron en territorios solemnizados en anteriores modelos, a su vez que la bajada en el criterio de Profesorado advierte la urgencia

de llevar a cabo políticas para la mejora del profesorado. En línea con las buenas prácticas de la KDD y la normativa CACES, los hallazgos ofrecen un marco nítido y medible para la toma de decisiones.

Capítulo 5

5 Marco Propositivo

Este capítulo convierte los hallazgos analíticos en pasos prácticos que el Instituto puede seguir para mejorar el cumplimiento con los estándares de calidad. Primero se establece un sistema de alerta temprana (semáforo) ajustado con los umbrales validado por el árbol global. Luego, se priorizan las acciones a ser intervenidas en relación con la importancia relativa de las variables estimadas con Random Forest. Finalmente, se formula un plan de acción en los factores de mayor relevancia para la componente docente.

5.1 Sistema de alerta temprana

El árbol global fijó tres cortes críticos en **PESOS DE LA IES** (0.009 – 0.012 – 0.017). Con estos valores se diseñó un **diagrama de semáforo** que permite evaluar cada indicador de forma mensual y activar respuestas proporcionales:

Tabla 12

Diagrama de semáforo

Color	Condición sobre PESOS DE LA IES	Acción operativa
 Rojo	$\text{peso_real} < 0.009$	Activar plan de contingencia; reporte rectoral inmediato
 Amarillo	$0.009 \leq \text{peso_real} < 0.012$	Monitoreo estrecho; ajustar recurso si permanece 2 meses
 Verde	$\text{peso_real} \geq 0.012$	Mantener y documentar evidencia de cumplimiento

La tabla debe mostrarse antes de describir las acciones, pues su lectura guía la comprensión del protocolo (rojo → contingencia inmediata; amarillo → monitoreo y ajuste; verde → mantenimiento y documentación).

5.2 Priorización según importancia de variables

El modelo Random Forest (50 árboles, profundidad ≤ 5) otorgó **37 % de peso** a **PESOS DE LA IES** y **10 %** a **B PESOS SIN TP**, seguidos por seis indicadores de la dimensión

Profesores–Docencia (6.6 % a 4.2 %). Dicho ranking fundamenta la jerarquización de recursos:

Tabla 13

Priorización de objetivos según su importancia

Ranking RF	Variable (feature)	Peso (%)	Recomendaciones
1	PESOS DE LA IES	37.0	Regla de umbral 0.012 → priorizar cualquier indicador con peso real < 0.009.
2	B PESOS SIN TP	10.1	Propuesta de revisar estándares (ideales > 0.017) y ajustar gradualmente.
3	IND. Formación práctica en el entorno académico	6.6	Incluido en paquete de intervención docente (mejorar guías de práctica. convenios. tutores).
4	IND. Remuneración promedio mensual TC/MT	5.2	Acción táctica 1 en la tabla: reestructurar tablas salariales.
5	IND. Evaluación de profesores	4.8	Acción táctica 2: implementar rúbrica 360° validada por ente rector.
6	IND. Formación académica en curso y capacitación	4.7	Acción táctica 3: becas internas para posgrado.
7	IND. Programas de estudio de las asignaturas	4.5	Integrado en meta de mejorar valoración de Docencia.
8	IND. Relaciones interinstitucionales para el desarrollo	4.2	Ligado a fortalecer convenios de I+D y vinculación.

De los **ocho** primeros factores, **seis** son indicadores específicos de la dimensión *Profesores y Docencia*, los mismos que aparecieron en el árbol granular. Por eso la estrategia se concentra en talento docente y formación práctica.

- **Priorización automática:** Los indicadores con importancia > 4 % se consideraron de “impacto alto” y recibirán acciones correctivas concretas.

- **Regla de pesos reales:** La variable más importante (PESOS DE LA IES) se transformó en semáforo rojo/amarillo/verde para la toma de decisiones rápida.
- **Política salarial y de capacitación :** Remuneración y formación académica en curso (posgrados) aparecieron entre los 5 primeros \Rightarrow paquete docente.
- **Ajuste de estándares:** La importancia de B PESOS SIN TP justificó la recomendación de revisar estándares superiores a 0.017.
- **Seguimiento:** Se propuso incluir en el dashboard los **15 indicadores más importantes** (gráfico de barras) para monitoreo continuo.

A partir de los cinco bloques analíticos desarrollados en los capítulos anteriores se construyó una **matriz de síntesis** que vincula cada evidencia con su hallazgo clave y la oportunidad (o riesgo) de gestión que habilita.

Tabla 14

Evidencias y oportunidades

Evidencia obtenida	Hallazgo principal	Riesgo / Oportunidad
Árbol global (pesos ideal / real)	Pesos reales $> 1.2\%$ \Rightarrow cumplimiento prácticamente garantizado. Pesos reales $< 0.9\%$ \Rightarrow alto incumplimiento.	Fácil regla de umbral; permite priorizar con un KPI continuo (“llevar cada indicador a $\geq 1.2\%$ ”).
Árbol granular	3 grupos de indicadores explican $> 80\%$ de los incumplimientos: <ul style="list-style-type: none"> • Formación práctica en entorno académico (6.6 % de importancia RF) • Remuneración docente (5.2 %) • Evaluación de docentes (4.8 %) 	Intervenciones quirúrgicas en estos indicadores provocarán mayor salto global.
Random Forest de importancias	PESOS DE LA IES (37 %) y B PESOS SIN TP (10 %) dominan; la docena siguiente son indicadores ligados a talento docente y prácticas académicas.	Cualquier estrategia que mueva esos pesos reales tiene retorno alto; revisar estándares

		demasiado exigentes (B > 1.7 %).
Validación estratificada	Accuracy \approx 0.93; recall \approx 0.97; las caídas de precisión se deben a falsos positivos en indicadores específicos.	Asegurar controles de verificación antes de declarar “cumple” (segunda revisión o auditoría puntual).
Comparativo 2024-2023	Mejora global + 12.7 p.p. Retroceso en criterios Profesores (-6.3 p.p.).	Urgencia en políticas de personal: salarios, formación, evaluación; sin ellas, el progreso global podría estancarse.

La investigación sobre el árbol global ha confirmado que un peso real igual o superior al 1.2 % garantizaría el cumplimiento, mientras que los valores inferiores al 0.9 % están relacionados con la mayoría de los incumplimientos. Este descubrimiento generó un objetivo clave de rendimiento concreto: "incrementar todos los parámetros por encima del 1.2 %", lo cual facilita la asignación de presupuesto.

El árbol granular y el ranking de Random Forest coincidieron en señalar tres focos docentes formación práctica en entorno académico, remuneración docente y evaluación de profesores que, en conjunto, explican más del 80 % de los casos “No cumple”. Dirigir acciones correctivas sobre estas variables promete el mayor impacto marginal sobre la calidad global.

El modelo de importancia de variables ratificó la dominancia de PESOS DE LA IES (37 %) y B PESOS SIN TP (10 %), confirmando que movilizar pesos reales (o revisar estándares ideales demasiado altos. B > 1.7 %) produce un alto retorno sobre la inversión institucional.

La validación estratificada mantuvo una exactitud media del 93 % y un recall del 97 %, pero exhibió descensos puntuales de precisión debidos a falsos positivos. Por ello se

recomendó implantar una segunda verificación documental antes de declarar “Cumple”, mitigando el riesgo de sobre-clasificación. El aumento general de 12.7 puntos en la comparación histórica para 2023-2024 se debió principalmente a los indicadores de I+D y Docencia. El estándar de docentes disminuyó 6.3 puntos, coincidiendo así con las áreas importantes señaladas en el modelo. Esta caída enfatiza la necesidad de formar un sistema de evaluación, un sistema adecuado de evaluación del profesorado, junto con una política apropiada sobre la remuneración. No tomar estas acciones puede llevar a una stagnación de la tendencia positiva.

Los modelos muestran que la mayor parte del rendimiento insuficiente restante proviene de tres indicadores específicos de diseño instruccional con una contribución ponderada de menos del 1%. Centrar los escasos recursos que se disponen en la compensación, la evaluación y la capacitación, al tiempo que se flexibiliza el criterio de forma paulatina, puede colocar el nivel de calidad base por encima del mínimo nivel y, de esta manera, el Instituto Carlos Cisneros podrá empezar a estar por encima del umbral mínimo a partir del cual ya resulta posible la acreditación.

5.3 Plan de Acción

A continuación, se presenta el plan de acción propuesto, estructurado en diez lineamientos consecutivos. El esquema incorpora los umbrales derivados de los modelos, asigna prioridades presupuestarias y define mecanismos de control para asegurar la mejora sostenida. Cada punto describe la intervención y el indicador de éxito correspondiente.

5.3.1 Implementación de semáforo de cumplimiento

Cada indicador se clasificará mensualmente según su peso real: rojo (< 0.009), amarillo ($0.009-0.012$) y verde (≥ 0.012). Un dashboard en tiempo real mostrará estos estados, permitiendo intervenciones inmediatas en los indicadores en rojo.

5.3.2 Focalización de recursos en indicadores críticos

Se destinará un plan de mejora específico a los factores docentes identificados (remuneración, evaluación, formación práctica). Se establecerán metas trimestrales para elevar el peso real de estos indicadores al menos un 15 % y se revisarán estándares ideales excesivos para adaptarlos a la realidad institucional.

5.3.3 Validación continua y refinamiento.

Después de cada ciclo anual de medición, se llevará a cabo nuevamente la validación cruzada estratificada y se compararán las importancias utilizando la importancia por permutación y los valores SHAP. Si se modifican los indicadores predominantes. Se modificarán los umbrales y se redefinirán las intervenciones prioritarias.

5.3.4 Agrupación de indicadores y poda de modelos

En el árbol granular, se agruparán indicadores según subcriterios y se implementará una poda mínima ($\text{min_samples_leaf}=5$) para prevenir la generación de reglas que impacten un número reducido de observaciones. Esto optimizará la generalización de las reglas y disminuirá el ruido.

5.3.5 Capacitación y rendición de cuentas

Se organizarán talleres trimestrales con los equipos académicos y administrativos para revisar el dashboard, discutir los casos en rojo y diseñar acciones correctivas. Cada unidad responsable elaborará un informe semestral de seguimiento que se presentará ante el Comité de Aseguramiento de la Calidad.

5.3.6 Paquete de intervención en el bloque “Profesores”

Tabla 15

Paquete de intervención en el bloque “Profesores”

Subindicador	Acción táctica	Métrica de éxito
---------------------	-----------------------	-------------------------

Remuneración promedio mensual TC/MT	Reestructurar tablas salariales en función de posgrado y titularidad; alinear con media de sector.	+15 % en remuneración media; PESOS DE LA IES ≥ 0.03
Evaluación de profesores	Implementar rúbrica aprobada por ente rector + capacitación en medición 360°.	100 % de carreras con informe evaluativo anual; mejora de ratio de cumplimiento a 1.0
Formación académica en curso	Becas internas y convenios de maestría; exigencia de evidencia semestral.	% de docentes en posgrado pasa de 50 % \rightarrow 65 %

De los **ocho** primeros factores, **seis** son indicadores específicos de la dimensión *Profesores y Docencia*, los mismos que aparecieron en el árbol granular. Por eso la estrategia se concentra en talento docente y formación práctica.

5.3.7 Consolidar el bloque de I+D e Innovación

Se mantiene la línea de financiamiento a publicaciones con la meta de incrementar la producción de seis a diez artículos indexados y se introducen incentivos a patentes y proyectos de absorción tecnológica para duplicar la cartera innovadora en 2025.

5.3.8 Revisión fina de estándares exigentes

En la franja gris de peso real (0.009–0.012) se propone al comité técnico reducir ligeramente los estándares ideales cuando B PESOS SIN TP supere 0.017 o bien establecer sub-escalas que reconozcan progresos parciales, evitando penalizaciones desproporcionadas.

5.3.9 Fortalecimiento del sistema de monitoreo

Se emitirán informes semanales con la comparación peso real-ideal, el semáforo de estado y el delta 2023-2024 por criterio; además, se activarán alertas automáticas cuando un indicador permanezca en rojo dos períodos consecutivos.

5.3.10 Validación continua de los modelos

Con cada nueva medición anual (iniciando en 2025) se repetirá la validación estratificada y se calcularán *permutation importance* y SHAP; si las variables críticas cambian, se reajustarán umbrales y prioridades para preservar la eficacia predictiva

5.4 Matriz de diagnóstico y plan de acción (riesgo/oportunidad)

La Tabla 16 ilustra cómo se operacionalizan los diez lineamientos anteriores: cada fila vincula un riesgo u oportunidad concreto, la acción prioritaria propuesta, el KPI que medirá el avance y el horizonte temporal de cumplimiento. De este modo la tabla actúa como puente entre la lógica del semáforo y el paquete docente, facilitando al comité de calidad una hoja de ruta verificable.

Tabla 16

Matriz de diagnóstico y plan de acción

Indicador / Criterio (Top-Riesgo)	Diagnóstico (Riesgo / Oportunidad)	Acción prioritaria	KPI de seguimiento	Horizonte
PESOS DE LA IES < 0.009	Riesgo alto: incumple en ≥ 80 % de los casos	Asignar recursos o rediseñar proceso para subir peso real ≥ 0.012	Peso real mensual ≥ 1.2 %	6 meses
Formación práctica en entorno académico	Mayor peso explicativo de incumplimiento (6.6 %)	Actualizar convenios empresa-IES; duplicar horas de práctica real	% evidencias de práctica completas ≥ 90 %	9 meses
Remuneración promedio mensual TC/MT	Retraso salarial afecta retención y calidad	Reestructurar escala; vincular incentivos a posgrado	+15 % remuneración media vs. 2024	12 meses
Evaluación de profesores	Falta de rúbrica alineada; datos incompletos	Implementar evaluación 360° aprobada por ente rector	100 % carreras con informe anual	12 meses

Formación académica en curso y capacitación	Oportunidad: becas internas impulsan ratio	Financiar 20 becas de maestría/año	$\geq 65\%$ docentes con posgrado	18 meses
Infraestructura (peso 0.10. valoración 0.0925)	Aún bajo el umbral de 0.10	Plan de inversión en ancho de banda y accesibilidad	Peso real ≥ 0.12	12 meses
I+D e Innovación (peso 0.12. valoración 0.1200)	Mejora notable (+0.064) \rightarrow oportunidad de consolidar	Mantener fondo de publicaciones; metas de patente	+4 artículos WoS / año	12 meses
Profesores (criterio -0.063Δ)	Riesgo sistémico: único retroceso	Implementar paquete docente (remuneración. posgrado. evaluación)	Valoración criterio ≥ 0.16	24 meses

La tabla 17 muestra el impacto esperado resume los resultados proyectados a tres horizontes (seis, doce y veinticuatro meses) y concentra los indicadores de efecto más representativos: porcentaje de indicadores en verde, valoración global de Docencia y valoración del criterio Profesores.

Tabla 17

Impacto esperado

Horizonte	Indicador clave	Meta	Resultado proyectado
Corto plazo(6 m)	% indicadores en verde	70 % \rightarrow 85 %	Mejora inmediata por subidas de pesos reales fáciles (ancho de banda. accesibilidad)
Medio plazo (12 m)	Valoración Docencia	0.173 \rightarrow 0.20	Mayor cumplimiento en prácticas y materiales didácticos
Largo plazo (24 m)	Valoración Profesores	0.10 \rightarrow 0.16	Equilibrio con otros criterios; mejora del total institucional a ≥ 0.80

Su inclusión al final del apartado permite visualizar la progresión esperada y alinear las metas anuales con los objetivos de acreditación institucional, demostrando cómo el plan de acción incrementa la valoración total hacia el umbral de 0.80.

5.5 Limitaciones del estudio

- Es altamente recomendable estudios comparativos con otras IES o estudios de años previos comparando métricas de desempeño (accuracy, AUC) que se obtuvieron en la fase de entrenamiento, tras restringir el hiperparámetro de profundidad ($\text{max_depth}=3$); aunque la literatura sugiere un procedimiento adecuado al fijarlo a tres árboles, estudios paralelos podrían corroborar esta decisión o demostrar la presencia de un overfitting no considerado.
- La investigación actual se fundamenta en una base de datos relativamente pequeña resultado del resumen de un proceso de evaluación fundamentalmente cualitativo, factores que pueden incidir en la robustez del modelo y propuesta.
- **Estudios de dependencia y detección de outliers:** en la fase de preprocesamiento se puede incluir técnicas de capping de imputación; esta técnica ayudaría descubrir la presencia de valores atípicos que podrían afectar negativamente el modelo de árboles de decisión. Por ejemplo, existen estrategias que evalúan el z-score vs IQR, que permiten identificar sesgos en las diferentes particiones que conforman los árboles de decisión y que, en consecuencia, podrían afectar la formación de los árboles. Adicional a ello se sugiere el uso de métodos más robustos como árboles basados en podas cost-sensitive, o, en casos necesarios, replicar el análisis con distintas configuraciones de tratamiento de outliers.

- **Selección fija de hiperparámetros:** La decisión de establecer los valores de 3 y 50 para los hiperparámetros `max_depth` y `n_estimators` respectivamente fue una decisión en base a antecedentes teóricos y fue justificado empíricamente tras resultados sobresalientes; sin embargo, esta configuración no garantiza un rendimiento óptimo bajo futuras variaciones en los datos o en el caso de una selección más amplia de indicadores. Para ello, en futuros estudios se podría considerar un grid search más extenso o la incorporación de técnicas bayesianas buscando un equilibrio entre complejidad y generalización.
- **Rigidez frente a cambios en los estándares del CACES:** este estudio se limita a los 42 indicadores propuestos por el CACES y vigentes para el año 2025; en caso de modificaciones en la normativa se ha de considerar ponderaciones, reglas de decisión y umbrales específicos para nuevos estudios, debido a que existe una limitada capacidad de una aplicación longitudinal del modelo sin un previo re-entrenamiento y adaptación de los parámetros a la nueva base de datos.

5.6 Trabajos a futuro

- **Validación externa y longitudinal:** Se recomienda incorporar datos de otras instituciones y de un rango más amplio de años; al aumentar este estudio se podría medir el modelo a largo plazo y en diferentes contextos. Esto garantizaría la robustez del modelo y su valía para la toma de decisiones.
- **Adaptación dinámica a nuevos estándares:** El diseño de un pipeline automatizado que detecte cambios en la documentación oficial del CACES sería un componente muy valioso para incorporar a este estudio; el uso de

técnicas como web scraping permitiría la recolección de nuevos datos. El modelo debería re-entrenarse y recalibrarse variando hiperparámetros y reduciendo el esfuerzo manual ante las actualizaciones regulatorias.

- **Exploración de modelos híbridos:** Se recomienda experimentar con técnicas de ensemble más flexibles como el uso de técnicas de gradiente boosting o stacking, por ejemplo. Potencialmente estas técnicas podrían mejorar la predicción sin el sacrificio de la interpretabilidad del modelo.
- **Implementación de un sistema de monitoreo continuo:** Se recomienda el desarrollo de un dashboard que sea interactivo y que esté alimentado de datos a tiempo real y permita al departamento de evaluación ser conscientes a tiempo real de posibles desviaciones en el cumplimiento de estándares y las consecuencias de la aplicación de las recomendaciones planteadas en esta investigación.

Conclusiones

Mediante la implementación de un sistema de análisis que utiliza aprendizaje automático, se ha logrado diagnosticar, explicar y priorizar el cumplimiento de los requerimientos del CACES en el Instituto Superior Universitario Carlos Cisneros. El modelo global fijó límites categóricos claros, gracias al contexto proporcionado por el modelo granular y el Random Forest a la granularidad y su interpretabilidad, se logró el objetivo de brindar una herramienta integral que permitiera mejorar y monitorear el desempeño institucional.

- Para cumplir con el primer objetivo, se recolectaron y procesaron datos históricos y actuales. Se combinaron los totales de 2023-2024 con cuarenta y un indicadores de 2024. No se encontraron nulos y se validó la correcta tipificación float64 en los cuantitativos. Con esto, se elaboró una base de datos consistente y confiable que mantuvo la comparabilidad histórica y la validez del modelo a través de las variables derivadas “fullratio”.
- Para el segundo objetivo específico, se establecieron, ejecutaron y validaron modelos predictivos. El árbol de decisión global alcanzó una exactitud media del 97.5 % en validación estratificada ratificando así, la solidez de los cortes 0.009 / 0.012 / 0.017. El árbol granular registró una exactitud media del 95 % y señaló tres indicadores docentes Formación práctica, La remuneración y la evaluación son responsables de más del 80 % de los incumplimientos. El Random Forest validó la superioridad de los PESOS DE LA IES (37 %) y B PESOS SIN TP (10 %). Fortaleciendo la confianza en las normas numéricas y aumentando la transparencia del proceso.
- En relación con el tercer objetivo específico, los resultados analíticos dieron lugar a recomendaciones estratégicas precisas: la implementación de un semáforo operativo que active respuestas cuando los pesos reales desciendan por debajo de 0.009; el desarrollo de un paquete docente que integre remuneraciones competitivas, evaluación 360° y becas de posgrado; y la creación de un plan de monitoreo que incluya dashboards semanales, validaciones anuales y revisión de estándares ideales superiores a 0.017. Estas acciones convierten la evidencia en métricas de seguimiento y permiten intervenciones focalizadas que facilitan la mejora continua.

En conjunto, el estudio cumplió los objetivos planteados al demostrar que la combinación de análisis predictivo y visualización estratégica puede guiar decisiones de alto impacto, reforzar la dimensión docente e investigativa y posicionar al ISU Carlos Cisneros en una trayectoria sostenible de cumplimiento con los estándares del CACES.

Recomendaciones

- **Asignar recursos según semáforo:** todo indicador con PESOS DE LA IES inferior a 0.009 debe recibir refuerzo inmediato para alcanzar, en seis meses, el umbral verde (≥ 0.012).
- **Impulsar un paquete docente prioritario:** incrementar remuneraciones al menos 15 %. implantar rúbricas de evaluación 360° y financiar becas de posgrado para que el 65 % del claustro curse estudios avanzados en dieciocho meses.
- **Revisar estándares ideales:** proponer al comité técnico la reducción gradual de B PESOS SIN TP cuando supere 0.017, evitando metas inalcanzables que penalicen avances parciales.
- **Monitorear y validar de forma continua:** desplegar *dashboards* semanales con estado de semáforo, peso real-ideal y delta interanual; repetir la validación estratificada y calcular *permutation importance* cada año para recalibrar umbrales.
- **Consolidar la mejora en I+D e Innovación:** mantener el fondo de publicaciones y así aumentar de manera concisa la producción científica en el instituto.

- **Refinar el modelo predictivo en 2025:** incorporar nuevas observaciones. aplicar poda mínima al árbol granular y emplear técnicas interpretables (SHAP) para confirmar la vigencia de los indicadores críticos.

Estas acciones convertirán los hallazgos analíticos en decisiones tangibles. elevarán la dimensión docente e investigativa y situarán al Instituto en una senda sostenible de cumplimiento con los estándares del CACES.

Referencias Bibliográficas

- Álvarez Polo, D. (2023). *Detección de ataques de intrusión con algoritmos de machine learning [Tesis de pregrado, Universidad de los Andes]. Universidad de los Andes.*
- Arias González, J. L. (2021). *Guía para elaborar la operacionalización de variables. Espacio I+D: Innovación más Desarrollo, 10(28).*
<https://doi.org/10.31644/IMASD.28.2021.a02>
- Asamblea Constituyente del Ecuador. (2008). *Constitución de la República del Ecuador. Registro Oficial 449.*
- Asamblea Nacional del Ecuador. (2018). *Ley Orgánica de Educación Superior, LOES. Registro Oficial Suplemento 297.*
- Basgall, M. J. (2022). *Análisis y diseño de técnicas de preprocesamiento de instancias escalables para problemas no balanceados en Big Data. Aplicaciones en situaciones de emergencias humanitarias [Tesis doctoral, Universidad de Granada]. Universidad de Granada.*
- CACES. (2020). *Reglamento de Evaluación Externa con Fines de Acreditación. Quito: Consejo de Aseguramiento de la Calidad de la Educación Superior.*
- CACES. (2021). *Modelo de Evaluación Externa con fines de acreditación para Institutos Superiores Técnicos y Tecnológicos. Quito: CACES.*
- CACES. (2022). *Guía metodológica para la autoevaluación institucional. Consejo de Aseguramiento de la Calidad de la Educación Superior.*
- CACES. (2023). *Modelo de Evaluación Externa con Fines de Acreditación para el Aseguramiento de la Calidad (Universidades y Escuelas Politécnicas). Quito: CACES.*
- Comisión Nacional de los Mercados y la Competencia. (2020). *Guía sobre cuantificación de daños por infracciones del derecho de la competencia. CNMC (España).*

- Díaz Martínez, M. A., Ahumada-Cervantes, M. de los Á., & Melo-Morín, J. P. (2021). Árboles de decisión como metodología para determinar el rendimiento académico en educación superior. *Revista Lasallista de Investigación*, 18(2), 94–104. <https://doi.org/10.22507/rli.v18n2a8>
- Díaz Martínez, N., López Cruz, P., & Ramírez Vega, A. (2021). Árboles de decisión: conceptos y aplicaciones en ciencia de datos. *Revista de Estadística Aplicada*, 9(2), 45-60.
- Fayyad, U. M., & Irani, K. B. (1992). On the handling of continuous-valued attributes in decision tree generation. *Machine Learning*, 8(1), 87-102.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery: An overview. *AI Magazine*, 17(3), 37–54.
- Flores Satalaya, J. M. (2024). El machine learning para abordar el abandono escolar: una revisión de los modelos más innovadores. *Ciencia Latina*, 8(6), 10992–11012. https://doi.org/10.37811/cl_rcm.v8i6.15824
- Gavilanes Simbaña, A., & Rojas Cevallos, A. (2023). Predicción del rendimiento académico... segundo semestre... Estadística – UCE, 2023 [Tesis de grado, Universidad Central del Ecuador]. Repositorio UCEDspace.uce.edu.ec.
- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (2.ª ed.). O'Reilly Media.
- Gonzales León, A. C. (2020). Perfilamiento y limpieza para optimizar la gestión de datos la plataforma web de una empresa de ventas, Lima 2020 [Tesis de grado, Universidad Privada Norbert Wiener]. Repositorio U. Wiener.
- ISU Carlos Cisneros. (2024). Informe anual de gestión institucional y aseguramiento de calidad. Riobamba: ISU Carlos Cisneros.

- Jiménez Galán, N. A. (2024). *Desarrollo de un modelo predictivo basado en minería de datos para anticipar el crecimiento de la cartera de clientes en una empresa proveedora de servicios de firma electrónica [Trabajo de titulación de ingeniería]. Universidad Tecnológica ECOTEC.*
- Luna, D., & Marín, R. (2018). *Diseño de evaluación de indicadores de calidad (KPIs) para una ruta de mantenimiento basada en árboles de decisión. En Actas de la XVIII Conferencia de la AEPIA – CAEPIA 2018.*
- Marín-Rodríguez, W., et al. (2023). *Modelo de clasificación para la deserción estudiantil en una universidad pública del Perú. EAI Endorsed Transactions on Scalable Information Systems, 10(5), e4.*
- Martínez-Pérez, J. A., & Pérez-Martín, S. (2023). *La Curva ROC. Medicina de Familia. SEMERGEN, 49(1), 1–3. <https://doi.org/10.1016/j.semerg.2022.101821>*
- Mendoza, J., & Castillo, R. (2023). *Análisis comparativo entre los estándares de calidad del Ministerio de Educación y del CACES [Trabajo de titulación]. Universidad X.*
- Organización de las Naciones Unidas para la Alimentación y la Agricultura. (1996). *Las bases y la necesidad de indicadores. FAO.*
- Pérez Peña, F., Cobaisse Ibáñez, M., Villagrán Pradena, S., & Alvarado, R. (2023). *Aspectos generales del uso de métodos mixtos para investigación en salud. Medwave, 23(10), e2767. <https://doi.org/10.5867/medwave.2023.10.2767>*
- Rivera Díaz, G. E. (2024). *Elaboración de un modelo de susceptibilidad frente a deslizamientos en el área urbana de Guayaquil [Proyecto integrador, ESPOL]. ESPOL.*
- Rodríguez González, J., & Ugalde Saborío, E. (2021). *Impacto de la estandarización y escalado: factor para predicción de costos en proyectos a través de una red neuronal*

- artificial. Ingeniare. Revista chilena de ingeniería*, 29(2), 265–274.
<https://doi.org/10.4067/S0718-33052021000200265>
- Romero-Duque, G. A., González-Prieto, C. A., Díaz-Barriosnuevos, M. A., & Rueda-Menjura, N. A. (2023). *Revisión y perspectivas para la construcción de bases de datos robustas con datos faltantes: caso aplicado a información financiera. Tecnura*, 27(75), 14–37. <https://doi.org/10.14483/22487638.18628>
- Ruiz Martínez, W. (2020). *El aprendizaje automático en la educación superior como herramienta para optimizar los índices de deserción y detectar los factores que la ocasionan. Hashtag Revista Digital*(16), 8–16.
- Shalev-Shwartz, S., & Ben-David, S. (2020). *Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press.*
- UNESCO. (2019). *Global Convention on the Recognition of Qualifications concerning Higher Education. Paris: UNESCO Publishing.*
- UNESCO. (2021). *Reimagining our futures together: A new social contract for education. Paris: UNESCO.*
- UNESCO-IESALC. (2023, 29 noviembre). *A survey will analyse global trends in tertiary education quality assurance. UNESCO International Institute for Higher Education in LAC. iesalc-unesco.org*
- Uriz Martín, M. (2021). *Mejora de los algoritmos de minería de datos: combinación de clasificadores, preprocesamiento y sus aplicaciones [Tesis doctoral, Universidad Pública de Navarra]. Universidad Pública de Navarra.*
- Vogel, S., et al. (2023). *Scikit-learn 1.4: New features and enhancements. Journal of Machine Learning Software*, 2(1), 1–9.
- Yates, D., Islam, Z., & Gao, J. (2019). *SPAARC: A fast decision tree algorithm. En R. Islam, Y. S. Koh, Y. Zhao, G. Warwick, D. Stirling, C.-T. Li, & Z. Islam (Eds.), Data Mining:*

16th Australasian Conference, AusDM 2018, Revised Selected Papers
(Communications in Computer and Information Science, Vol. 996, pp. 43-55).
Springer. https://doi.org/10.1007/978-981-13-6661-1_4

Zapeta Hernández, A., Galindo Rosales, G. A., Juan Santiago, H. J., & Martínez Lee, M.
(2022). Métricas de rendimiento para evaluar el aprendizaje automático en la
clasificación de imágenes petroleras utilizando redes neuronales convolucionales.
Ciencia Latina Revista Científica Multidisciplinar, 6(5), 4624–4638.
https://doi.org/10.37811/cl_rcm.v6i5.3420

Anexos

Anexo A: Capturas del Código realizado

```
import pandas as pd          # Lectura/escritura de datos y manejo de DataFrames
import numpy as np          # Operaciones numéricas y álgebra lineal
import matplotlib.pyplot as plt # Creación y personalización de gráficos 2D
import seaborn as sns       # Visualizaciones estadísticas avanzadas (barras, calor, etc.)

# ----- Modelado, validación y evaluación -----
from sklearn.model_selection import (
    train_test_split, # División entrenamiento-prueba
    KFold,            # Validación cruzada K-Fold simple
    StratifiedKFold, # K-Fold estratificado (mantiene proporción de clases)
    cross_val_score, # Métrica única en CV
    cross_validate   # Múltiples métricas en CV
)

from sklearn.tree import DecisionTreeClassifier, plot_tree
# DecisionTreeClassifier: construcción de árboles de decisión
# plot_tree: visualización de la estructura del árbol

from sklearn.ensemble import RandomForestClassifier
# RandomForestClassifier: ensamble de árboles para mayor robustez

from sklearn.metrics import (
    accuracy_score, # Exactitud
    precision_score, # Precisión
    recall_score,   # Sensibilidad
    f1_score,       # F1-score
    roc_auc_score,  # Área bajo la curva ROC
    classification_report # Resumen de métricas por clase
)
```

```
# Cargar el archivo Excel llamado "Resultados del estudio.xlsx"
df_weights = pd.read_excel("/content/Informe_resumen.xlsx")

# Mostrar información básica y primeras filas
df_weights.info()

df_weights.head(48)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 43 entries, 0 to 42
Data columns (total 3 columns):
 #   Column          Non-Null Count  Dtype
---  ---
 0   INDICADORES     43 non-null     object
 1   B PESOS SIN TP  43 non-null     float64
 2   PESOS DE LA IES 43 non-null     float64
dtypes: float64(2), object(1)
memory usage: 1.1+ KB
```

```

# Se imprimen detalles de la estructura y las primeras filas del DataFrame para verificar su contenido.
print("Tabla original:")
print(df_weights.info())
print(df_weights.head())

# Se crea una nueva columna "cumplimiento_ratio" que divide el peso real (PESOS DE LA IES) por el peso ideal (B PESOS SIN TP).
# Esto mide qué porcentaje de cumplimiento alcanza cada indicador.
df_weights["cumplimiento_ratio"] = df_weights["PESOS DE LA IES"] / df_weights["B PESOS SIN TP"]

# Se define "cumple" como 1 si el ratio es >= 0.60, y 0 en caso contrario. Este umbral determina quién cumple o no.
threshold = 0.60
df_weights["cumple"] = (df_weights["cumplimiento_ratio"] >= threshold).astype(int)

# Se muestra un ejemplo de cómo quedan los datos transformados, incluyendo la nueva columna "cumplimiento_ratio" y la etiqueta "cumple".
print("\nEjemplo de indicadores con su ratio y target:")
print(df_weights[["INDICADORES", "B PESOS SIN TP", "PESOS DE LA IES", "cumplimiento_ratio", "cumple"]].head(49))

```

```

# Se seleccionan las características (B PESOS SIN TP, PESOS DE LA IES) como X y la variable objetivo "cumple" como y.
X = df_weights[["B PESOS SIN TP", "PESOS DE LA IES"]]
y = df_weights["cumple"]

# Se divide el conjunto de datos en entrenamiento y prueba (80%-20%), fijando la semilla random para replicar resultados.
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

# Se crea y entrena un clasificador de árbol de decisión con max_depth=3 para evitar que crezca en exceso.

clf = DecisionTreeClassifier(random_state=42, max_depth=3)
clf.fit(X_train, y_train)

# Se evalúa el modelo en el conjunto de prueba: se imprime la accuracy y el reporte de clasificación.

y_pred = clf.predict(X_test)
print("\nResultados del modelo:")
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Reporte de clasificación:")
print(classification_report(y_test, y_pred))

# Se visualiza el árbol entrenado, mostrando las variables, las clases (No Cumple y Cumple) y redondeando las cajas.

plt.figure(figsize=(10,6))
plot_tree(clf, filled=True, feature_names=["B PESOS SIN TP", "PESOS DE LA IES"],
          class_names=["No Cumple", "Cumple"], rounded=True)
plt.title("Árbol de Decisión: Cumplimiento de Indicadores")
plt.show()

```

```

# SELECCIÓN DE CARACTERÍSTICAS:
# Se usan las dos columnas numéricas "B PESOS SIN TP" y "PESOS DE LA IES" para predecir el target.
X = df_weights[["B PESOS SIN TP", "PESOS DE LA IES"]]
y = df_weights["cumple"]

# CREACIÓN DEL MODELO DE ÁRBOL DE DECISIÓN:
# Se crea un DecisionTreeClassifier con una profundidad máxima de 3 (max_depth=3) para evitar
# que el árbol se sobre ajuste a datos pequeños.
clf = DecisionTreeClassifier(random_state=42, max_depth=3)

# APLICACIÓN DE VALIDACIÓN CRUZADA (CV):
# Se utiliza KFold con 5 folds, haciendo un shuffle de los datos y fijando una semilla para reproducibilidad.
cv = KFold(n_splits=5, shuffle=True, random_state=42)
# Se evalúa la exactitud del modelo en cada pliegue utilizando cross_val_score y se guarda cada resultado en 'scores'.
scores = cross_val_score(clf, X, y, cv=cv, scoring='accuracy')

# Impresión de los resultados de la validación cruzada: la accuracy en cada fold y el promedio global.
print("Resultados de la Validación Cruzada (Accuracy por Fold):")
print(scores)
print("Promedio de Accuracy con CV:", np.mean(scores))

```

Features y variable objetivo

```

X = df_weights[["B PESOS SIN TP", "PESOS DE LA IES"]]
y = df_weights["cumple"]

```

Configuración del modelo y validación cruzada estratificada

```

clf = DecisionTreeClassifier(max_depth=3, random_state=42)

```

```

# Estratificación para que cada fold contenga la misma proporción de
# clases "Cumple" y "No Cumple".

```

```

skf = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)

```

Métricas que queremos evaluar en cada pliegue

```

scoring = {
    'accuracy': 'accuracy',
    'precision': 'precision',
    'recall': 'recall',
    'f1': 'f1',
    'roc_auc': 'roc_auc'
}

```

Ejecución de la validación cruzada y reporte de resultados

```

cv_results = cross_validate(
    clf,
    X, y,
    cv=skf,
    scoring=scoring,
    return_train_score=False
)

```

Imprimir las métricas por pliegue y el promedio final

```

for metric in scoring.keys():
    print(f"{metric.capitalize()} por fold: {cv_results[f'test_{metric}']}")
    print(f"Promedio {metric}: {np.mean(cv_results[f'test_{metric}']):.3f}\n")

```

```

# Entrenamiento final con todos los datos (tras la validación cruzada)
clf_final = DecisionTreeClassifier(random_state=42, max_depth=3)
clf_final.fit(X, y)

plt.figure(figsize=(12, 7))
plot_tree(clf_final, filled=True, feature_names=X.columns, class_names=["No Cumple", "Cumple"], rounded=True)
plt.title("Árbol Final Entrenado con Todos los Datos")
plt.show()

#Transformación a Nivel Granular

# Se realiza one-hot encoding en la columna "INDICADORES" para crear variables dummy (0/1) por cada indicador.
df_model = df_weights.copy()
df_model_encoded = pd.get_dummies(df_model, columns=["INDICADORES"])

# Se escogen las características (features) descartando 'cumplimiento_ratio' y 'cumple',
# que no deben ser usadas como predictoras en este enfoque.
X = df_model_encoded.drop(columns=["cumplimiento_ratio", "cumple"])
y = df_model_encoded["cumple"]

print(f"\nShape de X_model: {X.shape}")
print(f"Columnas generadas: {list(X.columns)[:10]} ... (mostrando solo las primeras 10)")

# 3. Entrenar el Árbol de Decisión Granular

dt_classifier = DecisionTreeClassifier(random_state=42, max_depth=5)
dt_classifier.fit(X, y)

# Se predice sobre el mismo conjunto
y_pred = dt_classifier.predict(X)

# Métricas de desempeño en el mismo set de entrenamiento
accuracy = accuracy_score(y, y_pred)
report = classification_report(y, y_pred)
print(f"\nAccuracy del árbol (entrenado sobre todos los datos): {accuracy:.3f}")
print("Reporte de clasificación:")
print(report)

```

```

# Visualizar el Árbol de Decisión
# Se genera una figura donde se dibuja el árbol y se muestran las reglas y el resultado de cada nodo.
plt.figure(figsize=(20,10))
plot_tree(
    dt_classifier,
    filled=True,
    feature_names=X.columns,
    class_names=["No Cumple", "Cumple"],
    rounded=True,
    fontsize=8
)
plt.title("Árbol de Decisión - Clasificación de Indicadores (Enfoque Granular)")
plt.show()

```

```

# Separamos las características (X) y la columna objetivo (y).
# Eliminamos "cumplimiento_ratio" y "cumple" para no incluirlas como predictores.
X = df_model_encoded.drop(columns=["cumplimiento_ratio", "cumple"])
y = df_model_encoded["cumple"]

# Creamos el modelo de árbol de decisión; max_depth se puede ajustar según necesitemos.
clf = DecisionTreeClassifier(random_state=42, max_depth=5)

# Configuramos la validación cruzada con 5 divisiones (folds).
cv = KFold(n_splits=5, shuffle=True, random_state=42)

# Usamos 'accuracy' como métrica de desempeño; se pueden usar otras si es necesario.
scores = cross_val_score(clf, X, y, cv=cv, scoring='accuracy')

# Mostramos los resultados de cada fold y el promedio.
print("Exactitud en cada fold:", scores)
print("Promedio de exactitud:", np.mean(scores))

```

```
# Features (X) y etiqueta (y)

X = df_model_encoded.drop(columns=["cumplimiento_ratio", "cumple"])
y = df_model_encoded["cumple"]

# Modelo y CV estratificada

clf = DecisionTreeClassifier(max_depth=5, random_state=42)
skf = StratifiedKFold(n_splits=5, shuffle=True, random_state=42)

scoring = {
    'accuracy': 'accuracy',
    'precision': 'precision',
    'recall': 'recall',
    'f1': 'f1'
}

cv_results = cross_validate(clf, X, y, cv=skf, scoring=scoring,
                            return_train_score=False)

# Mostrar métricas

for metric in scoring.keys():
    scores = cv_results[f'test_{metric}']
    print(f"{metric.capitalize()} por fold : {scores}")
    print(f"Promedio {metric} : {np.mean(scores):.3f}\n")
```

```

# Preparar los datos de entrenamiento

X = df_model_encoded.drop(columns=["cumplimiento_ratio", "cumple"], errors='ignore')
y = df_model_encoded["cumple"]

# 2. Entrenar el modelo (Random Forest)

# Creamos un RandomForestClassifier para estimar la importancia de variables.
rf_model = RandomForestClassifier(random_state=42, n_estimators=50, max_depth=5)
rf_model.fit(X, y)

# Evaluar el desempeño en todo el conjunto
y_pred = rf_model.predict(X)
print("Exactitud (accuracy) en el conjunto completo:", accuracy_score(y, y_pred))
print("Reporte de clasificación:")
print(classification_report(y, y_pred))

# Extraer e imprimir la importancia de las variables

importances = rf_model.feature_importances_
feature_names = X.columns

# Crear un DataFrame para ordenar y mostrar las importancias
df_importances = pd.DataFrame({
    "Feature": feature_names,
    "Importance": importances
}).sort_values(by="Importance", ascending=False)

print("\nImportancia de variables (ordenadas):")
print(df_importances.head(20)) # Se muestran las 20 más relevantes

```

```

# Graficar las variables con mayor importancia

plt.figure(figsize=(12,8))
plt.barh(df_importances["Feature"].head(15),
         df_importances["Importance"].head(15),
         color="teal")
plt.gca().invert_yaxis() # Invierte el eje Y para mostrar la más importante arriba
plt.title("Top 15 - Importancia de Variables (Random Forest)")
plt.xlabel("Importancia Relativa")
plt.ylabel("Variable")
plt.tight_layout()
plt.show()

```

```

# Cargar el archivo Excel llamado "Resultados del estudio.xlsx"
df_criterio = pd.read_excel("/content/Informe_criterio.xlsx")

# Mostrar información básica y primeras filas
df_criterio.info()

df_criterio.head(87)

```

```

# Valoración del instituto 2024 (%) a float en la misma escala [0, 1].

def convert_valoracion_2024(x):
    if isinstance(x, str):
        x = x.strip()
        if x.endswith("%"):
            return float(x.rstrip("%").strip()) / 100
        else:
            return float(x)
    else:
        return x

df_criterio["Valoración del instituto 2024 (%)"] = df_criterio["Valoración del instituto 2024 (%)"].apply(convert_valoracion_2024)

# Calcular la diferencia (delta) entre la valoración 2024 y la de 2023 para cada criterio.
df_criterio['Delta'] = df_criterio["Valoración del instituto 2024 (%)"] - df_criterio["Valoración del instituto 2023 (%)"]

# Mostrar la tabla comparativa y visualizar con un gráfico de barras.
print("Tabla Comparativa de Valoraciones (2023 vs 2024) por Criterio:")
print(df_criterio)

plt.figure(figsize=(10,6))
sns.barplot(x='Criterio', y='Delta', data=df_criterio, palette="viridis")
plt.axhline(0, color='gray', linestyle='--')
plt.title("Variación (Delta) en la Valoración del Instituto: 2024 - 2023")
plt.ylabel("Delta (2024 - 2023)")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

# Entrenar un árbol de decisión para clasificar los criterios en "Cumple" o "No Cumple"
# Definamos que un criterio 'cumple' si su valoración en 2024 es mayor o igual a un umbral
umbral = 0.10
df_criterio['cumple'] = (df_criterio["Valoración del instituto 2024 (%)"] >= umbral).astype(int)

# Para el modelo, usaremos dos variables: "Peso modelo (%)" y "Delta"
X = df_criterio[['Peso modelo (%)', 'Delta']]
y = df_criterio['cumple']

clf_criterio = DecisionTreeClassifier(random_state=42, max_depth=3)
clf_criterio.fit(X, y)

plt.figure(figsize=(10,6))
plot_tree(clf_criterio, filled=True, feature_names=X.columns, class_names=["No Cumple", "Cumple"], rounded=True)
plt.title("Árbol de Decisión por Criterio")
plt.tight_layout()
plt.show()

```

Anexo B: Enlace al código de la investigación en Google Colab

https://colab.research.google.com/drive/1rMRHJ7qI9Br4K8TxZBrR_F7pB-jXj00E?usp=sharing

Anexo C: Enlace al informe de autoevaluación del ISU Carlos Cisneros 2024

https://carloscisnerosrio-my.sharepoint.com/:f/g/personal/carolina_llanga_istcarloscisneros_edu_ec/EuS8gQ6G1VHjyUMFVm79MoBClhUiXMZCDu2wtyS474c7A?e=cBu23x

Anexo D: Enlace al informe de autoevaluación del ISU Carlos Cisneros 2023

https://carloscisnerosrio-my.sharepoint.com/:f/g/personal/carolina_llanga_istcarloscisneros_edu_ec/EuS8gQ6G1VHjyUMFVm79MoBClhUiXMZCDu2wtyS474c7A?e=cBu23x